The Cerebellum. 2005; 4: 1–9

Taylor & Francis
Taylor & Francis Group

REVIEW

Genomics, morphogenesis and biophysics: Triangulation of Purkinje cell development

MALCOLM J. SIMONS[1] & ANDRÁS J. PELLIONISZ[2]

[1]*Haplomic Technologies Pty. Ltd., Victoria, Australia, and* [2]*HelixoMetry, Sunnyvale, California, USA*

Abstract
The cerebellar Purkinje cells (P-cells) comprise an organelle that is suitable for combined analysis by morphology and genomics, using biophysical tools. In some unknown way, genomic information specifies the development of P-cells. One of us (AJP) has previously proposed that fractal processes associated with DNA are in a causal relation to the fractal properties of organelles such as P-cells (FractoGene, 2002, patent pending). This fractal postulate predicts that the dendritic arborization of P-cells will be less complex in lower order vertebrates. The prediction can be tested by systematic comparative neuroanatomy of the P-cell in species for which genome sequences permit inter-species comparison. The *Fugu rubripes* (Fugu), *Danio rerio* (Danio) and other species are lower order vertebrates for which genome sequences are available and tests could be conducted. Consistent with the fractal prediction, P-cell dendritic arbor is primitive in Fugu, being much less complex than in *Mus musculus* and in *Homo sapiens*. Genomic analysis readily identified PEP19/Pcp4, Calbindin-D28k, and GAD67 genes in Fugu and in Danio that are closely associated with P-cells in *Canis familiaris, Rattus norvegicus, Mus musculus* and *Homo sapiens*. Gene L7/Pcp2 exhibits strongest association with P-cells in higher vertebrates. L7/Pcp2 shows strong protein residue homology with genes greater than 600 residues and including 2–3 GoLoco domains, designated as having G protein signaling modulator function (AGS3-like proteins). Fugu has a short gene with a single GoLoco domain, but it has greatest homology with the AGS3-like proteins. No similar short gene is present in Danio or in Xenopus. Classical L7/Pcp2 is only detected in higher vertebrates, suggesting that it may be a marker of more recent evolutionary development of cerebellar P-cells. We expect that a new generation of data mining tools will be required to support recursive fractal geometrical, combinatorial, and neural network models of the genomic basis of morphogenesis.

Key words: *Biophysics, genomics, morphogenesis, cerebellum, Purkinje cell*

Introduction

In ways that are largely unknown, genomic information is presumed to specify the development of cellular systems, distributed as well as discrete. For most of the half century since the discovery of DNA structure it has been assumed that the projected 100,000+ genes were sufficiently responsible for the genesis of dispersed cell systems and of organelles and organs. It has been a major surprise to find that the number of genic units in humans is only of the order of 20,000, and possibly an even greater challenge to comprehend that a similar number is present in most eukaryotic species whose genomes have been sequenced. While recognizing that alternative splicing is exhibited by many genes, augmenting the number of protein varieties, it is becoming widely appreciated, even in generalist publications, that previous concepts of genomic information limited to gene-based dogma need to be reviewed (1,2). It is fast becoming recognized that there is genetics beyond genes in which the information content of introns and other non-coding sequences can be expected to contribute important roles. As recently as 2003, Gibbs (1) recorded the assertion by Mattick (3) that "the failure to recognize the importance of introns 'may well go down as one of the biggest mistakes in the history of molecular biology'". One of us (MJS) did not make this mistake. By 1989, 11 years before essential completion of the Human Genome sequencing project in 2000, patents were filed based on the discovery that intron/non-coding sequence variation was sufficiently non-random for haplotype analysis in unrelated subjects (4–6). Recognition of the utility of linkage disequilibrium-based haplotype structure was a paradigm shift from previous linkage-based pedigree analysis. The central issue was that if the so-called Junk DNA had structure then, under

Correspondence: Dr Malcolm Simons, Haplomic Technologies Pty. Ltd., PO Box 5201, Geelong North, Victoria, 3215, Australia. E-mail: drsimons@haplomics.com

(Received 2 May 2005; accepted 28 September 2005)

ISSN 1473-4222 print/ISSN 1473-4230 online
DOI: 10.1080/14734220500378581

Darwinian Theory, it could be expected to have function. Proof of application was provided for the most polymorphic and complex component of the genome, the major histocompatibility complex (#6p21.3), in the context of HLA genetic tissue typing (7,8). The non-random haplotypic structure of non-coding DNA has been the basis for genome-wide gene discovery by haplotype mapping for more than a decade, and underpins the current international HapMap project. Also in 1989, the second author (AJP), working in a totally different field, and unknown to MJS, recognized that attention would need to be given to the wider genome to understand genetically determined, fractal geometrical development of P-cells (9), suggesting that a key to fractal geometry of the P-cell may lie in the genome, while cautioning that an understanding may lie 'far in the future' (9).

With expansion of genomic horizons beyond genes into non-protein coding DNA, it is timely to consider conceptual frameworks that have the potential of accelerating entry into the post-Gene era. Conceptual frameworks are most useful if they offer the opportunity of mathematization as well as of description, both of the genomic informatics and of derivative cellular and organismal development.

Linguistic features of non-coding DNA were suggested in the 1990s, prompting the possibility of fractal-like characteristics in the genome (10–14). The claims appear to be methodology based, since other investigations reported a failure to detect ''well-defined scaling or fractal exponents'' (15) or any signs of hidden language in non-coding DNA (16).

In 1977, Mandelbrot conjectured that the P-cell might be modelled as a non-Euclidean fractal structure since P-cell arborization has the fractal-like features of scale free, self-similar patterns suggestive of generation by an iterative process (17). In 1989, prompted by Mandelbrot's musing (17, p. 162), one of us (AJP) developed a fractal model of the P-cell (9).

We recognize that a mathematical framework is required for holistic understanding of the contribution of both non-coding and coding genomic components to organelle and organ development. Presently there is a limited choice of conceptual frameworks that also provide a process and a platform. In 2002, one of us (AJP) postulated a causal relationship between fractal genetic structure and fractal P-cell arborization as illustrative of organelle morphogenic complexity (see 'FractoGene Fig. 6' from ref. 18, shown as Figure 3 in this paper). Thus, our preferred model for a systematic approach to the intrinsic mathematics of genome-directed P-cell arborization development is that of fractal geometry. The simplest notion is that structural DNA is itself fractal, sufficient to encode the information required for development of discrete and distributed cellular systems. Whether or not structural DNA is itself fractal, we propose that nucleic (DNA-based), epinucleic (beyond DNA), epistatic (interaction between two or more genes influencing a single phenotype) and other genetic and genomic mechanisms are considered as conferring fractal characteristics to morphogenic development.

We propose a role for triangulation, a term which has a specific meaning in geometry as the minimum information required to define the precise location of a point in a plane. Genomic contribution and morphological development are two such points. Application of biophysical tools to space-time domain provides the third component to relate the two points. We feel that such an interactive strategy is required to supplement if not supplant current reductionist paradigms that, in the extreme, assume causality between single genes and a phenotype.

Triangulation by mathematization of biology (biophysics) has already proven useful when membrane electrophysiology (19) was interrelated with Neuro-morphology by computer modelling of the P-cell (20,21). That triangulation successfully explained electrophysiological findings such as how the climbing fiber-evoked 'complex spikes' arose from the complex P-cell dendritic tree (20). The present task of relating Neuro-morphogenesis and genomics through biophysical concepts and methods is even more challenging. One approach is through the pathogenetic spinocerebellar ataxias (SCA1-25). The SCAs are a clinically heterogeneous group of disorders that were addressed in a recent special issue of *The Cerebellum* (2005), overviewed by Manto (22). Those SCA associated with CAG expansions, particularly Friedreich SCA1, have already been a focus of genetic attention (23), including special focus on the P-cell (24–26). Morphological studies coupled with biophysical modelling of the P-cell have been elegantly presented (21) as depicted in Figure 1, left side.

Of all neurons, P-cells exhibit the most complex dendritic arborization and a most remarkable diversity, yet conserving a pattern of arborization such that morphologists recognize this cell virtually independent of the studied species. The visibly geometrical branching pattern, which is curiously confined to a wafer-thin layer, is strikingly similar across different species, yet distinct in its complexity. The greatest dendritic complexity was first observed in humans almost a Century ago (27), see right side of Figure 1). Subsequent studies have revealed that the P-cell in frog, mouse, guinea pig (etc) has a less complex pattern than in the human (see frog P-cell histology (28), frog P-cell modelling (29), P-cell histology for the guinea pig (30) and guinea pig P-cell modelling (21), see left side of Figure 1.). The well-known but hitherto unexplained exception from an apparent trend is the Mormiryd (an electric fish)

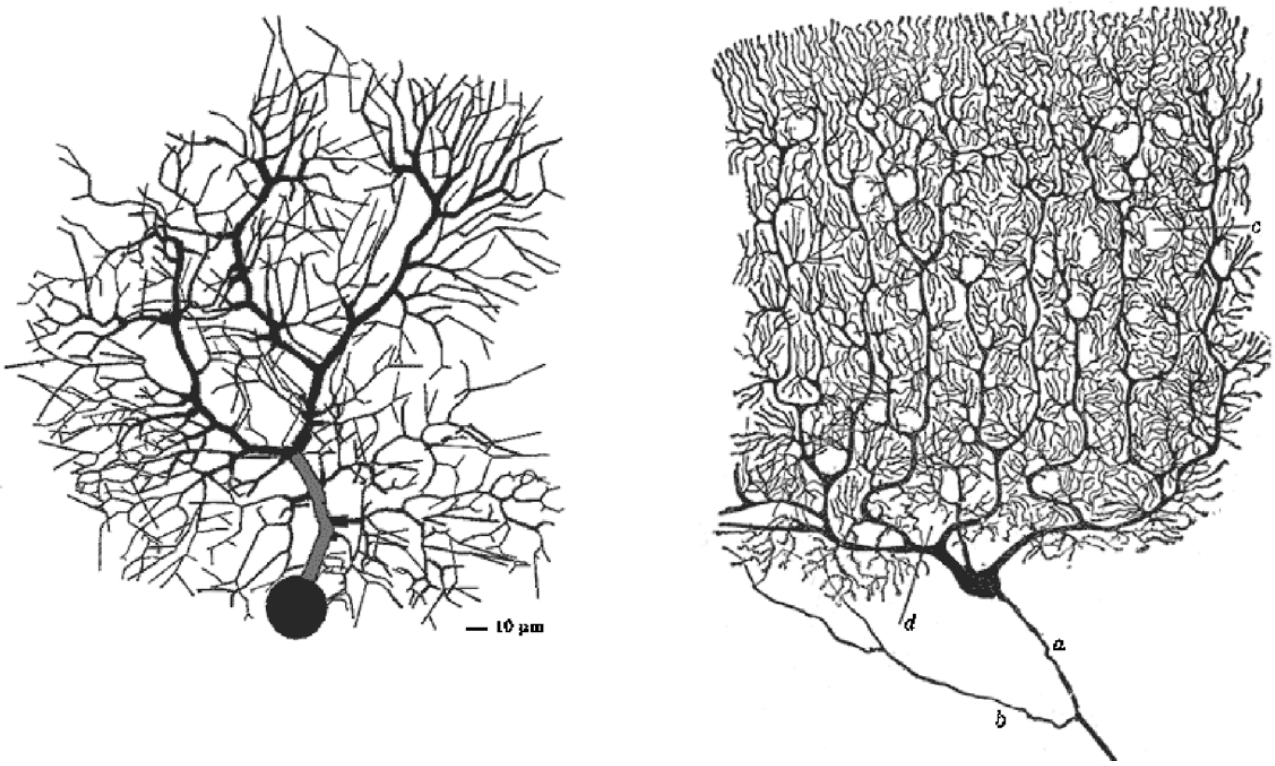

Figure 1. The cerebellar P-cell in the guinea pig (left (21) and on the right, in the human (27)). While the diagram on the left is a modern computer reconstruction of fluorescent-stained P-cell and thus exhibits all details, the classical diagram (27) on the right is a drawing based on Golgi silver stain preparation, and thus the actual P-cell dendritic arbour could be even more complex than shown.

in which the 'Gigantocerebellum' exhibits a giant P-cell of astounding refinement (31).

The genomes of many species have been sequenced. They provide a new opportunity for combining comparative anatomy and genomics, guided by mathematical conceptual frameworks. Hitherto, neurobiological and biophysical studies of P-cells and networks revealed an understanding of the space-time sensorimotor coordination by the cerebellar neural network (32). As yet, the genomic aspects of P-cell development have not been brought into similar focus. Recently, the genomes of the poisonous sea-water puffer fish *Takifugu rubripes* (hereinafter Fugu) and the non-poisonous fresh-water puffer fish *Tetraodon nigriviridis* (fam. Tetradontidae) have been revealed to be much smaller than higher vertebrates, comprising some 0.365 and 0.385 Gigabases, respectively. They are in stark contrast to the size of the human genome at 3.1 sequenced Gigabases, and even smaller than the 2.6 Gigabases of the mouse genome. The genome of another lower order fish, *Danio rerio* (common name – zebrafish, hereinafter Danio), has also been sequenced as smaller than that of human, at 1.56 Gigabases (all data are from the University of California at Santa Cruz Database). The availability of the genome sequences of Fugu (and of Danio and other species) promises to offer insights into genomic mechanisms underlying organelle development.

Based on the prediction of 'FractoGene, Fig.6' (18) (reproduced in Figure 3. of this paper) that Fugu should exhibit a primitive P-cell arbor, while Danio should show an intermediate complexity, one of us suggested (AJP to Prof. G. Székely on 4 July 2003) that a preliminary study be conducted to reveal the facts of neuroanatomy in the Fugu, and more recently in the Danio. While it is simplistic o imagine that total genome size would correlate linearly with specific organelle complexity, we were curious to examine P-cell presence and dendritic complexity in Fugu and Danio. That study is to be reported separately (33). The authors have permitted us to refer to their results as 'unpublished observations' for the Fugu. The Danio results (only partially available) and other comparisons are expected to blossom as a new branch of research.

The P-cell is a specific morphological development platform to be put into the 'crossfire of triangulation' similar to earlier syntheses by biophysics of Neuro-morphology and electrophysiology. In particular, we wish to model P-cell morphogenesis based on the genetic control that we assume is inherent in genomic information. Here we begin investigation of the P-cell using the three dimensions of morphological analysis, genomics and biophysics.

Genomics

We have searched the public genome databases for Fugu and (in anticipation of completed morphological studies also for Danio) for molecules that are candidate markers of cerebellar P-cells in that they are expressed with varying specificity in P-cells.

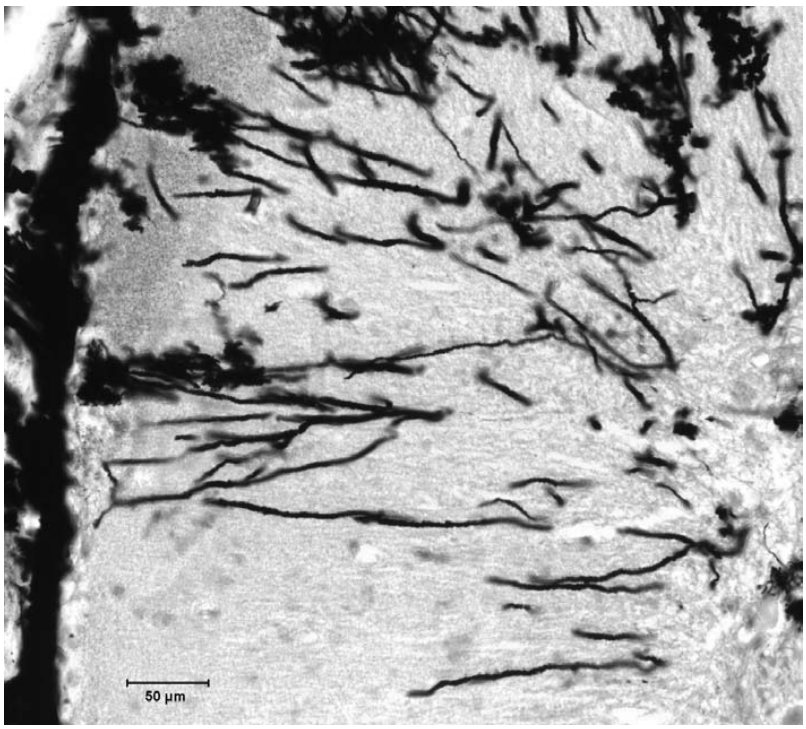

Figure 2. Golgi-stained cerebellar P-cell in *Fugu rubripens* (Courtesy of Székely et al. (33)). Note the primitive template of 'arbor'. Calibration bar in the upper left corner shows 50 micrometers.

Seventeen years ago the earliest report of what became known as L7/Pcp2 was found to be one of three cloned genes associated with P-cell degeneration in the recessive mouse mutant pcd (PCD5) (34).

The L7/Pcp2 marker has been detected only in P-cells and in retinal bipolar neurons. It appears to be the most specific marker of P-cells. L7/Pcp2 is a GoLoco domain protein that functions as a cell-type specific modulator of Galpha(i) and Galpha(o) in G-protein-mediated cell signaling. L7/Pcp2 mRNA transcripts have been reported as localizing within the proximal and distal branches of dendrites and in the proximal part of P-cell axons. The strong association of L7/Pcp2 protein and mRNA with P-cells has led to the suggestion of ''likely importance in controlling the development and/or motor control function of Purkinje cells'' (35).

The second clone identified 17 years ago was Calbindin-D28k (38). While this molecule is present in neuronal structures outside the cerebellum, among cerebellar elements it is a specific marker of P-cells. Homology BLAST search (NCBI, NIH) of Fugu and Danio protein databases readily identified Calbindin sequences for both species (Fugu – CAF99552, CAF96430, CAG00001; Danio – NP_957012, NP_957005, NP_001005776, AAH59479, AAH59467, AAH83168).

The third of the three genes to be reported was PCD6, an unidentified protein of greater than 500aa. 14 years later (36), the gene mutated in the original pcd mouse was identified as Nna1 (Fernandez-Gonzalez 2002). The Nna1 protein is 1,160 aa in mice, and has strong homology to proteins present in Fugu (CAG08401 – 997aa; CAF98247 – 538aa; CAG00089 – 774aa), Danio (NP_001004113, AAH80248 – 885aa), and Xenopus (AAH77808 – 1225aa; AAH77561 – 678aa).

Other molecules that are expressed in P-cells include PEP19/Pcp4 and glutamic acid decarboxylase 67 (GAD67).

The PEP19/Pcp4 gene product of 62 aa is one of the smallest translation products identified in eukaryotes. Using the human sequence as reference (NP_006189), an unnamed protein of 63 aa is readily identified in *Tetraodon nigriviridis* (CAF93207) as full length homology. The NCBI, NIH BLAST 'select organism' lists *Takifugu rubripes*, but we have not observed any listing for this organism in the BLAST output. Further, when the *T. nigiviridis* sequence (CAF93207) is BLASTed against Takifugu ribripes as the selected single organism, no hits were observed. Using the same human sequence as query, an unnamed 72 aa protein (CAE17624) is identified in Danio, with homology of less than full length, extending from aa 22 to the 3' end. Among lower vertebrates, the human gene sequence identified only Fugu and Danio. No sequences were revealed in Xenopus or Gallus.

GAD67 is present in both the Fugu and Danio databases in many entries (Fugu: CAF97647, CAG07380, CAG09942, CAG05466; Danio: AAH47851, AAC24327, AAD22710, NP_919400).

In view of the ease in which the P-cell markers PEP19/Pcp4, Calbindin-D28k, GAD67, and Nna1

homologues are detectable in Fugu and Danio databases, we were surprised that the most P-cell specific marker, L7/Pcp2, does not appear to be present in Danio, and is not readily recognizable in Fugu.

At a nucleotide level, BLASTing the promoter region of L7/Pcp2 and exon 2 reveals a homology with sequences restricted to vertebrate species (mouse, rat, dog and cow). By contrast, human exon1 and intron1 nucleotide sequences show no homology with any other species.

BLASTing with a range of vertebrate L7/Pcp2 protein sequences regularly identified all human, murine, rat, canine and bovine homologues with strong identity (Homo: XP_058956, Q8IVA1, AAN52488, AAN52487, AAH38715, XP_058956; Mus: AAH14694, NP_032816, P12660, AAH24853, AAH28982, AAN52485, AAA02989, BAC25015, B34955, AAB19316; Rattus: XP_221787; Canis: XP_542114; Bos: XP_600613).

In humans, L7/Pcp2 exists in two isoforms of 99 and 136 residues as a single GoLoco domain protein that modulates the activation of guanine nucleotide-binding regulatory protein (G protein) subunits Galpha (i) and Galpha(o) in cell signaling. It was therefore of considerable interest that, lower on the L7/Pcp2 list of revealed peptide sequences, were those designated as having G protein signalling modulator function (activator of G-protein signalling 3 [AGS3]-like proteins). These latter proteins, mostly greater than 600 aa in length, included 2–3 GoLoco sequences with which the L7/Pcp2 GoLoco region aligned. It was with this group that Fugu (CAG12405, CAF91737, CAG07707, CAG11554) and Danio (AAH83520, NP_001005936, NP_001007779, AAU14175, AAH54918, NP_956732) sequences were grouped. We do not know whether the third clone (PCD6) (36) to encode a protein greater than 500aa was an AGS3-like protein.

There was one exception to Fugu protein molecules of long length that was identified by vertebrate L7/Pcp2 BLAST query. CAF91737 accession sequence is a 78 residue peptide for which the most homologous sequences were G-protein signalling modulators 1,2 and 3 (activator of G-protein signalling 3-like). Lower on the listing of homologous sequences were the L7/Pcp2 sequences of the vertebrates. The domain showing greatest homology (Homo Q8IVA1 aa 32–53) overlapped with the GoLoc domain (aa 28–42, of 99). However, the corresponding residue positions in the Fugu peptide were at the 3' end of the peptide, at positions 55–76 of the 78 residues. When this domain in Fugu CAF91737 was used for BLAST query, the most similar sequences were Pcp2 peptides of the vertebrates. In view of the 3' end position of the GoLoco domain in the Fugu protein, it is notable that the domain is similarly positioned in bovine Pcp2 (aa72–93 of 100), different from all the other Pcp2 molecules. Nonetheless, the primary homology appears to be with AGS3-like molecules, rather than with L7/Pcp2 itself.

The apparent absence of observable cerebellar abnormalities in L7/Pcp2-null mice does not support a necessary role in P-cell development (37,38). The genomic database finding that Fugu and Danio seem to lack the prototypic L7/Pcp2 molecule is in keeping with the non-essentiality of the gene in P-cell development. There is evidence of close interaction between L7/Pcp2 and Galpha(o). Colocalization in P-cells suggests a functional role in regions of synaptic activity. It is therefore intriguing that, despite lacking a prototypic L7/Pcp2 molecule, both Fugu and Danio do have molecules exhibiting the GoLoco domains presumably involved in G-protein signalling modulation.

Discussion

The present study is a first step towards a triangulation by mathematization of the genomic contribution to morphological development guided by biophysical tools. Towards this goal we chose the P-cell as a platform.

The FractoGene postulate of a causal relationship between fractal genetic function and P-cell arborization that prompted the present study was detailed in 2002 (18).

'FractoGene', which is Figure 3 of this paper, is an actual copy of Figure 6 of the Patent Application (18). The legend to the original figure in the Patent stated: ''*Elements* of Prior Art are combined here, to make a complete circle between the demonstrably fractal-like neuron of Purkinje cell and the demonstrably fractal-like sequences of (only illustrative, not necessarily actual Purkinje cell) DNA base-pairs. From this demonstration diagram it is obvious that the four A,C,T,G base-pair segments are not identical, only self-similar''.

There were two cited *elements*:

(1) The fractal P-cell model, lifted from the publication (9), and
(2) An illustrative DNA sequence as one example from the wide array of 'repetitive DNA sequences' in the literature (http://bssv01.lancs.ac.uk/ADS/BIOS336/336L2.html).

The 'first element' (fractal P-cell model – top row of Figure 3. of this paper) shows, in four stages, the increasing complexity of the dendritic tree from B,C,D, to E, representing the fractal steps towards development of the Golgi-stained Guinea Pig P-cell (F) (30). As described in the original publication, the model (9) utilized an L-string replacement algorithm. It can be seen that the stem (corresponding to stage A that is not shown) and the Y-shaped fractal template remains conserved through the four stages. By contrast, all peripheral twigs are replaced

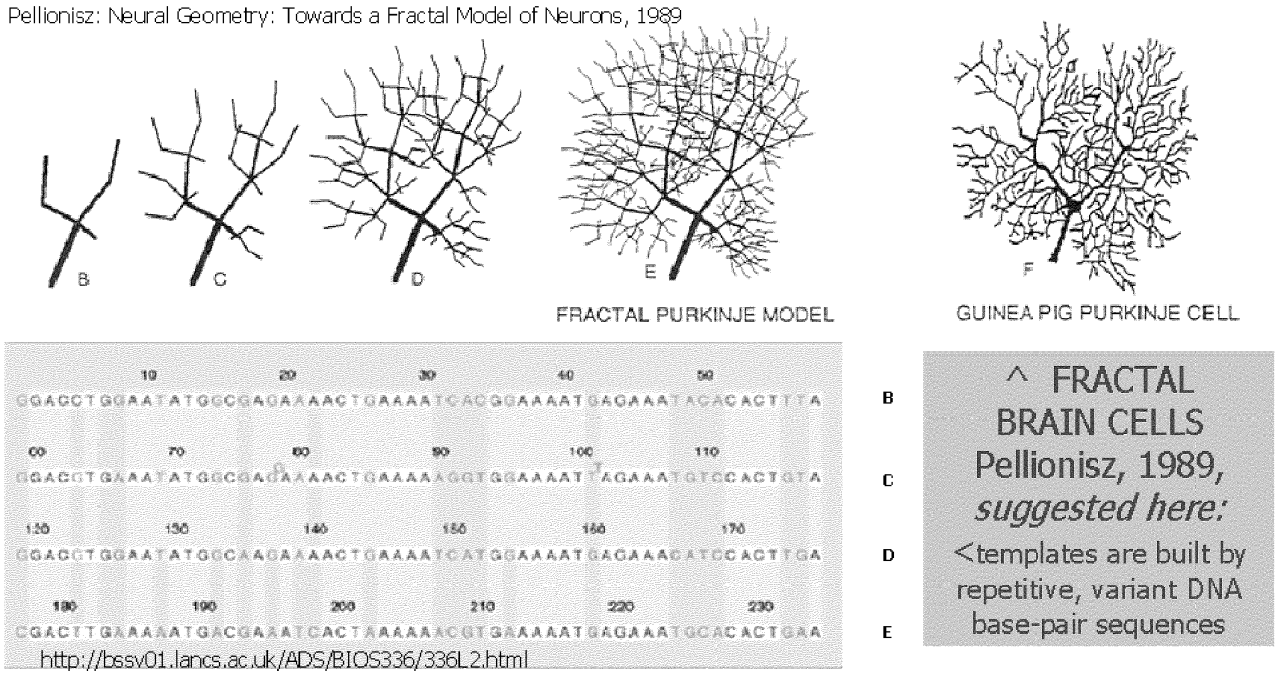

Figure 3. Postulates of the 'FractoGene' diagram, an actual copy of Figure 6 of the Patent Application 'FractoGene: Utility to use self-similar repetitions in the language-like genetic information as fractal sets'. US Patent Application (18) 1 August 2002. FractoGene postulated a causal relationship between fractal genetic function and P-cell arborization that prompted preliminary morphological study of *Fugu rubripes* (33), and of the (to be completed) study of *Danio rerio*, as its genome became available.

by the fractal template (or 'primitive') of stage B. Branchlet characters such as length and thickness need additional information that are not strictly defined but rather are parameters in the template.

An observation in 2002, thirteen years after publication of the fractal organelle model (9), of sequence elements that are self-similar but not identical (bottom left field, four consecutive DNA nucleotide sequences) prompted the FractoGene concept to explain the additional information not inherent in the fractal template, namely that the 'fractal-like' properties of P-cells, as well as of other organelles and organs (e.g., Cardial coronary arteries (39)), and the 'fractal-like' self-similar repetitions known to be in DNA since 1994 (12), are in a *causal relationship*, with fractal sets of DNA generating fractal anatomical (structural proteins), physiological (metabolic proteins) and related pathological formations.

FractoGene predicted that the Fugu P-cell arbor would take the form of the 'B' stage depicted in the 'FractoGene' diagram (Figure 6 of ref. 18, Figure 3 of this paper).

It is worthy of mentioning that the prediction of the 'FractoGene' diagrams for the several subsequent stages of fractal development are not just visually descriptive. Rather, the fractal model is mathematically deterministic. This means that the basic information is contained in the 'L-string fractal primitive' of 'B-stage' (18).

The morphological findings reported elsewhere and summarized here support the prediction. The P-cell patterning of Fugu is clearly the least developed of all organisms reported to date (Figure 4). Danio is predicted to have an intermediate level of P-cell arbor structure.

The FractoGene concept, with here-presented experimental support for the first prediction that Fugu should exhibit a rudimentary dendritic arbor, seeks also to explain the expected link between the progressive structural properties of P-cells, on the one hand, and the proteins expressed by these cells, on the other. Embodying iterative and recursive components in an evolving process, FractoGene offers a framework for explanation of the presence (Calbindin-D28k, PEP19/Pcp4, GAD67) or absence (L7/Pcp2) of proteins, and of their underlying genes.

It is clear that proteins with GoLoco domains exist in lower order species, including Fugu and Danio. Others have reported co-localization of GoLoco and L7/Pcp2 protein expression in the cerebellum, suggesting a ''functional role in regions of synaptic activity'' (24). Our finding that L7/Pcp2 may have evolved in higher vertebrates having more complex P-cell dendritic trees, and may not be present at earlier stages of evolution, offers approaches to functional correlates of genic presence and expression.

The haplostructural organization of non-coding DNA bespeaks functional correlates. The FractoGene concept envisages that such non-coding 'regulatory' and other elements will contribute an essential role to genic expression and function, and predicts new experimental approaches to non-coding functional sequence identification. Revealing the interrelation

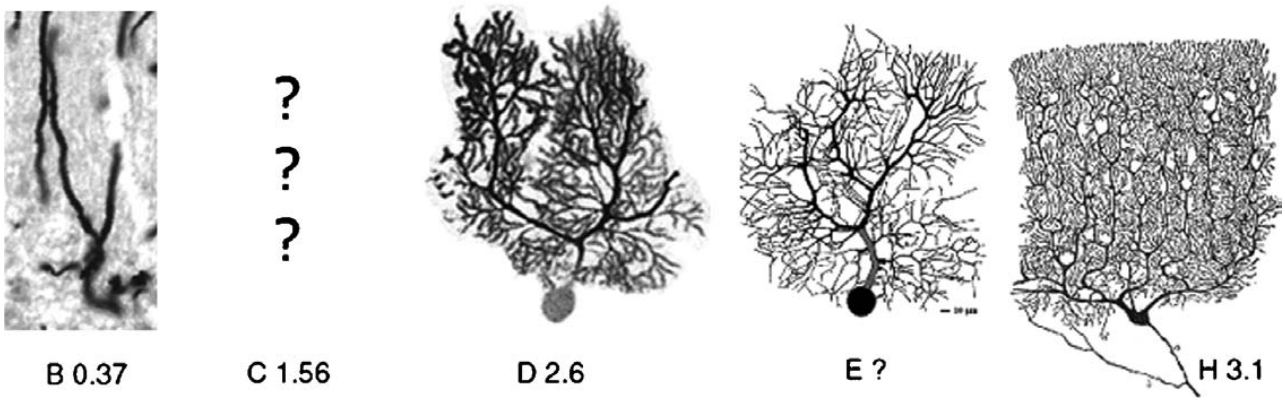

Figure 4. Sketch of the emerging field of comparison of the complexity of the dendritic trees of P-cells, their genomic analysis, calling for biophysical synthesis. Insert B shows the P-cell in the Fugu rubripes (B is courtesy of Székely (33)), in which the genome size is 0.37 Gigabases. C will show the P-cell in Danio rerio (as it becomes available, according to studies at an early stage to exhibit an interim complexity) in which the genome size is 1.56 Gigabases. D shows the dendritic arbor of the P-cell in the mouse (genome size is 2.6 Gigabases). Insert D is fluorescent-stained photo, courtesy of Prof. Helen Blau (40). E shows a computer-reconstruction of the P-cell in the guinea pig (21). The genome size in the guinea pig is not known to date, but its sequencing was slated (at Broad Institute and MIT) among other species. Insert H shows the P-cell of the human (27). The genome size in the human is 3.1 Gigabases.

between non-coding genetic function and protein coding sequence expression is a major goal of the post-Gene era.

The complexity of the genomic-morphological nexus is indicated by the apparent absence of a primary candidate gene L7/Pcp2 for P-cell development in Fugu, yet the gene is also undetectable in two genera of Xenopus (laevis and.tropicalis) in which P-cell dendritic arbor is distinct (28). Again, in Danio, in which L7/Pcp2 is also not evident, preliminary observations indicate an intermediate stage of P-cell arbor development. The more recent evolutionary emergence of basket cell neurons in Xenopus (the so-called 'Basic Cerebellar Circuit' notion, based on the lack of basket and Golgi cells (41)) may be relevant to resolution of these current disparities.

If L7/Pcp2 could be demonstrated in Fugu/Danio/Xenopus by histological, molecular biological or other procedures, then current sequence homology tools would need to be supplemented by a new generation of genomic search tools, based on significant leaps in algorithmic developments, such as triggered by FractoGene (18).

The complexity of the relationship between genomic function and morphological development requires experimentation to be guided by concepts such as that of FractoGene, which specifies a process (recursive instructions) and, here, a platform (P-cell).

Any viable alternative to the FractoGene model, such as a combinatorial approach[42] to DNA, would also need to identify a suitable process and platform.

However, it can easily be imagined that the fractal approach can be extended to combinatorial concepts. Components of the DNA in most animals are not a single monofractal, as appears to be the case in some simple plants, but are almost certainly multifractal. This seems evident from the different fractal properties of their separable organelles (lung, coronaries, intestines, etc.). Therefore, the FractoGene concept might have to be augmented with methods of sequence analysis.

It can also be predicted that neural network algorithms derived from biological organisms (43) and patents (44), will serve as novel computational elements to measure the self-similarity by fractals. Indeed, a recent neuronal network study shows already that fractal, neural network, multidimensional/combinatorial approaches to mathematical treatment of recursive and iterative biological processes have commenced (45). Interestingly this recent initiation is applied to sensorimotor sequences, which were the historical precursors to geometrization of biology (32). Application of fractal and combinatorial approaches together with neural network algorithms to genomic information analysis can be expected as the next development.

We regard it to be essential in this post-gene era that biological studies are guided by some mathematical (geometrical) model of genome-morphodevelopment interaction. While the imperative of mathematization may not be yet compelling to some morphologists and even genomists, the cardinal importance of mathematical tools is rapidly becoming understood by practitioners of nanotechnology attempting to create protein-based novel materials. Nanotechnologists recognize that protein-based Nano-construction requires both a quantitative understanding of the full genome, including introns and non-coding DNA, and appropriate quantitative 'blueprints' (46). For general essays on this issue, see (47,48). Indeed, if protein-based Nanotechnology mathematical tools are not provided by Genomics, then Nanotechnology can be expected to generate such tools on their own, outside of a true interdisciplinary cooperation (49–51).

Acknowledgement

We are thankful for the use of unpublished observations of Dr G. Székely and Dr N. Matsumoto (Department of Anatomy, University of Debrecen, Hungary, and Kyushu Institute of Technology, Japan). These authors wish to acknowledge the help of Mr J. Takata and I. Takami, (Marine Laboratory Nagasaki) with the collection of the animals.

Note added in proof

Our attention has been directed to a net-publication (Rothemund PWK, Papadakis N, Winfree E. 2004. Algorithmic self-assembly of DNA Sierpinski triangles: PLoS Biol 2(12):e424–7 December 2004) that mentions the likely utility of fractal approaches to genomic DNA analysis advocated herein.

The publication supports our proposal, ref.18 (2002), for which ref. 50 (2003) is an appropriate peer-reviewed citation, including vital nanotechnology aspects of fractal applications. The publication extrapolates to the field of fractal algorithms from claims made in 1998 by of one of the authors (Rothemund) that DNA is a ''molecular Turing machine'' capable of fabricating von Neumann's ''universal constructor''. There is no mention of fractals in the 1998 claims. Their publication now acknowledges that ''examination of self-assembly in modern organisms reveals many mechanisms beyond those considered here, including ... interactions with genetic regulatory networks''.

It is reassuring to note that our proposal for Information Technology-based decoding of the DNA using fractal analysis is beginning to be considered by pioneers of mathematization of biology in the field of Neural Networks, who are now turning their attention to mathematization (geometrization) of the full genome. John Hopfield (Institute of Integrative Genomics at Princeton) leads endorsement of the potential utility of fractal approach to DNA. The fractal approach to DNA was also endorsed by Benoit Mandelbrot earlier in 2004 by his Keynote Lecture at the IEEE-organized, HP-supported Computational Systems Bioinformatics Conference, ECSB2004, Stanford, August 18, 2004.

References

1. Gibbs WW. The unseen genome: Gems among the junk. Sci Am. 2003;289(5), 48–53.
2. Battistutti WB. Cross-talking between cells – the future of functional genomics? In: The future of drug discovery, London: Business Briefings; 2004. pp 54–58.
3. Mattick J, In, ''Genius of Junk'' (Producer Sonia Pemberton, Australian Broadcasting Company), 2003. Full transcript available at http://www.abc.net.au/catalyst/stories/s898887.htm#transcript.
4. Simons MJ. Intron sequence analysis method for detection of adjacent and remote locus alleles as haplotypes; 1993. US Patent # 5,192,659.
5. Simons MJ. Intron sequence analysis method for detection of adjacent and remote locus alleles as haplotypes; 1997. US Patent # 5,612,179.
6. Simons MJ. Genomic mapping method by direct haplotyping using intron sequence analysis; 1998. US Patent # 5,851,762.
7. Simons MJ, Limm TM, Naughton MJ, Quinn DL, McGinnis MD, Ashdown ML. Strategy for definition of DR/DQ haplotypes in the 4AOHW cell panel using noncoding sequence polymorphisms. Hum Immunol. 1993;38(1), 69–74.
8. Limm TM, Ashdown MJ, Naughton MJ, McGinnis MD, Simons MJ. HLA-DQA1 allele and suballele typing using noncoding sequence polymorphisms. Application of 4AOHW cell panel typing. Hum Immunol. 1993;38(1), 57–68.
9. Pellionisz AJ. Neural Geometry: Towards a fractal model of neurons. In: Cotterill RMJ, editor. Models of brain function, Cambridge: Cambridge University Press; 1989. pp 453–464.
10. Oliver JL, Bernaola-Galvan P, Guerrero-Garcia J, Roman-Roldan R. Entropic profiles of DNA sequences through chaos-game-derived images. J Theor Biol. 1993;160(4), 457–70.
11. Buldyrev SV, Goldberger AL, Havlin S, Peng CK, Stanley HE, Stanley MH, Simons M. Fractal landscapes and molecular evolution: Modeling the myosin heavy chain gene family. Biophys J. 1993;65(6), 2673–9.
12. Flam F. Hints of a language in junk DNA. Science. 1994;266:1320.
13. Mantegna RN, et al. Linguistic features of noncoding DNA sequences, Phys Rev Lett., 1994;73:3169–3172.
14. Oliver JL, Roman-Roldan R, Perez J, Bernaola-Galvan P. SEGMENT: Identifying compositional domains in DNA sequences. Bioinformatics. 1999;15(12), 974–9.
15. Chatzidimitriou-Dreismann CA, Streffer RM, Larhammar D. Lack of biological significance in the 'linguistic features' of noncoding DNA – a quantitative analysis. Nucleic Acids Res. 1996;24(9), 1676–81.
16. Bonhoeffer S, Herz AV, Boerlijst MC, Nee S, Nowak MA, May RM. No signs of hidden language in noncoding DNA. Phys Rev Lett. 1997;76(11).
17. Mandelbrot BB. The fractal geometry of nature. New York: WH Freeman; 1977.
18. Pellionisz A. FractoGene: Utility to use self-similar repetitions in the language-like genetic information as fractal sets. US Patent Application (1 August 2002).
19. Hodgkin AL, Huxley AF. A quantitative description of membrane current and its application to conduction and excitation in nerve. J Physiol. 1952;117:500–544.
20. Pellionisz A, Llinás R. Computer model of cerebellar Purkinje cells. Neuroscience. 1977;2:37–48.
21. De Schutter E, Bower JM. An active membrane model of the cerebellar Purkinje cell: I. Simulation of current clamps in slice. J Neurophysiol. 1994;71:375–400.
22. Manto MU. The wide spectrum of spinocerebellar ataxias (SCAs). Cerebellum. 2005;4(1), 2–7.
23. Santos MM, Miranda CJ, Levy JE, Montross LK, Cossee M, Sequeiros J, Andrews N, Koenig M, Pandolfo M. Iron metabolism in mice with partial frataxin deficiency. Cerebellum. 2003;2(2), 146–53.
24. Redd KJ, Oberdick J, McCoy J, Denker BM, Luo Y. Association and colocalization of G protein alpha subunits and Purkinje cell protein 2 (Pcp2) in mammalian cerebellum. J Neurosci Res. 2002;70(5), 631–7.
25. Barski JJ, Lauth M, Meyer M. Genetic targeting of cerebellar Purkinje cells: History, current status and novel strategies. Cerebellum. 2002;1(2), 111–18.
26. Koeppen AH. The pathogenesis of spinocerebellar ataxia. Cerebellum. 2005;4(1), 62–73.
27. Cajal yRamon. Histologie du systeme nerveux de l'homme et des vertebras. Paris: Maloine, 2 vols; 1911.

1 28. Hillman DE. Morphological organization of frog cerebellar cortex: A light and electron microscopic study. 1969.
29. Pellionisz A. Modeling of neurons and neuronal networks. In: Schmitt FO, Worden FG, editors. The Neurosciences, IVth Study Program.MIT Press, 1979. pp 525–46.
2 30. Tank DW, Sugimori M, Connor JA, Llinás R. Spatially resolved calcium dynamics of mammalian Purkinje cells in cerebellar slice. Science. 1988;242:633–828.
31. Nieuwenhuys R, Nicholson C. Aspect of the histology of the cerebellum in mormyrid fishes. In: Llinas R, editor. Neurobiology of cerebellar evolution and development, Chicago: AMA; 1969. pp 135–69.
32. Pellionisz A, Llinás R. Tensor network theory of the metaorganization of functional geometries in the CNS. Neuroscience. 1985;16:245–74.
33. Székely G, Nobuyoshi Matsumoto N. The Purkinje cell of the Fugu rubripes: A Golgi study. Unpublished; 2004.
34. Nordquist DT, Kozak CA, Orr HT. cDNA cloning and characterization of three genes uniquely expressed in cerebellum by Purkinje neurons. J Neurosci. 1998;12:4780–9.
35. Zhang X, Zhang H, Oberdick J. Conservation of the developmentally regulated dendritic localization of a Purkinje cell-specific mRNA that encodes a G-protein modulator: Comparison of rodent and human Pcp2(L7) gene structure and expression. Brain Res Mol Brain Res. 2002;105(1–2):1–10.
36. Fernandez-Gonzalez A, La Spada AR, Treadaway J, Higdon JC, Harris BS, Sidman RL, Morgan JI, Zuo J. Purkinje cell degeneration (pcd) phenotypes caused by mutations in the axotomy-induced gene, Nnal. Science. 2002;295(5561), 1904–6.
37. Vassileva G, Smeyne RJ, Morgan JI. Absence of neuroanatomical and behavioral deficits in L7/pcp-2-null mice. Brain Res Mol Brain Res. 1997;46(1–2):333–7.
38. 38., Mohn, AR, Feddersen RM, Nguyen MS, Koller BH. Phenotypic analysis of mice lacking the highly abundant Purkinje cell- and bipolar neuron-specific Pcp2 protein. Mol Cell Neurosci. 1997;9(1), 63–76.
39. Van Beek H, Bassingthwaighte JB. Fractal model of the coronaries of the human heart. Sci Am. 1990;46.
40. Weimann JM, Johansson CB, Trejo A, Blau HM. Stable reprogrammed heterokaryons form spontaneously in Purkinje neurons after bone marrow transplant. Nature Cell Biol. 2003;11(5), 959–66.
41. Llinas R. Frog cerebellum: Biological basis for a computer model. Math Biosci. 1971;11:137–51.
42. Pevzner PA. Combinatorial methods for DNA physical mapping. IEEE Comp Med Biol. 1994;13:146–50.
43. Anderson J, Pellionisz A, Rosenfeld E. Neurocomputing-2. MIT Press, 1989.
44. Pellionisz A, Llinás R. Sensorimotor coordinator; 1984. US Patent #4,450,530.
45. Nishimoto R, Tani J. Learning to generate combinatorial action sequences utilizing the initial sensitivity of deterministic dynamical systems. Neur Netw. 2004;(7):925–33.
46. Mao C, LaBean TH, Reif JH, Seeman NC. Logical computation using algorithmic self-assembly of DNA triple crossover molecules. Nature. 2000;407:493–6.
47. Seeman NC. Nanotechnology and the Double Helix. Sci Am. 2004;June:65–75.
48. Battistutti WB, Phueringer A. New strategies for understanding functional genomics. Business Briefings, Technology – Genomics and Pharmacogenomics. London: Business Briefings; 2005. pp 64–6.
49. Shu D, Huang L, Hoeprich S, Guo P. Construction of phi29 DNA-packaging RNA (pRNA) monomers, dimers and trimers with variable sizes and shapes as potential parts for nano-devices. J Nanosci Nanotechnol. 2003;3:1–8.
50. Pellionisz A. FractoGene design-tool for protein-based self-assembling nanostructures, materials and applications. Invited Keynote Lecture at Nano-Bio Technology Session, Proceedings of the 204th Meeting of The Electrochemical Society, 2003. pp 1195.
51. Shu D, Huang L, Guo P. A simple mathematical formula for stoichiometry quantification of viral and nanobiological assemblage using slopes of log/log plot curves. J Virol Methods. 2004;115:19–30.