



# Plurigon: three dimensional visualization and classification of high-dimensionality data

Bronwen Martin<sup>1†</sup>, Hongyu Chen<sup>2†</sup>, Caitlin M. Daimon<sup>1</sup>, Wayne Chadwick<sup>2</sup>, Sana Siddiqui<sup>2</sup> and Stuart Maudsley<sup>2\*</sup>

<sup>1</sup> Metabolism Unit, Laboratory of Clinical Investigation, National Institute on Aging, National Institutes of Health, Baltimore, MD, USA

<sup>2</sup> Receptor Pharmacology Unit, Laboratory of Neuroscience, National Institute on Aging, National Institutes of Health, Baltimore, MD, USA

## Edited by:

Firas H. Kobeissy, University of Florida, USA

## Reviewed by:

Anshu Bhardwaj, Council of Scientific and Industrial Research, India

Natalia Polouliakh, Sony Computer Science Laboratories Inc., Japan  
Mariette Awad, American University of Beirut, Lebanon

## \*Correspondence:

Stuart Maudsley, Receptor Pharmacology Unit, Laboratory of Neuroscience, National Institute on Aging, National Institutes of Health, 251 Bayview Blvd., Suite 100, Baltimore, MD 21224, USA  
e-mail: maudsleyst@mail.nih.gov

<sup>†</sup> These authors have contributed equally to this work.

High-dimensionality data is rapidly becoming the norm for biomedical sciences and many other analytical disciplines. Not only is the collection and processing time for such data becoming problematic, but it has become increasingly difficult to form a comprehensive appreciation of high-dimensionality data. Though data analysis methods for coping with multivariate data are well-documented in technical fields such as computer science, little effort is currently being expended to condense data vectors that exist beyond the realm of physical space into an easily interpretable and aesthetic form. To address this important need, we have developed *Plurigon*, a data visualization and classification tool for the integration of high-dimensionality visualization algorithms with a user-friendly, interactive graphical interface. Unlike existing data visualization methods, which are focused on an ensemble of data points, *Plurigon* places a strong emphasis upon the visualization of a single data point and its determining characteristics. Multivariate data vectors are represented in the form of a deformed sphere with a distinct topology of hills, valleys, plateaus, peaks, and crevices. The gestalt structure of the resultant *Plurigon* object generates an easily-appreciable model. User interaction with the *Plurigon* is extensive; zoom, rotation, axial and vector display, feature extraction, and anaglyph stereoscopy are currently supported. With *Plurigon* and its ability to analyze high-complexity data, we hope to see a unification of biomedical and computational sciences as well as practical applications in a wide array of scientific disciplines. Increased accessibility to the analysis of high-dimensionality data may increase the number of new discoveries and breakthroughs, ranging from drug screening to disease diagnosis to medical literature mining.

**Keywords:** *Plurigon*, three dimensional, data visualization, data classification, multivariate data vectors, algorithms, systems biology, bioinformatics

## INTRODUCTION

With the advent of large-scale repositories of scientific knowledge and the increasing prevalence of information science, researchers from multiple disciplines are often faced with the task of collecting, manipulating, and disseminating high-dimensionality data. Currently, the model of choice for dealing with such data is the vector space model, a system of representing any entity with a list of identifiers in vector form. For example, high-throughput screening assay data from PubChem presents each compound in their database as a 881-dimensional binary vector representing the absence or presence of various features (elements, ring systems, atom pairing, nearest neighbors) in the compound. Similarly, a list of experimental results from any test subject can be equated to a vector with a dimensionality equal to the cardinality of the number of tests performed.

Such a model is undoubtedly invaluable for its condensation of data into a computable space. However, with increasing dimensionality, various numerical analysis challenges such as sparseness, statistical insignificance, and computational difficulty arise (Bellman, 1957). For classification, Vapnik-Chervonenkis theory

and the Hughes effect state that theoretical classification rate decreases as dimensionality increases (Hughes, 1968). To address these issues, data is often preprocessed with feature extraction and dimensionality reduction techniques. Reducing dimensionality allows data to be processed more quickly and leads to reductions in noise, sparseness, and redundancy.

Additionally, as vectors often extend the scope of physical space, data visualization becomes an issue. Unfortunately, since presenting information in graphical ways is not necessary for our ability to extract answers, the field of data visualization lags far behind its sister disciplines of data analysis and data mining (Fayyad et al., 2002). This occurrence has led to a system in which only highly-specialized experts in esoteric technicalities can interpret the data at hand. If high-dimensionality data hopes to gain a larger audience however, advanced data visualization is instrumental in allowing end users with a rudimentary understanding of complex mathematics to interpret the graphical metaphors of the data at hand.

There are currently several excellent examples of high-dimensionality data visualization software including Ggobi

(<http://www.ggobi.org/>), Visumap (<http://www.visumap.net/>) and Iris from Ayasdi (<http://www.visumap.net/>). It is clear from the applications of these programs (Wurtele et al., 2003; Landauer et al., 2004; Nicolau et al., 2011; Lum et al., 2013) that the inclusion of the exquisite capacity for human-centric visual appreciation and recognition into complex data analysis is a fertile ground for future research into high-dimensionality data interpretation. *Plurigon* serves to further this current interest and provide a synergistic alternative to these already useful applications.

*Plurigon* provides unification of high-dimensionality algorithms with visual human interfaces by converting a vector that exceeds physical space into an easily interpretable and highly interactive three-dimensional object. Feature extraction can then be performed on the *Plurigon*, either visually or computationally, for classification or machine learning. The usage of these features for data analysis is unexplored but promising due to the ease of *Plurigon* visual interpretation. The most unique trait of the *Plurigon* is that it is currently the only data visualization technique that places an emphasis on individual data vectors as opposed to an ensemble of different data vectors. This aspect of *Plurigon* provides an alternative to other forms of high-dimensional data visualizers. For example, genes or proteins are likely to act in two different modes, at times there may be strong individual actions (e.g., amyloid precursor protein mutations in Alzheimer's disease; Maudsley and Mattson, 2006) while at other times a specific gene/protein may act in a collective manner with other genes/proteins (Mootha et al., 2003). In most physiological systems a combination of these two functional modes is likely to be apparent, and especially in the presence of relatively few data points, *Plurigon* may provide a valuable alternative to ensemble visualization. In addition, the actual physiological actions of gene transcripts or proteins are highly contextual, i.e., a gene or a protein may possess a wide range of potential functionalities, but depending on the activity of other functionally-related or physically proximal factors, this spectrum of activity may be both qualitatively and quantitatively affected. By creating a data-derived physical object we intend to allow the influence of each individual piece of data with each other to create a form that encodes all potential interactions via the revelation of a recognizable series of topologies. These structures therefore may be characteristic of the actual "gestalt" output of the altered series of genes/proteins in the physiological paradigm.

With the *Plurigon*, data mining and knowledge discovery are more easily accessible to everyone—providing for integrated solutions between the biological and computational sciences. Increased accessibility to the analysis of high-dimensionality data may increase the number of new discoveries and breakthroughs in science, ranging from drug screening to literature mining.

## MATERIALS AND METHODS

### DETERMINATION OF VERTICES ON THE PLURIGON

For an input of  $n$ -values corresponding to a vector in  $n$ -dimensional space, a *Plurigon* can be generated without loss as a set of spherical coordinates ( $r$ ,  $\theta$ ,  $\varphi$ ). While the radius captures the magnitude of each value  $\theta$ ,  $\varphi$  captures its location

in the original vector. Transformation of a data set into a *Plurigon* structure requires three steps: generation of a prototype structure with equal radii, remapping every point in the prototype to reflect actual data values, and iterative smoothing of the resulting *Plurigon* to remove sharp edges and unaesthetic qualities.

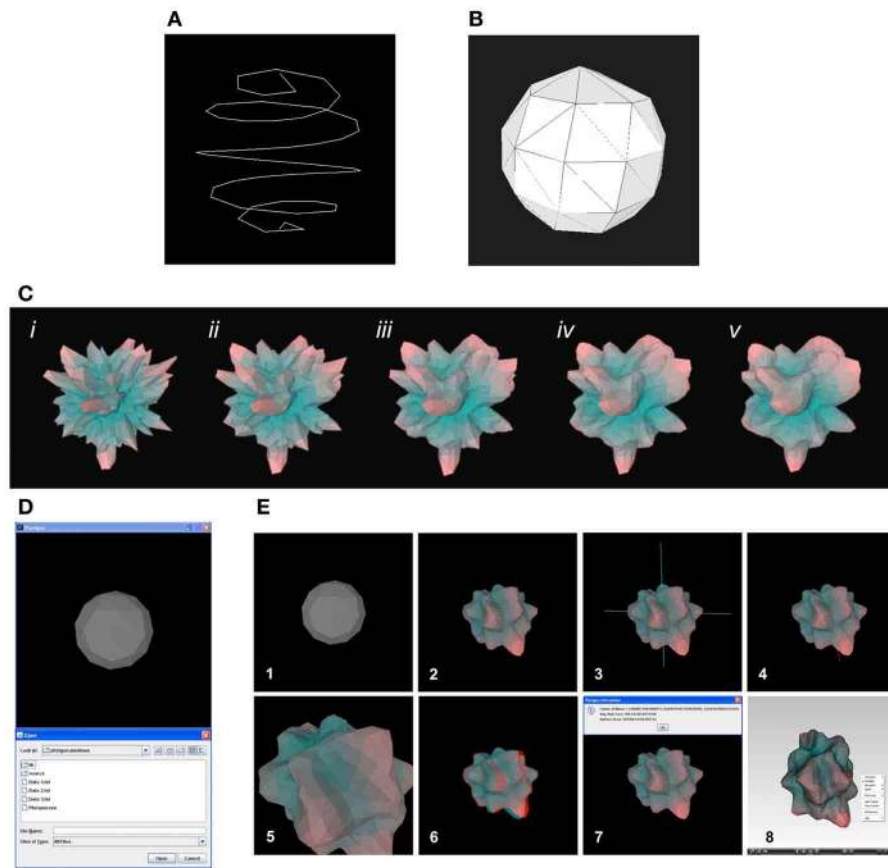
The vertices of the prototype *Plurigon* are generated by spacing  $n$  points on the prototype's circumsphere as far as possible. Unfortunately, this is a non-trivial task. Due to Euclid's proof that there are only five platonic solids, perfectly spaced points on a cube can only be achieved for dimensions 4, 6, 8, 12, and 20. In all other dimensions, perfect spacing cannot be achieved; however, there are a number of methods for approximating a distribution that minimizes the variance in distance between points. It is important to note that the naïve method of choosing points at equally spaced intervals of  $\theta$  and  $\varphi$  is insufficient because data points are much more concentrated near the sphere's poles (Cook, 1957). As such, current methods for spacing vertices on a sphere include hypercube rejection, creation of a simulation involving electron repulsion, and spiral tracing (Smith, 1984; Rakhmanov et al., 1995; Saff and Kujilaars, 1997; Thomsen, 2007).

For its ability to run in linear time, we use a slight improvement, created by Thomsen (2007), upon the methodology developed by Saff et al. for spacing points (Saff and Kujilaars, 1997), in which a larger spacing between the highest and lowest point better promotes point sparseness. This method falls into the category of spiral tracing, where a spiral is constructed with the endpoints as the sphere's poles and vertices placed at equal distances along the line segment (**Figure 1A**).

### GENERATION OF A POLYGONAL MESH FROM THE VERTICES

Optimal generation of faces from the prototype's vertices requires performing the Delaunay triangulation on the set of points (Delaunay, 1934; Lee and Schachter, 1980). Briefly, the Delaunay triangulation of a set of points,  $P$ , in two dimensions is defined as the triangulation,  $T$ , in which no point in  $P$  rests in any circumcircle of any triangle in  $T$ . Although Delaunay triangulation is always possible in two dimensions, when considering extensions to higher dimensions, triangulations are often impossible or not unique. Fortunately, in the case of points of equal radius from a sphere's center of mass (COM), Delaunay triangulations are not only always possible, but also computable via taking the convex hull of the collection of points. Since there are no points within the sphere, a convex hull should encompass all vertices on the *Plurigon* prototype.

Generation of a convex hull in  $n$ -dimensions is well documented in the field of computer science. Since the inception of the Jarvis march ( $O(n^2)$ ) method for computing convex hulls in 1973, a variety of algorithms have been discovered and employed with much lower time-complexity (Graham, 1972; Jarvis, 1973). Here we use the QuickHull ( $O(n \log n)$ ) divide-and-conquer algorithm outlined in Barber et al. for the construction of the prototype's triangular faces (Barber et al., 1996). The resulting three dimensional polyhedron contains  $n$  vertices,  $n$  triangular faces, and can be contained in a circumsphere such that all vertices rest on its surface (**Figure 1B**).



**FIGURE 1 | Generation of *Plurigon* structure and its general manipulation.**

**(A)** Initial *Plurigon* backbone creation. An illustration of vertex placement through spiral tracing. A set of 50 points was placed on the sphere at approximately equal distances from each other. **(B)** *Plurigon* polygonal basic structure. Convex hull generated from the vertices shown in **(A)**. This is the completed version of the *Plurigon* prototype; radial distance to the core remains constant for all vertices. **(C)** Laplacian smoothing in progress. Data was taken from a subset of gene expression values from murine genomic expression data. Iterations shown are *i*, *ii*, *iii*, *iv*, and *v*, afterwards, the movement of points becomes negligible, so iteration is stopped. **(D)** Initial *Plurigon* interface. The basic start-up *Plurigon* is depicted in an image window. **(E)** Simple and advanced *Plurigon* operations. Pressing “o” initiates the ability to choose a specific file to be depicted (1). Loading and pre-processing of data text file results in the generation of the basic color-coded *Plurigon* (2). Rotation of the *Plurigon* in all three dimensions is achieved using the up/down and left/right cursor keys. Addition of any other visualization features onto the

*Plurigon* does not affect the rotational capacity. Pressing “x” generates the superimposition of *x*, *y*, and *z* axes onto the *Plurigon* (3). Pressing “x” while the three axes are present toggles the axes off. This action format is conserved for all other forms of *Plurigon* visualization. Pressing “c” superimposes the vector position for the *Plurigon* center of mass (COM) (4). This COM is represented by a red line. Pressing “+” or “-” generates an ability to zoom in and out of the *Plurigon* (5). A 3-dimensional (3-D) viewing version of the *Plurigon* is generated by pressing the number “3” (6). Pressing “3” again while in the 3-D mode removes this visualization format. Simple output of basic *Plurigon* structural information is achieved by pressing “i” (7). The ability to save a TIFF picture file of the window view of the *Plurigon* is achieved by pressing “s.” For each of the functions, sequential superimposition upon the *Plurigon* can be achieved using the respective key functions. For export to further 3-D viewing applications a.vrml/wrl file of the *Plurigon* can be generated by pressing “v” (8). The image depicted is viewed using a Cortona-3D viewing application ([www.cortona3d.com/Products/Viewer/Cortona-3D-Viewer.aspx](http://www.cortona3d.com/Products/Viewer/Cortona-3D-Viewer.aspx)).

### CREATION OF THE *PLURIGON* FROM A PROTOTYPE

The purpose of creating a prototype is to ensure that the number of vertices on the resulting *Plurigon* is the same as *n*, the dimensionality of the data vector, during calculation of the convex hull. This was calculated with a unit vector to ensure that no points are located in the interior of the sphere. After calculation of the convex hull, the radii, *r*, of the prototype must be replaced by the individual input data values. The resulting surface should still be continuous, though during this phase it will likely be excessively turbulent and undulating for aesthetic and interpretable viewing.

To address this issue, the surface of the *Plurigon* must be smoothed and normalized until distinct topographic features, i.e., troughs, hills, peaks, and crevices, can be viewed. To this end, iterative Laplacian smoothing is applied to the polygonal mesh. Laplacian smoothing is a widely used technique in a variety of scientific fields (Briere and George, 1995; Amenta et al., 1997; Canann et al., 1997). Since its inception, many optimizations and improvements have improved the aesthetic outcomes of the smoothing process. Most notably, it has become possible to smooth a surface while maintaining Delaunay triangulation (Herrmann, 1976; Field, 1988; George and Borouchaki, 1998).

Since maintenance of the *Plurigon's* Delaunay triangulation is crucial, a version of the smoothing algorithm outlined in (Field, 1988) was used. After a certain amount of iterations, where the movement of vertices becomes negligible, iteration is terminated and a highly smooth surface with distinct contours can be observed (**Figure 1C**).

### EXTRACTION OF RUDIMENTARY FEATURES

As an example of *Plurigon's* feature extraction capabilities, the latest release of *Plurigon* facilitates the automated calculation of a small number of built-in global features. The COM is calculated by converting spherical coordinates to cartesian coordinates and then computing the mean of the  $x$ -,  $y$ -, and  $z$ -values. Average radius to the *Plurigon's* core, i.e., the origin, is calculated by the mean of all radial values. Finally, surface area is computed by applying Heron's formula to each of the *Plurigon's*  $n$  faces. Other advanced features that can be extracted from *Plurigons* that may facilitate integration with other interaction-based technologies such as tangible user interfaces (TUIs) (Ratti et al., 2004), but are not included in the program for their length of compute time include: angular momentum for spinning *Plurigons*, linear momentum for moving *Plurigons*, and recognition of specific local analogous features (valleys, hills, peaks, crevices) with automated pattern recognition. With ease of human visualization and interpretation, however, specific feature extraction algorithms can be generated to suit the specific experiment at hand. It is interesting to note, however, that feature extraction may be infeasible for certain applications. Despite the superficial simplicity of *Plurigon* structures for the visual appreciation of complex datasets, simple machine-based feature extractions can rapidly become computationally impractical. For example, the triangulation of the *Plurigon* into tetrahedrons cannot be computed in polynomial time and is NP-hard (Freund and Orlin, 1985).

### FUNCTIONAL FEATURES OF PLURIGON

The *Plurigon* interface exists as a Java application in either Windows, Mac OSX, or Linux formats. *Plurigon* is freely available for download from the National Institute on Aging/National Institutes of Health website (<http://www.irdp.nia.nih.gov/bioinformatics/plurigon.html>) (**Figure S1**). For review purposes, we have uploaded a Windows and Mac version of *Plurigon* for the reviewers and editor to test. The *Plurigon* application can be controlled entirely by keyboard (**Figures 1D,E**). Pressing "o" initiates the ability to choose a specific file to be depicted (**Figure 1D**). Loading and pre-processing of data text file results in the generation of the basic *Plurigon* (**Figure 1E**). Rotation of the *Plurigon* in all three dimensions is achieved using the up/down and left/right cursor keys. Addition of any other visualization features onto the *Plurigon* does not affect the rotational capacity. Pressing "x" generates the superimposition of color-coded (pink, yellow and blue)  $x$ ,  $y$ , and  $z$  axes onto the *Plurigon*. Pressing "x" while the three axes are present toggles the axes off. This action format is conserved for all other forms of *Plurigon* visualization. Pressing "c" superimposes the vector position for the *Plurigon* COM. This COM is represented by a red line. Pressing "+" or "-" generates the ability to zoom in and out

of the *Plurigon*. A 3-dimensional (3-D) viewing version of the *Plurigon* can be generated by pressing the number "3." Pressing "3" again while in the 3-D mode removes this visualization format. Output of basic *Plurigon* structural information is achieved by pressing "i." The output information box details the three dimensional coordinates of the calculated COM ( $x$ ,  $y$ ,  $z$ —format), the average radius of the plurigon structure from the central core of the platform (Avg.Rad.Core) and the total plurigon surface area (Surface Area). For future versions of *Plurigon* additional feature extraction tools will be developed (see Conclusions). The ability to save a TIFF picture file of the window view of the *Plurigon* is achieved by pressing "s." For each of the functions, sequential superimposition upon the *Plurigon* can be achieved using the respective key functions. For export to further 3-D viewing applications a.vrml/.wrl file of the *Plurigon* can be generated by pressing "v." The image depicted can be viewed using a Cortona-3D viewing application ([www.cortona3d.com/Products/Viewer/Cortona-3D-Viewer.aspx](http://www.cortona3d.com/Products/Viewer/Cortona-3D-Viewer.aspx)).

### MURINE HYPOTHALAMIC TRANSCRIPTOMIC INVESTIGATION

Wildtype C57BL6 mice were housed and employed in accordance with the Animal Care and Use Committee (ACUC) regulations at the NIH National Institute on Aging. Briefly, mice, three per age (3, 6, 12, and 18 months of age) and gender group (male, female), were humanely sacrificed and their hypothalamic tissue was excised rapidly and snap frozen for Illumina Fluorescent Gene Array analysis as described previously (Martin et al., 2012b).

### HETEROZYGOUS GENOTYPE TRANSCRIPTOMIC INVESTIGATION

Four month old male wildtype C57BL6 (WT) and G protein-coupled receptor kinase-interacting transcript 2 (GIT2)-heterozygous (GIT2<sup>-/+</sup>, aka HET) mice were housed and employed in accordance with the ACUC regulations at the National Institute on Aging. Briefly, multiple mice from both genotype groups (WT or HET) were humanely sacrificed and tissue extracts from the following organs were prepared for gene array analysis: hypothalamus, hippocampus, skeletal muscle, liver, pituitary gland and testes. For *Plurigon* analysis a similar tissue-based range of transcriptomic data was simultaneously analyzed for both genotypes under study, i.e., WT and HET.

### ANTI-NEURODEGENERATIVE TRANSCRIPTOMIC INVESTIGATION

Clonal human neuronal cells, SH-SY5Y, were employed to study the pro-neurotrophic actions of the tri-cyclic antidepressant amitriptyline (AMI). SH-SY5Y cells (American Type Culture Collection) were maintained in a humidified 5% CO<sub>2</sub> atmosphere at 37°C as described previously (Chadwick et al., 2010). We have previously demonstrated that in an aging-neurodegenerative murine model AMI exerts strongly neurotrophic pharmacological activity. We therefore employed *Plurigon* to investigate and analogize the activity of AMI compared to endogenous classical neurotrophic peptides such as brain-derived neurotrophic factor (BDNF) and nerve growth factor (NGF). Transcriptomic responses to AMI (10 nM), BDNF (10 ng/mL) or NGF (10 ng/mL) stimulation (8 h) of SH-SY5Y cells were assessed as previously described (Martin et al., 2009; Chadwick et al., 2012). AMI-hydrochloride, BDNF and NGF were all obtained



from Sigma Aldrich (St. Louis, MO). In addition to assessing the transcriptomic responses of these human neuronal cells to AMI, BDNF, and NGF we also assessed the same activity responses in SH-SY5Y cells pre-treated with a chronic minimal peroxide exposure protocol designed to mimic age-related neurodegeneration (Chadwick et al., 2010; Martin et al., 2012a). This oxidative insult consists of a chronic (7 days) exposure to a survivable and minimal concentration (10 nM) of the oxidizing agent hydrogen peroxide. Transcriptomic responses to these three ligands, AMI, BDNF, and NGF, were measured as described previously. In addition to transcriptomic effects, protein expression profiles for various proteins were assessed using selective antibody-based western blotting and immunoprecipitation procedures described previously (Maudsley et al., 2000). Western blotting procedures were as performed previously (Martin et al., 2009): the sources of the primary antibodies employed in this study are detailed in **Table S1**. Subcellular fractionation of SH-SY5Y cell proteins was performed to separate intracellular proteins between Golgi and endoplasmic reticular compartments as described in Ko and Puglielli (2009).

## RESULTS AND DISCUSSION

### DESCRIPTION OF PLURIGON GENERATION AND THE USER INTERFACE

The *Plurigon* software application aims to facilitate the transformation of high-volume data into a simpler, more appreciable structure. We term this resultant three-dimensional structure a “plurigon.” To create this data structure we use spiral tracing, where a spiral is first generated with the endpoints as the sphere’s poles and vertices placed at equal distances along the line segment (**Figure 1A**). We then use a divide-and-conquer algorithm to generate a solid figure with triangular faces on the three-dimensional solid contained in a circumsphere with all the vertices resting on its surface (**Figure 1B**). After data input and data-magnitude color-coding, surface smoothening and normalization is applied to generate a more aesthetic *Plurigon* with more easily-appreciable contoured topography (**Figure 1C**). A detailed description of the *Plurigon*-generating computational steps is outlined in the Materials and Methods section and a flowchart of the functional data transition through *Plurigon* is depicted in **Figure S2**. *Plurigon* is available as a lightweight, standalone Java application, and as with other visualizers such as Ggobi, is available in versions for Windows, Mac OSX, and Linux. Untagged data can be uploaded into the *Plurigon* program with a.txt file containing precisely three floating-point numbers per line delimited by newlines. For direct comparison of comparable data, e.g., similar denominating factors with variable numerators, appropriate pre-processing should be performed by the user. Memory and computational requirements for *Plurigon* are markedly low. The algorithm for *Plurigon* generation outlined in the Materials and Methods section is highly scalable because of its logarithmic time-complexity. As a result, graphics rendering is typically smooth even on very large data sets (~20–40,000 features) and with low-end computers (<256 MB Random Access Memory). Rendering is typically instantaneous with frame rates of over 30 in practice. *Plurigon* is accessible completely from the keyboard. User functionality for visualization, portability, and interaction is extensive, as described in section Functional Features

of *Plurigon*. *Plurigons* are automatically normalized to emphasize differences within data sets. Additionally, peaks, plateaus, and troughs within the *Plurigon* are color-coded with varying shades of red (highest value scores), gray, and cyan (lowest value scores), respectively.

### VISUALIZATION OF HIGH DIMENSIONALITY DATA SETS

Visualization of high-dimensionality data sets provides for the clear communication of information, unites various scientific disciplines, and makes otherwise arcane or highly-cryptic information more easily accessible. The current state of data visualization revolves around the representation of data matrices—an ensemble of data points. For example, when data is sparse and slight information loss is not an issue, a common practice is to perform Principal Components Analysis (PCA) to extract three or fewer components and then graph the resulting data points in a three-dimensional scatterplot (Wold et al., 1987; Kim et al., 2007). Other methods have been developed for even larger data sets. Two online visualization suites, Mondrian (<http://mondrian.pentaho.com/>) and Ggobi, both employ polylines with parallel coordinates, a common method for visualizing multivariate data by drawing  $n$  equally-spaced, parallel, vertical lines and plotting data along each line (Swayne et al., 1998; Theus, 2002; Buja et al., 2003; Cook and Swayne, 2007). Despite the considerable attention given to the relationships between data vectors, each individual vector is often neglected. For example, Ggobi represents a single data vector as a linear chain, which can quickly become chaotic and unintelligible once dimensionality exceeds a certain threshold. *Plurigon* provides for an alternative visualization mechanism by placing the emphasis not on the ensemble of data points, but rather on the individual data points themselves. The chaos of a single Ggobi linear chain is not present in the compact, spherical representation of the *Plurigon* even with dimensionalities in the tens of thousands. The effective employment of multiple distinct strategies for data investigation is likely to provide adequate solutions for a wide range of scientists. For *Plurigon* emphasis is dedicated to the prominence of a single, high-dimensionality data vector, whether it be a test subject, a compound, a document, or a gene. Strong deformities in topography could potentially be identified as important features. Despite the emphasis on individual points, however, there still exists a capacity for visualizing multiple subsets of vectors in a data matrix by comparing multiple *Plurigons* simultaneously for structural analogies or by subtracting or averaging data vectors with a data preprocessor before visualization.

### PRACTICAL EMPLOYMENT OF PLURIGON

As *Plurigon* represents a versatile tool that can be applied to multiple forms of biomedical investigation we chose to use it to interrogate three data from three distinct but functionally-related sets of transcriptomic data. The three experimental paradigms we used to investigate the implementation of *Plurigon* are all associated with the study of aging and age-related neurodegeneration. The first paradigm (*Plurigon Investigation of Murine Hypothalamic Aging Profiles*) involves scrutiny of gender-based aging in the hypothalamus (a core central nervous system tissue for the control of aging). The second paradigm (*Plurigon Investigation of*

Complex Murine Genotypic Profiles) challenges the ability of *Plurigon* to identify and classify large transcriptomic datasets from animal tissues from wildtype and heterozygous transgenic mice with reduced expression of a protein crucial for aging, GIT2 (G protein-coupled receptor kinase-interactor 2: Chadwick et al., 2012). The final paradigm (Plurigon Investigation of Novel Drug-Response Profiles) uses *Plurigon* to investigate novel anti-neurodegenerative drug mechanisms.

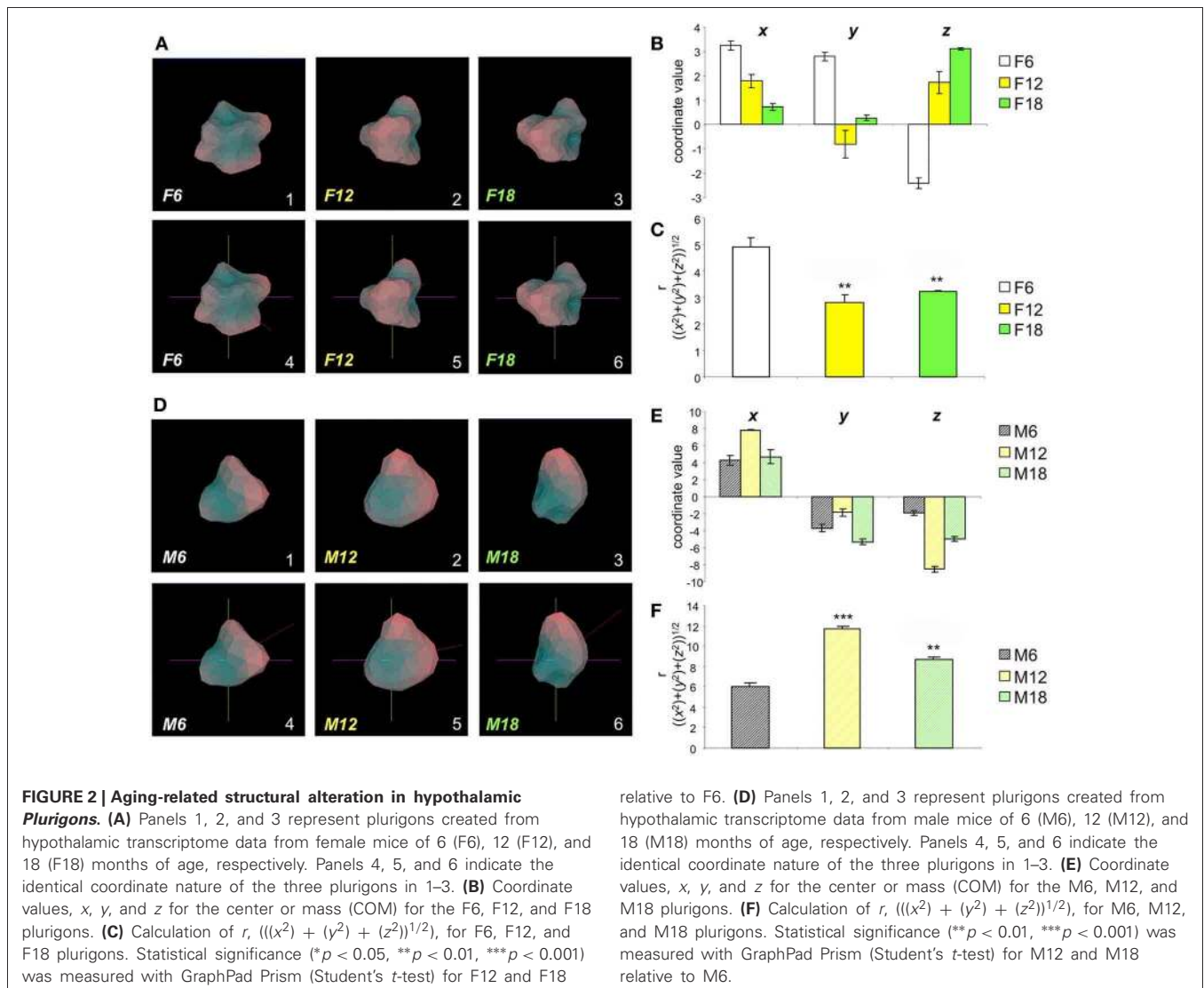
### Plurigon investigation of murine hypothalamic aging profiles

We have previously identified the hypothalamus as one of the most important physiological loci for both aging and neurodegenerative activities (Chadwick et al., 2012; Martin et al., 2012b). In the current study, we employed hypothalamic gene transcriptomic data from wildtype mice, of both genders, across a broad age range (6, 12, 18 months) (Figure 2; Tables S2–S4 Female, Tables S5–S7 Male). After uploading of the transcripts consistently regulated across all three mouse ages (compared to 3 month old controls) for both genders there is a clear

hypothalamic *Plurigon* structural “evolution” with the increasing age (Figure 2A: Female; Figure 2D: Male). As a simple indicator of the utility of our *Plurigon* program, we have just taken one output function, i.e., the COM calculation, to demonstrate an analytical function of our program. Using the Cartesian  $r$ , transform for the three spherical coordinates ( $x, y, z$ ) we found that there are statistically significant changes ( $p < 0.05$  see Figures 2; B,C Female, E,F Male) in the COM output with age in both females and males. Therefore, with relatively simple feature extraction *Plurigon* is able to generate easily-appreciable models of simple transcriptomic data associated with highly complex biological functions such as aging. The addition of further feature extraction modes will undoubtedly aid further structural differentiation pipelines.

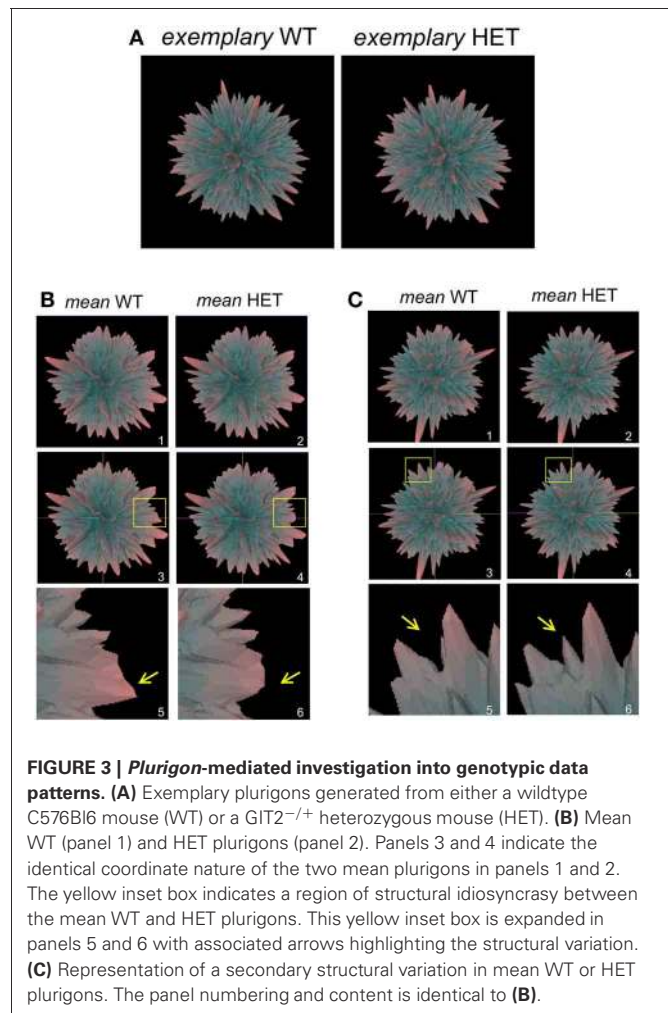
### Plurigon investigation of complex murine genotypic profiles

In addition to employing *Plurigon* for relatively small datasets we also introduced a considerably larger data burden to *Plurigon* to assess its ability to separate highly complex transcriptomic



data. To this end we also endeavored to assess whether *Plurigon* could effect a data group separation that was not achievable using PCA to three dimensions. We have therefore compared two heterogeneous, but functionally-related, groups of data, i.e., transcriptomic data from multiple murine tissues from either wildtype (WT) mice or mice with a heterozygous condition for the aging-regulator, *GIT2* (Chadwick et al., 2012). *GIT2* has been demonstrated to exert multiple cellular effects in mice (Schmalzigaug et al., 2009; Chadwick et al., 2010, 2012; Wang et al., 2012) therefore reduction in the expression of this important gene is likely to cause systemic effects in the animal, therefore there should be multiple differences in the transcriptomes between WT ( $n = 17$ ) and the HET ( $n = 19$ ) mice. Therefore, we employed *Plurigon* to attempt to separate these two sets of data (Table S8 WT, Table S9 HET). For these two genotypic examples our data corpus for each set has a dimensionality of over 20,000. Generating and scrutinizing individual plurigons for each mouse dataset (36 in total), therefore we initially chose to investigate whether systemic differences could be visually distinguished between “mean” plurigons for each genotype (Figure 3). We found that visual user scrutiny of both exemplary individual and the mean plurigons for each genotype facilitated the identification of complex and genotype-specific structural features (Figures 3A–C, S3–S7). We were also able to re-trace the origins of structural features observed in the mean plurigons back to individual animal plurigon data (Figure 4). Using visual user-based scrutiny it is clear that additional levels of data description can be generated with three dimensional plurigons, i.e., structural encryption and topological analogy. The presence or absence of view-obscuring factors can present or reveal distinctive features to the viewer that are discriminatory between complex datasets (Figure 3). Visual scrutiny is particularly good at quickly disseminating critical differences from less significant differences, even in large data sets. *Plurigon* outperforms PCA in this regard by virtue of being lossless. Distinctive features on the *Plurigon* can theoretically be mapped back to the individual dimensions by a second pass through the generation process. In this respect a counterpart for PCA does not exist.

We also found that structural topologies were also useful as characteristic agents, especially when compared to PCA-based analyses. Upon applying standard PCA algorithms, down to three dimensions, to the 36 mouse datasets we found the two distinct groups, WT or HET were non-linearly separable (Table S10: Figure S8). Unfortunately one of the main drawbacks of PCA is its inability to handle non-linearity, however human visual objects categorization however can help in this issue. Hence while PCA was unable to distinguish between the two groups we found that running a linear SVM (support vector machine) classifier on the genotypic data perfectly separates the two groups in 36 dimensions, suggesting that there are discriminatory structural plurigon features even in this reduced state. As the plurigons themselves are highly complex structures we computationally discovered structural motifs that support this genotype discrimination. Firstly we reduced the complexity of the WT/HET plurigons (with PCA) down to 36 dimensions for simpler visualization. Secondly we algorithmically selected three structural regions that could result in near-perfect separation of the genotypic data.

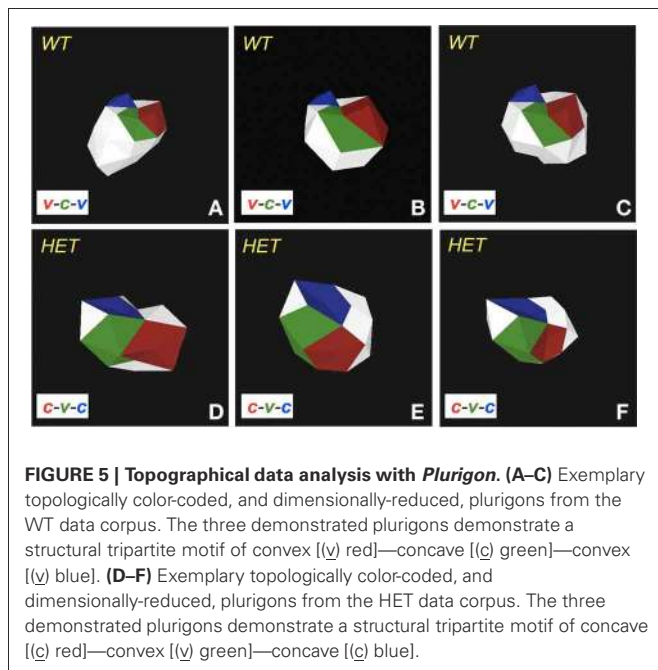
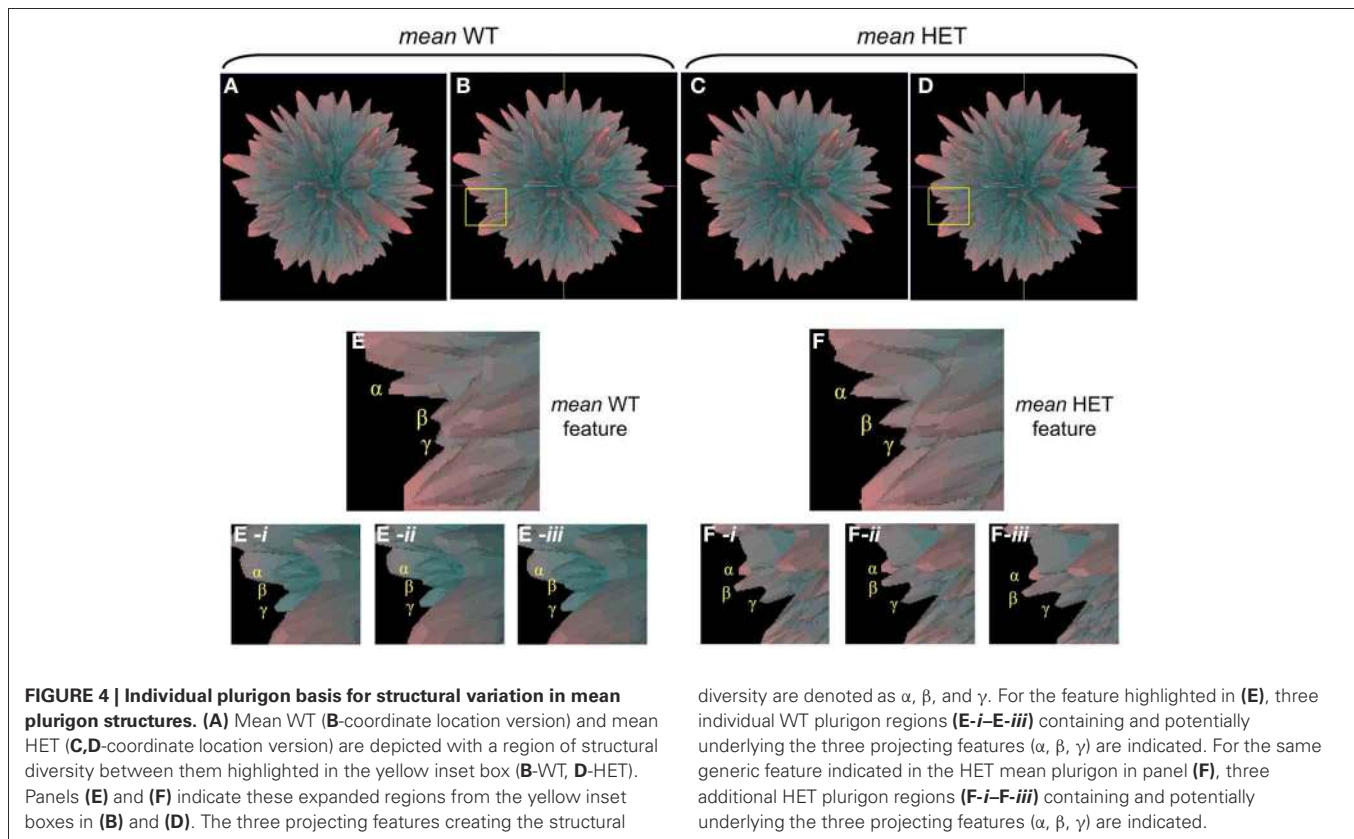


Analyzing the three dimensional nature, convex (denoted by  $v$ ) or concave (denoted by  $c$ ), of the three specific regions we found that their relationship to each other engendered the near-perfect data separation (Figure 5). The order of structure of these three faces (highlighted in red, green or blue) was therefore able to strongly discriminate the WT group ( $v-v-c$  or  $v-c-v$ ) from the HET group ( $c-v-c$  or  $v-v-v$ ) (Figure 5: Table S11). Therefore, in this circumstance it seems that using the structural nature and topographical data extraction of *Plurigon* can offer some assistance where classical three dimensional PCA is limited. Linked to this, one important feature of *Plurigon* is that it can retain an arbitrary number of dimensions (in this case 36) while visualizable PCA is limited to two or three dimensions. Both PCA- and *Plurigon*-mediated analysis possess benefits and flaws and the implementation of them will depend on the eventual user preference, nature of the dataset and also the specific questions asked of the data.

#### **Plurigon investigation of novel drug-response profiles**

Over the past decade our appreciation of the potential complexity of drug responses has necessitated the development of higher-density drug-response data (Maudsley et al., 2005, 2011b;





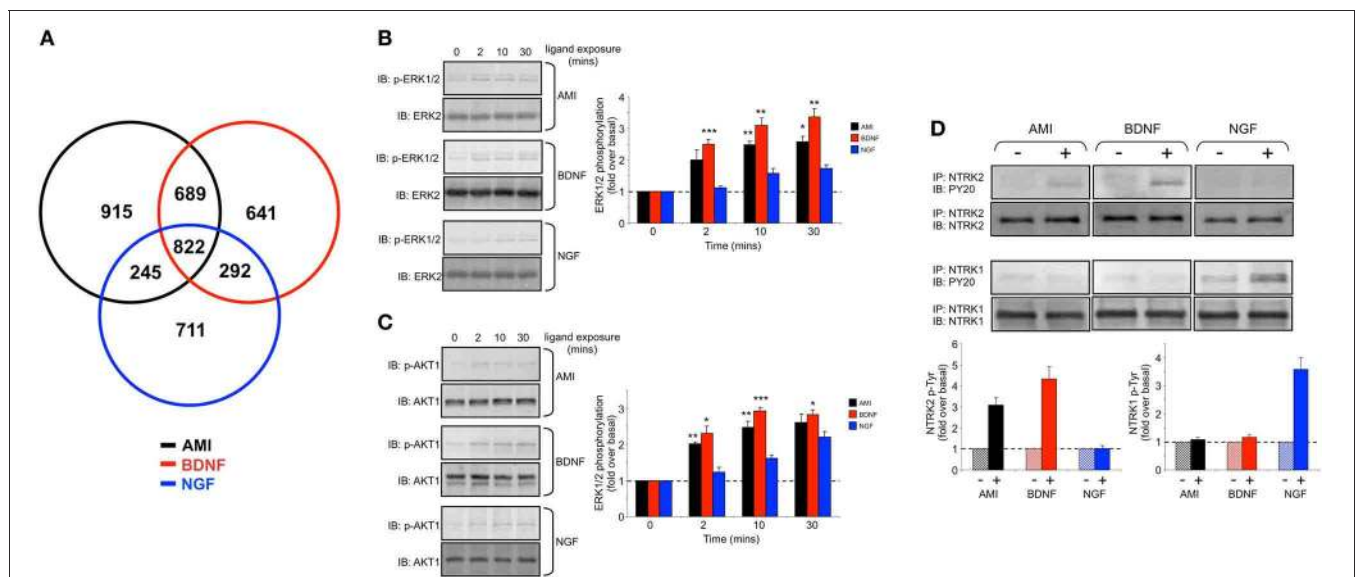
Luttrell and Kenakin, 2011). Potentially one of the most efficient ways of generating this high-density drug response data is the use of transcriptomics (Ruf et al., 2003; Gesty-Palmer et al., 2013). It is now known that many “small molecule” drugs, for

G protein-coupled receptors or other targets, can indeed exert strong transcriptomic and genomic responses (Golan et al., 2009; Chadwick et al., 2011) that were previously considered to be solely regulated by growth factor receptor systems. We have previously shown that the small molecule tricyclic AMI is able to exert beneficial therapeutic actions in aged mice that possess a pathophysiological signature of Alzheimer’s disease (Oddo et al., 2003; Chadwick et al., 2011). The actions of AMI seemed to be strongly associated with the activation of neurotrophin-like signaling pathways. Neurotrophins are large endogenous growth factors that, via activation of their cognate receptors (of the receptor tyrosine kinase family), entrain strongly neuroprotective signaling pathways (e.g., v-akt murine thymoma viral oncogene homolog 1 (AKT-1) phosphorylation) as well as regulating signaling pathways (e.g., extracellular signal-regulated kinase 1/2 (ERK1/2) phosphorylation) and the expression of markers (post-synaptic density protein 95 (PSD95)) linked with neuronal development. The three primary neurotrophin ligands are NGF, BDNF, and neurotrophin-3 (NT3). Respectively these ligands activate their cognate neurotrophic tyrosine kinase receptors (NTRK), i.e., NTRK1, NTRK2 and NTRK3. The majority of research into these neurotrophin systems has focused on the therapeutic benefits of NGF and BDNF signaling. In Alzheimer’s disease animals we previously found that the beneficial actions of AMI appeared to stem from its ability to mimic the actions of BDNF, rather than that of NGF. This finding was unexpected as AMI is considered a “small” drug molecule (0.277 kDa) while BDNF is approximately 13 kDa in mass, therefore despite their



completely different physico-chemical properties they seemed to share strong functional analogy in the central nervous system of aged and diseased mice. We therefore chose to investigate, at the transcriptomic response level, whether *Plurigon* would be able to assist in the characterization of structures indicative of specific forms of pharmacological activity, such as neuroprotective and neurodevelopmental mechanisms. Stimulation of human clonal neuronal cells, SH-SY5Y, with AMI, BDNF or NGF at the 8 h time point (Table S15-AMI, Table S16-BDNF, Table S17-NGF) we found after creating the plurigons for these three ligands that strong visual distinctions were apparent (Figures 7A,B, S9, S10). In each of these example plurigons the structural idiosyncrasies demonstrated a strong similarity between the AMI and BDNF structures, both of which were distinguishable from the NGF structures. With respect to the potential molecular mechanism of BDNF mimicry and the therapeutic effects observed in Alzheimer's disease animals (Chadwick et al., 2011) we found that chronic cellular treatment (24 h) with AMI resulted in a disruption of the cellular disposition (measured using subcellular biochemical separation) of components (anterior pharynx defective 1 homolog A (APH1A) and nicastrin (NCSTN)) of the amyloid precursor protein (APP) processing complex,  $\gamma$ -secretase (Figure S11A). Aberrant processing of APP is a hallmark of the Alzheimer's disease pathological process. In addition to changes in cellular distribution of components of the  $\gamma$ -secretase complex we found alterations in NTRK2 and the multifunctional adapter protein  $\beta$ -arrestin 2 (ARRB2). We

In contrast, NGF cellular stimulation (20 min) resulted in the increase of phosphotyrosine content of the NTRK1 receptor (Figure 6D). These data therefore demonstrate that despite considerable physico-chemical differences AMI is able to mimic many of the functional effects of BDNF. We then assessed whether this distinction could be translated into distinct *Plurigon* structures. Using transcripts commonly activated between AMI, BDNF and NGF (Table S15-AMI, Table S16-BDNF, Table S17-NGF) we found after creating the plurigons for these three ligands that strong visual distinctions were apparent (Figures 7A,B, S9, S10). In each of these example plurigons the structural idiosyncrasies demonstrated a strong similarity between the AMI and BDNF structures, both of which were distinguishable from the NGF structures. With respect to the potential molecular mechanism of BDNF mimicry and the therapeutic effects observed in Alzheimer's disease animals (Chadwick et al., 2011) we found that chronic cellular treatment (24 h) with AMI resulted in a disruption of the cellular disposition (measured using subcellular biochemical separation) of components (anterior pharynx defective 1 homolog A (APH1A) and nicastrin (NCSTN)) of the amyloid precursor protein (APP) processing complex,  $\gamma$ -secretase (Figure S11A). Aberrant processing of APP is a hallmark of the Alzheimer's disease pathological process. In addition to changes in cellular distribution of components of the  $\gamma$ -secretase complex we found alterations in NTRK2 and the multifunctional adapter protein  $\beta$ -arrestin 2 (ARRB2). We

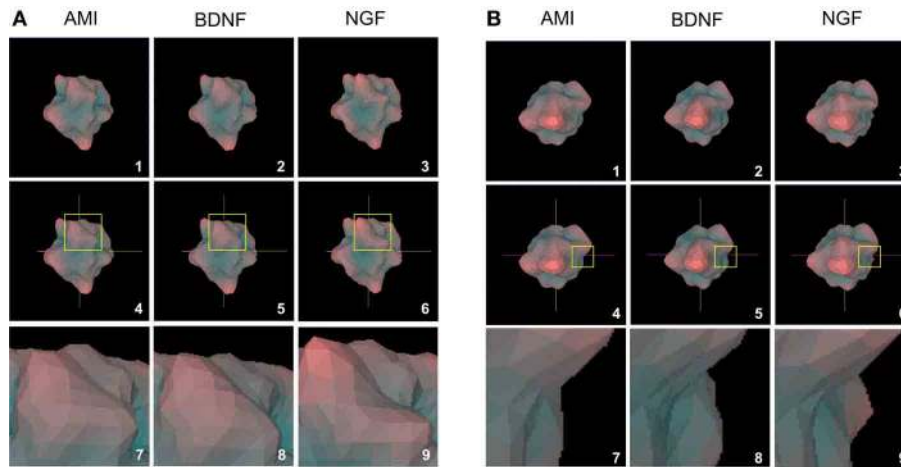


**FIGURE 6 | Amitriptyline demonstrated functional signaling similarities with brain-derived neurotrophic factor in human neuronal cells. (A)**

Venn diagram analysis of the gene transcripts significantly altered in their expression (compared to vehicle-treated control) in human SH-SY5Y cells 8 h after stimulation with amitriptyline (AMI-black), brain-derived neurotrophic factor (BDNF-red) and nerve growth factor (NGF-blue).

**(B)** Time course data for AMI (10 nM)-, BDNF (10 ng/mL)- or NGF (10 ng/mL)-mediated extracellular signal-regulated kinase 1/2 (ERK1/2) phosphorylation which denotes kinase activation (measured with p-ERK1/2 phosphospecific antibody). Statistical significance (Student's *t*-test) (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ) for AMI and BDNF compared to NGF

was measured at each time point from three independent experiments. **(C)** Time course data for AMI-, BDNF- or NGF-mediated v-akt murine thymoma viral oncogene homolog 1 (AKT1) phosphorylation which denotes kinase activation (measured with p-AKT1 phosphospecific antibody). Statistical significance (Student's *t*-test) (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ) for AMI and BDNF compared to NGF was measured at each time point from three independent experiments. **(D)** Treatment of SH-SY5Y cells (20 min) with AMI (10 nM) or BDNF (10 ng/ml) results in the increase of phosphotyrosine content [measured using anti-phosphotyrosine (p-Tyr) antibody—PY20] of immunoprecipitated (IP) NTRK2, while only NTRK1 is significantly phosphorylated by NGF.



**FIGURE 7 | Structural analogy between AMI and BDNF plurigons. (A)**

Panels 1, 2, and 3 depict representative plurigons for AMI-, BDNF- and NGF-stimulated transcriptomic datasets. Panels 4–6 depict the coordinate locations for plurigons in panels 1–3. In each plurigon window (4–6) a region of structural divergence between the plurigon structures in

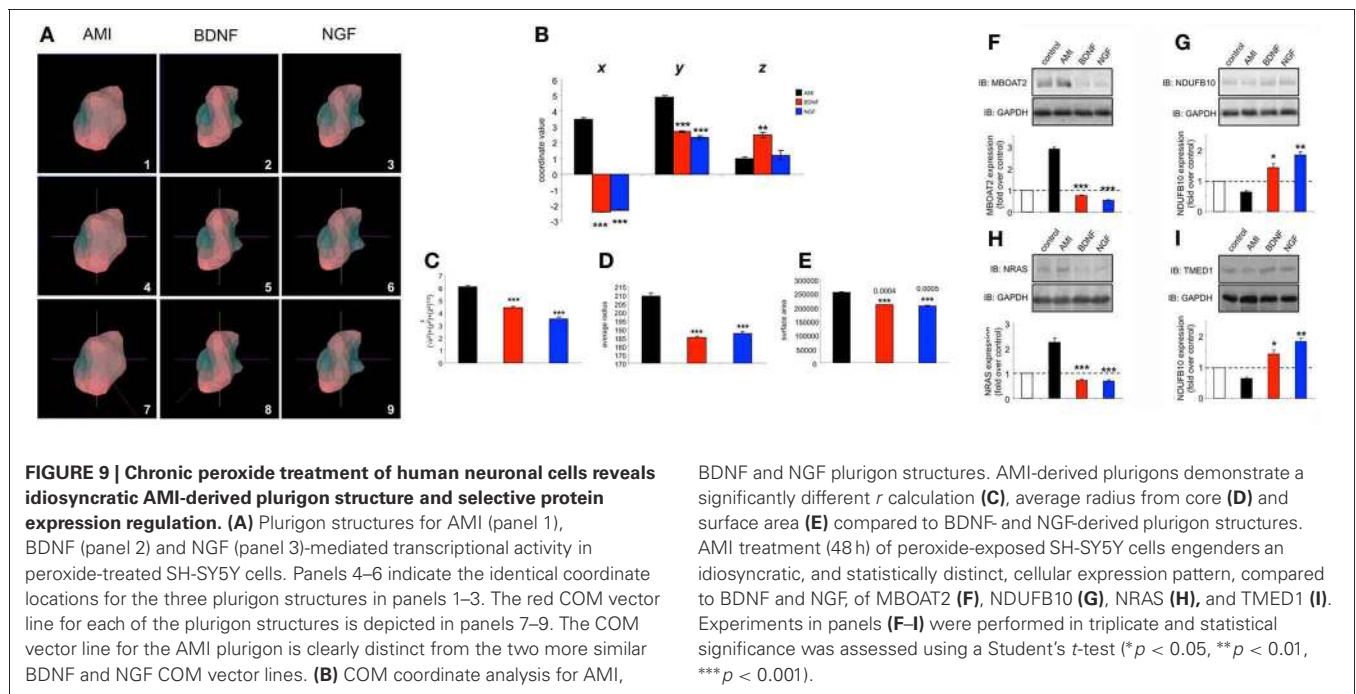
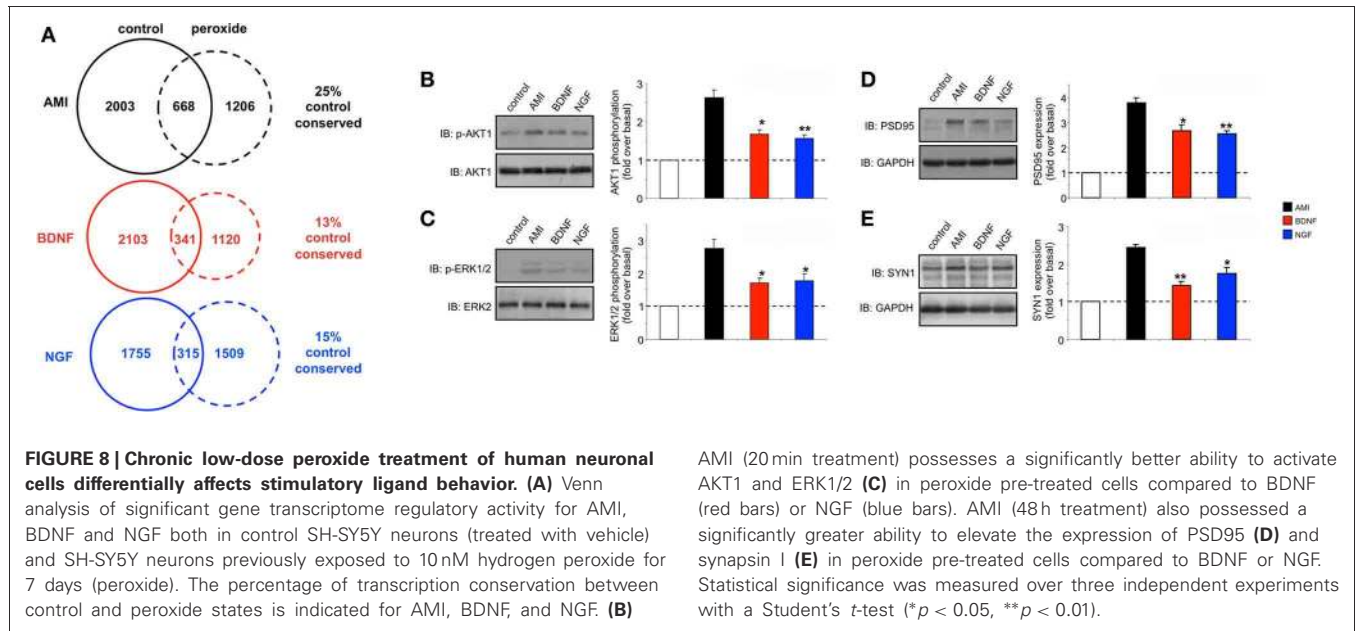
highlighted in a yellow box. This region is expanded (panels 7–9) to demonstrate the similarity in structural region between AMI and BDNF and their diversity from NGF. **(B)** Distinct secondary example of structural analogy between AMI and BDNF plurigons compared to NGF. The panel enumeration pattern is identical to **(A)**.

subsequently found that acute AMI treatment (20 min) affected the physical interaction between NTRK2 and APP, APH1A and ARRB2 (**Figure S11B**), suggesting that changes in this multiprotein complex may be associated with the beneficial therapeutic actions of AMI.

BDNF possesses pharmacological activities that are highly sought after for novel pharmacotherapeutics for aging-related neurodegenerative disorders such as Alzheimer's disease (Mattson et al., 2004; Nagahara and Tuszynski, 2011). However, many therapeutic strategies are not evaluated in appropriate pharmacogenomic settings, i.e., most neurotherapeutic activity is assessed in non-diseased, young healthy tissues or cells. The genomic and transcriptomic effects of aging and disease are considerable and can exert potent effects on drug efficacy (Lesko and Schmidt, 2012; Liou et al., 2012). We recently developed an *in vitro* mechanism to artificially “age” neuronal cells lines (Chadwick et al., 2010) as the extraction and analysis of central nervous tissues from aged experimental animals is highly problematical due to cellular stress and rapid degradation. We therefore applied this minimal peroxide exposure (7 days, 10 nM hydrogen peroxide treatment) to SH-SY5Y cells and repeated the AMI, BDNF and NGF transcriptomic stimulation protocols (**Table S18-AMI**, **Table S19-BDNF**, **Table S20-NGF**). Upon inspection of the transcriptomic responses in these “aged,” peroxide treated cells we found that in the “aged” cells AMI was able to activate a transcriptomic response more similar (25% conserved) to that in non-peroxide-treated control cells than either BDNF (13% conserved) or NGF (15% conserved) (**Figure 8A**). Perhaps linked to this we found that AMI possessed a superior ability, compared to BDNF or NGF (all 10 min stimulation), to activate neuroprotective (AKT1, **Figure 8B**) and neurodevelopmental (ERK, **Figure 8C**) signaling functions. In addition to these rapid signaling effects of AMI we found that with longer-term exposure (48 h), AMI demonstrated

an enhanced ability, compared to BDNF or NGF, to increase the expression of the neurosynaptic developmental markers (indicative of neurotrophin-like activity), post-synaptic density 95 (PSD95) and synapsin I (SYN1) (**Figures 8D,E**, respectively). To investigate whether this cellular context, and ligand-specific, alteration of pharmacological activity could be associated with plurigon structure we generated comparable plurigons from transcripts commonly activated by AMI, BDNF and NGF in these peroxide-treated “aged” cells (**Table S21-AMI**, **Table S22-BDNF**, **Table S23-NGF**) (**Figure 9A**). In the “aged” neuronal cells the AMI plurigon structures now demonstrated a more singular structure compared to those for BDNF or NGF (**Figures 9B–E**: COM coordinates,  $r$ , average radius, surface area). As we had previously shown that the actions of AMI appear to mimic the beneficial capacities of BDNF we applied bioinformatic annotation to the AMI-controlled transcripts more similarly regulated (ranked by expression score similarity) to BDNF as opposed to NGF (**Table S24**). Using the NIH DAVID Analytical Suite (<http://david.abcc.ncifcrf.gov/>) to generate KEGG (Kyoto Encyclopedia of Genes and Genomes (<http://www.genome.jp/kegg/>)) pathway output from these AMI-regulated, BDNF-like, transcripts, we found that the only signaling pathway significantly populated ( $p < 0.05$ ) by this gene list was the hsa047223: Neurotrophin signaling pathway ([http://www.genome.jp/kegg-bin/show\\_pathway?hsa047223](http://www.genome.jp/kegg-bin/show_pathway?hsa047223)) (**Table S25**).

We assessed, at the protein level, these cell context-specific actions of AMI, by validating several transcript protein products of genes differentially regulated by all three ligands that showed a preference for BDNF-like activity (**Table S24**). We found that 48 h after exposure, AMI was able to cause highly selective and unique alterations in the expression of MBOAT2 (membrane bound O-acyltransferase domain containing 2: **Figure 9F**), NDUFB10 (NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 10: **Figure 9G**), NRAS (neuroblastoma RAS viral (v-ras) oncogene



homolog: **Figure 9H**) and TMED1 (transmembrane emp24 protein transport domain containing 1: **Figure 9I**). Therefore, with respect to the actions of AMI in the “aged” neuronal tissue, we have seen that a strong alteration in its plurigon structure (**Figure 9**) is associated with selective signaling (**Figure 8**) and protein expression changes (**Figure 9**) while still retaining a strongly pro-neurotrophic signaling activity (**Table S25**). Therefore, *Plurigon*, in association with classical protein biochemistry and informatic techniques is able to assist in the identification of activity patterns of “small molecule” compounds that

possess a strong neuroprotective activity even in aged/damaged neuronal tissue.

### CONCLUSIONS

Given the increasing use of high-dimensionality data in many disciplines, including specifically biomedical research, the development of our application *Plurigon* serves as a mechanism to assist human-aided feature extraction. *Plurigon* may also be capable of bringing a larger audience to high-dimensionality data and represents a mechanism to open doors for integrated solutions

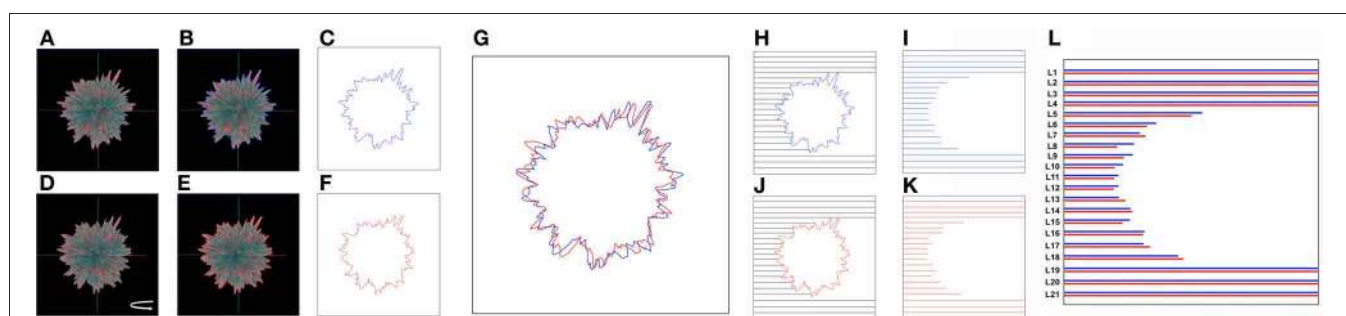


to biological and biomedical problems. *Plurigon* also serves as a potential alternative to other standard forms of complex data analysis such as PCA and also as a complementary system to other excellent high-dimensionality suites such as Ggobi and Ayasdi Iris. One specific feature of *Plurigon* that may assist in comparative studies is its strong standardized structural platform. This is in contrast to the more “freeform” nature of visualizers such as Iris. A strong expected structural platform may facilitate simpler comparisons between minimally-distinct datasets, but in other circumstances it is also possible that a more freeform topological data interpretation may be superior. The nature of the data corpus, the personal preferences of the user and the required form of answers from the data should all combine to influence the choice of application used. Due to its ability to condense multivariate data into a visually interpretable form, *Plurigon* can lay the groundwork for novel methods, experimental design, and new discoveries in a variety of scientific fields ranging from molecular biology to computational linguistics, genomics to proteomics, bioinformatics to pharmacology.

Here we have described the generation and implementation, in three varied but synergistic biomedical work paradigms, i.e., hypothalamic age patterns, genotypic analysis of age-modifying proteins such as *GIT2* (Chadwick et al., 2012) and the pharmacogenomic investigation of small molecule neurotrophic ligands. Our current appreciation of the full gamut of *Plurigons'* utility will no doubt be expanded once the application is transferred to a diverse range of investigators in different disciplines. Therefore, it is highly likely, and desirable, that further significant advancements in analysis, classification and feature extraction will be made in the future. The relative simplicity of the *Plurigon* concept underlies the enormous potential for comparative data analysis.

For example we propose that outline-transition analysis, i.e., the effect of feature “*crypticism*” (obscuration or revelation of outline features based on differential visual aspects and scale: **Figure 10**) in the plurigon may be an interesting avenue for further feature extraction. Another example for potential future development for *Plurigon* data outputs is the synergism of physical *plurigon* structures with TUIs. Hence, in addition to the standard *Plurigon* output file formats, the derived.vrml or.wrl files can be employed for physical 3-dimensional rendering/printing of a plurigon of choice (**Figure 11**). The generation of an actual physical output of *Plurigon*, with 3-dimensional printing using a Z Corp ZPrinter 650 (<http://www.sculpteo.com/en/>) may allow future users to easily connect their high-dimensionality data to machine-learning applications using TUIs (Ishii et al., 2012) that can extract more physically-relevant object data outputs.

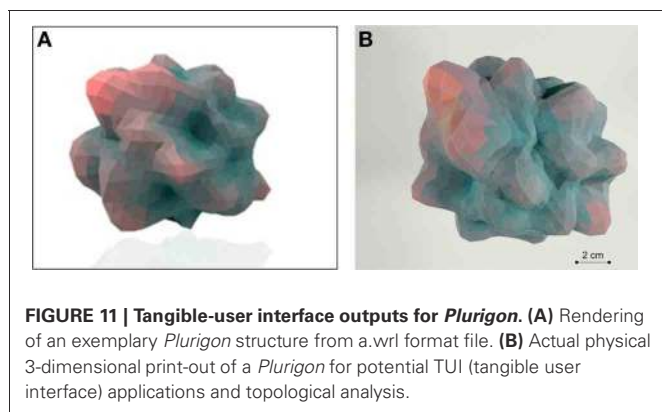
When data dimensionality reaches a certain threshold, it becomes challenging even for advanced computers to process such data. Aside from the sheer infeasibility of computation, problems of noise, sparseness and statistical insignificance increase dramatically as the number of dimensions increases. To this end, a common practice is to preprocess data with feature extraction and dimensionality reduction techniques such as PCA, Latent Semantic Analysis (LSA), or Semidefinite Embedding (SDE). A collective group of different feature extraction methods can be effective for the classification of a broad spectrum of entities (Orlov et al., 2008). Extraction is often performed after various mathematical transforms have been applied to the original data set to increase the number of total usable features. As such, feature extraction from *Plurigon* may therefore potentially be useful as additional input features for applied machine learning algorithms. One potential advantage that *Plurigon* possesses is the ability for humans to easily participate in the genesis of diverse



**FIGURE 10 | Future developments for *Plurigon* feature output and analysis.**

Outline recognition of cryptic feature generation. As *Plurigon* generates a reliable-oriented three-dimensional figure it is likely that surface features may indeed have an impact upon the appreciable visual nature of the plurigon through obscuring other data points. Such an effect therefore generates an entirely new level of data that other applications may not due to their free-form structure or lack of three-dimensional rendering. To attempt to quantify this effect, and also create additional extractable data features for detailed data classification, we propose the following methodology. **(A)** A single exemplary plurigon is chosen and for a given set of coordinates for the axes an outline of the edge of plurigon can be generated, indicated by the presence of the blue line in panel **(B)**. This trace can then be extracted in **(C)**. With the same exemplary plurigon as in **(A)**, slight rotation around the vertical  $y$  axis **(D)** helps generate a novel

plurigon outline **(E—red line)** that also can be extracted **(F)**. Even with small rotations (in axis required) the plurigon outline can clearly change **(G—superimposition of (C) and (F) traces)**, due to both loss of features from the visual field or also via obscuration of newly-oriented features with respect to the plane of the viewer. To extract quantifiable features from such effects line-length scanning can be used from any direction required in the extracted trace box. In this case scanning line-length is determined by the termination of the scan line (black) at the blue **(H)** or red **(I)** plurigon trace outline. Extraction and matching of these scan lines **(J,K)**; compared with minor vertical shift to improve comparison in **(L)**) can be used to discriminate divergent features of the plurigon based on this perspective-regulated analysis. The number of numerical features such as this can be easily scaled up or down based on the plurigon rotation angles and the orientation and density of the trace outline scan lines.



feature extraction processes from familiar and potentially tangible sources. With efficient data visualization, a coordination of efforts can be made for the efficient extraction of features either directly from the *Plurigon* structure, or from the original data set itself.

One potential important venue of future research for validating the use of “data texturizers” such as *Plurigon*, or the excellent Ayasdi Iris (Lum et al., 2013), as tools for visualization and classification is the field of drug screening and discovery. In our current work we have attempted to demonstrate that transitional changes in plurigon structures are tightly associated with complex changes in gene expression patterns, cellular signaling activity and protein-protein interaction. In addition to this we found that pharmacogenomic alterations in ligand (AMI) activity were mirrored in the subtle changes in plurigon structure (Figure 9) and that these changes were strongly associated with the maintenance of desirable drug activity patterns (Table S25). In addition to aiding pharmacological research by interpreting functional signaling patterns, the implementation of *Plurigon* to disease diagnoses and classification may be an important future use. Many disease processes, e.g., hypertension, Alzheimer’s disease or diabetes (Maudsley et al., 2011a), are often considered in a monolithic sense. While the eventual diagnosis, e.g., elevated blood pressure, compromised memory function, or excess glucose in the urine, may seem to unify patients presenting those data, it has been steadily proven over recent decades, assisted by assay multiplexing and high-content analytical tools, that these disorders can be fractionated into multiple distinct subtypes (Maudsley et al., 2011a; Israel et al., 2012; Kota et al., 2012; Zeller et al., 2012). Therefore, rather than considering them singular entities they are actually the output function of hyper-complex molecular signatures. As the ability to gather genome- or proteome-wide levels of patient data becomes a reality, our need to analyze and classify disease sub-groups becomes ever more important. The ability to separate these complex disease phenotypes will significantly enhance our ability to differentially and specifically diagnose and treat these various “sub-conditions” more effectively compared to monolithic treatment strategies. *Plurigon* therefore may be able to assist in generating a readily comparable platform for patient data classification and disease spectrum analysis.

With the increasing prevalence of the Semantic Web (<http://www.w3.org/standards/semanticweb/>) and drug-target databases such as PubChem, Linked Open Drug Data (LODD), and Chem2Bio2RDF, data mining for novel drug discovery is particularly promising (Shadbolt et al., 2006; Chen et al., 2010; Samwald et al., 2011; Wang et al., 2011). Using mass analytical techniques such as genomics and protein mass spectrometry, the generation of high-dimensionality data is now routine in biomedical science. Currently, a significant degree of informatics development has taken place for the compilation of multivariate drug-target response data in high-throughput assays, standardized binary representation of compounds (PubChem fingerprints), and rudimentary feature extraction on compounds (atom pair similarities) (Cao et al., 2008). Therefore, at the present time, with the creation of high-dimensionality data analyzers, such as *Plurigon*, Ggobi and Iris following the widespread implementation of mass data collection pipelines (genome/proteome-wide), we may soon see the generation of highly predictive and nuanced structure-activity-relationships (SAR) specifically modeled against pathophysiological scenarios. We have already pioneered this approach, in a considerably low-scale manner, for the discovery and whole-animal SAR analysis of palliative agents for Huntington’s disease (Martin et al., 2012b). In this study we were able to predict whole-animal therapeutic efficacy (using multiple biomedical indices) from complementary pathophysiological-pharmacogenomic-pattern investigation. In the future therefore, small molecular nuances may be able to be linked to subtle plurigon deformations induced by subtle “systems-level” transcriptional or proteomic responses in the experimental animal or even in clinical patients.

It is important to note that data texturizers such as *Plurigon* should rationally be employed, in a user-defined manner, alongside, and not instead of, conventional data analysis methods such as PCA, LSA, and SDE. Some of these techniques though do possess inherent deficiencies, such as PCA’s inability to handle complex non-linearity, while in contrast many groups are beginning to appreciate the tremendous categorizing power of the human eye linked to our excellent shape-recognition capacities. We have shown with our present data that there may indeed be circumstances where novel applications such as *Plurigon* can assist where the implementation of PCA is not ideal. However, the specific use of one tool or application over another is entirely dependent on a combination of user preference and suitability of the datasets in use. With potential future user-created *Plurigon* developments additional and novel extracted features may serve as useful discriminatory identifiers in cases where traditional features may not.

## ACKNOWLEDGMENTS

This work was supported by the Intramural Research Program of the National Institute on Aging, National Institutes of Health. We would also like to thank Doug Hansen for photographic assistance and Kevin G. Becker, William H. Wood 3rd, Elin Lehrmann and Yongqing Zhang for their assistance with microarray data collection.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: [http://www.frontiersin.org/Systems\\_Biology/10.3389/fphys.2013.00190/abstract](http://www.frontiersin.org/Systems_Biology/10.3389/fphys.2013.00190/abstract)

**Figure S1 | Website platform for Plurigon application home.** The Plurigon application is available in Windows-PC, Mac OSX and Linux formats.

**Figure S2 | Flowchart diagram for the generation and operation of the Plurigon application.**

**Figure S3 | Structural diversity between mean WT and mean HET plurigon structures.** Panels 1 and 2 depict mean WT and HET plurigons, respectively. Panels 3 and 4 demonstrate the identical coordinate location of the plurigons in panels 1 and 2. The yellow inset box highlights the region of structural diversity between WT and HET. Panels 5 and 6 depict expanded views of the highlighted box in panels 3 and 4. Yellow arrows indicate the regions of structural diversity.

**Figure S4 | Structural diversity between mean WT and mean HET plurigon structures.** Panels 1 and 2 depict mean WT and HET plurigons, respectively. Panels 3 and 4 demonstrate the identical coordinate location of the plurigons in panels 1 and 2. The yellow inset box highlights the region of structural diversity between WT and HET. Panels 5 and 6 depict expanded views of the highlighted box in panels 3 and 4. Yellow arrows indicate the regions of structural diversity.

**Figure S5 | Structural diversity between mean WT and mean HET plurigon structures.** Panels 1 and 2 depict mean WT and HET plurigons, respectively. Panels 3 and 4 demonstrate the identical coordinate location of the plurigons in panels 1 and 2. The yellow inset box highlights the region of structural diversity between WT and HET. Panels 5 and 6 depict expanded views of the highlighted box in panels 3 and 4. Yellow arrows indicate the regions of structural diversity.

**Figure S6 | Structural diversity between mean WT and mean HET plurigon structures.** Panels 1 and 2 depict mean WT and HET plurigons, respectively. Panels 3 and 4 demonstrate the identical coordinate location of the plurigons in panels 1 and 2. The yellow inset box highlights the region of structural diversity between WT and HET. Panels 5 and 6 depict expanded views of the highlighted box in panels 3 and 4. Yellow arrows indicate the regions of structural diversity.

**Figure S7 | Structural diversity between mean WT and mean HET plurigon structures.** Panels 1 and 2 depict mean WT and HET plurigons, respectively. Panels 3 and 4 demonstrate the identical coordinate location of the plurigons in panels 1 and 2. The yellow inset box highlights the region of structural diversity between WT and HET. Panels 5 and 6 depict expanded views of the highlighted box in panels 3 and 4. Yellow arrows indicate the regions of structural diversity.

**Figure S8 | PCA-mediated alignment of WT and HET murine multi-tissue transcriptomic data.** The three-dimensional PCA plot depicts the data points provided in **Table S10**. Red plus signs indicate WT data and the blue crosses indicate HET data.

**Figure S9 | Structural diversity between AMI, BDNF and NGF plurigon structures.** Panels 1, 2 and 3 depict AMI-, BDNF- and NGF-derived plurigons, respectively. Panels 4, 5 and 6 demonstrate the identical coordinate location of the plurigons in panels 1–3. The yellow inset box highlights the region of structural diversity between the three plurigons.

Panels 7–9 depict expanded views of the highlighted box containing the feature demonstrating similarity between AMI and BDNF compared to NGF in panels 4–6.

**Figure S10 | Structural diversity between AMI, BDNF and NGF plurigon structures.** Panels 1, 2 and 3 depict AMI-, BDNF- and NGF-derived plurigons, respectively. Panels 4, 5 and 6 demonstrate the identical coordinate location of the plurigons in panels 1–3. The yellow inset box highlights the region of structural diversity between the three plurigons. Panels 7–9 depict expanded views of the highlighted box containing the feature demonstrating similarity between AMI and BDNF compared to NGF in panels 4–6.

**Figure S11 | AMI-mediated subcellular expression and protein-protein interaction changes in human neuronal cells. (A)** Subcellular fractionation (fraction lanes 1–9: Golgi region = VAMP4-enriched lanes 4–8; endoplasmic reticulum region = calreticulin-enriched lanes 1–5) of human SH-SY5Y cells treated with AMI for 48h reveals alterations in expression patterns of the amyloid precursor protein (APP) processing factors nicastrin (NCSTN) and anterior pharynx defective 1 homolog A (APH1A), the NTRK2 receptor and the multifunctional adaptor protein  $\beta$ -arrestin 2 (ARRB2). **(B)** AMI treatment (10 nM) of SH-SY5Y cells results in the altered co-immunoprecipitation of APP (reduced), APH1A (increased) and ARRB2 (increased) with the NTRK2 receptor (representative western blots and associated histograms below). Experiments in panel **(B)** were performed in triplicate and statistical significance was assessed using a Student's *t*-test (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ).

**Table S1 | Primary antibody sources and employed dilution factors for selective western blotting.**

**Table S2 | Conserved hypothalamic transcriptomic expression data from 6 month old female mice.** The transcriptomic data from separate (F6-1, F6-2, F6-3) 6 month old mice was compared to 3 month old gender-matched control hypothalamic expression values.

**Table S3 | Conserved hypothalamic transcriptomic expression data from 12 month old female mice.** The transcriptomic data from separate (F12-1, F12-2, F12-3) 12 month old mice was compared to 3 month old gender-matched control hypothalamic expression values.

**Table S4 | Conserved hypothalamic transcriptomic expression data from 18 month old female mice.** The transcriptomic data from separate (F18-1, F18-2, F18-3) 18 month old mice was compared to 3 month old gender-matched control hypothalamic expression values.

**Table S5 | Conserved hypothalamic transcriptomic expression data from 6 month old male mice.** The transcriptomic data from separate (M6-1, M6-2, M6-3) 6 month old mice was compared to 3 month old gender-matched control hypothalamic expression values.

**Table S6 | Conserved hypothalamic transcriptomic expression data from 12 month old male mice.** The transcriptomic data from separate (M12-1, M12-2, M12-3) 12 month old mice was compared to 3 month old gender-matched control hypothalamic expression values.

**Table S7 | Conserved hypothalamic transcriptomic expression data from 18 month old male mice.** The transcriptomic data from separate (M18-1, M18-2, M18-3) 18 month old mice was compared to 3 month old gender-matched control hypothalamic expression values.



**Table S8 | Wildtype (WT) murine transcriptomic data from multiple isolated tissues.** Datasets are arranged alphanumerically with the following nomenclature: WT1 to WT16.

**Table S9 | GIT2 heterozygous (HET) murine transcriptomic data from multiple isolated tissues.** Datasets are arranged alphanumerically with the following nomenclature: HET1 to HET19.

**Table S10 | Standard PCA analysis of WT and HET genotypic datasets.** PCA was applied to reduce data dimensionality to three primary dimensions, D1, D2 and D3. Using these three dimensions the two genotype groups appeared non-separable.

**Table S11 | Plurigon structural concavity/convexity analysis.** The structural nature of three coincident facets of each dataset plurigon are color coded (red, green or blue) and their physical nature (convex (VEX) or concave (CAVE)) is denoted.

**Table S12 | Human neuronal cell (SH-SY5Y) transcriptomic effects induced by amitriptyline exposure.** For each amitriptyline (AMI) regulated transcript the official Gene Symbol, z-ratio and transcript description are indicated.

**Table S13 | Human neuronal cell (SH-SY5Y) transcriptomic effects induced by brain-derived neurotrophic factor exposure.** For each brain-derived neurotrophic factor (BDNF) regulated transcript the official Gene Symbol, z-ratio and transcript description are indicated.

**Table S14 | Human neuronal cell (SH-SY5Y) transcriptomic effects induced by nerve growth factor exposure.** For each nerve growth factor (NGF) regulated transcript the official Gene Symbol, z-ratio and transcript description are indicated.

**Table S15 | Conserved neuronal transcriptomic expression data from amitriptyline-treated clonal human cells.** Transcriptomic gene regulation was measured from amitriptyline (AMI)-treated SH-SY5Y neuronal cell batches (AMI-1, AMI-2, AMI-3). The transcriptomic data from the three separate experiments was compared to gene expression levels in vehicle (phosphate-buffered saline)-treated control SH-SY5Y cells.

**Table S16 | Conserved neuronal transcriptomic expression data from brain-derived neurotrophic factor-treated clonal human cells.** Transcriptomic gene regulation was measured from brain-derived neurotrophic factor (BDNF)-treated SH-SY5Y neuronal cell batches (BDNF-1, BDNF-2, BDNF-3). The transcriptomic data from the three separate experiments was compared to gene expression levels in vehicle (phosphate-buffered saline)-treated control SH-SY5Y cells.

**Table S17 | Conserved neuronal transcriptomic expression data from nerve growth factor-treated clonal human cells.** Transcriptomic gene regulation was measured from nerve growth factor (NGF)-treated SH-SY5Y neuronal cell batches (NGF-1, NGF-2, NGF-3). The transcriptomic data from the three separate experiments was compared to gene expression levels in vehicle (phosphate-buffered saline)-treated control SH-SY5Y cells.

**Table S18 | Transcriptomic effects of amitriptyline in human neuronal cells (SH-SY5Y) cells pre-treated with chronic minimal levels of hydrogen peroxide.** For each amitriptyline (AMI) regulated transcript in the peroxide-treated SH-SY5Y cells the official Gene Symbol, z-ratio and transcript description are indicated.

**Table S19 | Transcriptomic effects of brain-derived neurotrophic factor in human neuronal cells (SH-SY5Y) cells pre-treated with chronic minimal**

**levels of hydrogen peroxide.** For each brain-derived neurotrophic factor (BDNF) regulated transcript in the peroxide-treated SH-SY5Y cells the official Gene Symbol, z-ratio and transcript description are indicated.

**Table S20 | Transcriptomic effects of nerve growth factor in human neuronal cells (SH-SY5Y) cells pre-treated with chronic minimal levels of hydrogen peroxide.** For each nerve growth factor (NGF) regulated transcript in the peroxide-treated SH-SY5Y cells the official Gene Symbol, z-ratio and transcript description are indicated.

**Table S21 | Conserved neuronal transcriptomic expression data from pre-peroxidated amitriptyline-treated clonal human cells.** Transcriptomic gene regulation was measured from amitriptyline (AMI)-treated SH-SY5Y neuronal cell batches (AMI-1, AMI-2, AMI-3) that were pre-conditioned to mimic oxidative neurodegeneration with a continuous 7 day treatment with 10 nM hydrogen peroxide. The transcriptomic data from the three separate experiments was compared to gene expression levels in vehicle (phosphate-buffered saline)-treated control SH-SY5Y cells that were also treated with 10 nM hydrogen peroxide as well.

**Table S22 | Conserved neuronal transcriptomic expression data from pre-peroxidated brain-derived neurotrophic factor-treated clonal human cells.** Transcriptomic gene regulation was measured from brain-derived neurotrophic factor (BDNF)-treated SH-SY5Y neuronal cell batches (BDNF-1, BDNF-2, BDNF-3) that were pre-conditioned to mimic oxidative neurodegeneration with a continuous 7 day treatment with 10 nM hydrogen peroxide. The transcriptomic data from the three separate experiments was compared to gene expression levels in vehicle (phosphate-buffered saline)-treated control SH-SY5Y cells that were also treated with 10 nM hydrogen peroxide as well.

**Table S23 | Conserved neuronal transcriptomic expression data from pre-peroxidated nerve growth factor-treated clonal human cells.** Transcriptomic gene regulation was measured from nerve growth factor (NGF)-treated SH-SY5Y neuronal cell batches (NGF-1, NGF-2, NGF-3) that were pre-conditioned to mimic oxidative neurodegeneration with a continuous 7 day treatment with 10 nM hydrogen peroxide. The transcriptomic data from the three separate experiments was compared to gene expression levels in vehicle (phosphate-buffered saline)-treated control SH-SY5Y cells that were also treated with 10 nM hydrogen peroxide as well.

**Table S24 | Similarity score calculations for BDNF-like AMI-regulated transcripts.** A relative similarity score for AMI-regulated transcripts that possessed a regulatory mode more akin to BDNF than compared to NGF was calculated as the follows. The modulus of the ((AMI score-NGF score)/(AMI score-BDNF score)) was obtained and transcripts possessing a modulus of over 1 were considered to be more AMI-BDNF correlated than AMI-NGF correlated.

**Table S25 | KEGG pathway analysis for AMI-regulated, BDNF-like, transcripts in peroxide-treated human neuronal cells.** Using the NIH-DAVID Bioinformatic Suite (<http://david.abcc.ncifcrf.gov/>) we annotated the AMI-regulated, BDNF-like, transcripts stimulated by AMI treatment of artificially "aged" neuronal tissue. For each human signaling pathway the probability of geneset enrichment (*P*-Value), the enrichment relative to a scaled input from the human genome (Fold Enrichment) and the specific genes in the input list that populate ( $n \geq 2$ ) the assessed pathways. The only significantly populated signaling pathway was the hsa047223: neurotrophin signaling pathway.

## REFERENCES

- Amenta, N., Bern, M., and Eppstein, D. (1997). "Optimal point placement for mesh smoothing," in *8th ACM-SIAM Symposium on Discrete Algorithms* (New York, NY: ACM Press), 528–537.
- Barber, B. C., Dobkin, D. P., and Huhdanpaa, H. T. (1996). The Quickhull algorithm for convex hulls. *ACM Trans. Math. Softw.* 22, 469–483. doi: 10.1145/235815.235821
- Bellman, R. E. (1957). *Dynamic Programming*. Princeton, NJ: Princeton University Press.
- Briere, E., and George, P. L. (1995). Optimization of tetrahedral meshes. *IMA Vol. Math. Appl.* 75, 97–127.
- Buja, A., Lang, D. T., and Swayne, D. F. (2003). Ggobi: evolving from Xgobi into an extensible framework for interactive data visualization. *J. Comput. Stat. Data Anal.* 43, 423–444. doi: 10.1016/S0167-9473(02)00286-4
- Canann, S. A., Liu, Y. C., and Mobley, A. V. (1997). Automated 3D surface meshing to address today's industrial needs. *Finite Elem. Anal. Des.* 25, 185–198. doi: 10.1016/S0168-874X(96)00060-1
- Cao, Y., Charisi, A., Cheng, L. C., Jiang, T., and Girke, T. (2008). ChemmineR: a compound mining framework for R. *Bioinformatics* 24, 1733–1734. doi: 10.1093/bioinformatics/btn307
- Chadwick, W., Martin, B., Chapter, M. C., Park, S. S., Wang, L., Daimon, C. M., et al. (2012). GIT2 acts as a potential keystone protein in functional hypothalamic networks associated with age-related phenotypic changes in rats. *PLoS ONE* 7:e36975. doi: 10.1371/journal.pone.0036975
- Chadwick, W., Mitchell, N., Caroll, J., Zhou, Y., Park, S. S., Wang, L., et al. (2011). Amitriptyline-mediated cognitive enhancement in aged 3×Tg Alzheimer's disease mice is associated with neurogenesis and neurotrophic activity. *PLoS ONE* 6:e21660. doi: 10.1371/journal.pone.0021660
- Chadwick, W., Zhou, Y., Park, S. S., Wang, L., Mitchell, N., Stone, M. D., et al. (2010). Minimal peroxide exposure of neuronal cells induces multifaceted adaptive responses. *PLoS ONE* 5:e14352. doi: 10.1371/journal.pone.0014352
- Chen, B., Dong, X., Jiao, D., Wang, H., Zhu, Q., Ding, Y., et al. (2010). Chem2Bio2RDF: a semantic framework for linking and data mining chemogenomic and systems chemical biology data. *BMC Bioinformatics* 11:255. doi: 10.1186/1471-2105-11-255
- Cook, D., and Swayne, D. F. (2007). *Interactive and Dynamic Graphics for Data Analysis: With R and GGobi*. New York, NY: Springer-Verlag. doi: 10.1007/978-0-387-71762-3
- Cook, J. M. (1957). Technical notes and short papers: rational formulae for the production of a spherically symmetric probability distribution. *Math. Tables Aids Comput.* 11, 81–82.
- Delaunay, B. (1934). Sur la sphère vide, *Izvestia Akademii Nauk SSSR, Otdelenie. Matematicheskikh i Estestvennykh Nauk* 7, 793–800.
- Fayyad, U. M., Wierse, A., and Grinstein, G. G. (2002). *Information Visualization in Data Mining and Knowledge Discovery*. Burlington, MA: Morgan Kaufmann.
- Field, D. A. (1988). Laplacian smoothing and delaunay triangulations. *Commun. Appl. Numer. Methods* 4, 709–712. doi: 10.1002/cnm.1630040603
- Freund, R. M., and Orlin, J. B. (1985). On the complexity of four polyhedral set containment problems. *Math. Progr.* 33, 139–145. doi: 10.1007/BF01582241
- George, P. L., and Borouchaki, H. (1998). *Delaunay Triangulation and Meshing, Application to Finite Elements*. Paris: Edition Hermes.
- Gesty-Palmer, D., Yuan, L., Martin, B., Wood, W. H. 3rd., Lee, M. H., Janech, M. G. et al. (2013).  $\beta$ -arrestin-selective G protein-coupled receptor agonists engender unique biological efficacy *in vivo*. *Mol. Endocrinol.* 27, 296–314. doi: 10.1210/me.2012-1091
- Golan, M., Schreiber, G., and Avissar, S. (2009). Antidepressants, beta-arrestins and GRKs: from regulation of signal desensitization to intracellular multifunctional adaptor functions. *Curr. Pharm. Des.* 15, 1699–1708. doi: 10.2174/138161209788168038
- Graham, R. (1972). An efficient algorithm for determining the convex hull of a finite planar set. *Inf. Process. Lett.* 1, 132–133. doi: 10.1016/0020-0190(72)90045-2
- Herrmann, L. R. (1976). Laplacian-isoparametric grid generation scheme. *J. Eng. Mech. Div.* 102, 749–756.
- Hughes, G. F. (1968). On the mean accuracy of statistical pattern recognizers. *IEEE Trans. Inf. Theory* 14, 55–63. doi: 10.1109/TIT.1968.1054102
- Ishii, H., Lakatos, D., Bonanni, L., and Labrune, J. B. (2012). Radical atoms: beyond tangible bits, toward transformable materials. *Interactions* 19, 38–51. doi: 10.1145/2065327.2065337
- Israel, M. A., Yuan, S. H., Bardy, C., Reyna, S. M., Mu, Y., Herrera, C., et al. (2012). Probing sporadic and familial Alzheimer's disease using induced pluripotent stem cells. *Nature* 482, 216–220.
- Jarvis, R. A. (1973). On the identification of the convex hull of a finite set of points in the plane. *Inf. Process. Lett.* 2, 18–21. doi: 10.1016/0020-0190(73)90020-3
- Kim, H., Park, H., and Drake, B. L. (2007). Extracting unrecognized gene relationships from the biomedical literature via matrix factorizations. *BMC Bioinformatics* 8:S6. doi: 10.1186/1471-2105-8-S9-S6
- Ko, M. H., and Puglielli, L. (2009). Two endoplasmic reticulum (ER)/ER Golgi intermediate compartment-based lysine acetyltransferases post-translationally regulate BACE1 levels. *J. Biol. Chem.* 284, 2482–2492. doi: 10.1074/jbc.M804901200
- Kota, S. K., Meher, L. K., Jammula, S., Kota, S. K., and Modi, K. D. (2012). Genetics of type 2 diabetes mellitus and other specific types of diabetes; its role in treatment modalities. *Diabetes Metab. Syndr.* 6, 54–58. doi: 10.1016/j.dsx.2012.05.014
- Landauer, T. K., Laham, D., and Derr, M. (2004). From paragraph to graph: latent semantic analysis for information visualization. *Proc. Natl. Acad. Sci. U.S.A.* 101, 5214–5219. doi: 10.1073/pnas.0400341101
- Lee, D. T., and Schachter, B. J. (1980). Two algorithms for constructing a Delaunay triangulation. *Int. J. Parallel Progr.* 9, 219–242. doi: 10.1007/BF0097785
- Lesko, L. J., and Schmidt, S. (2012). Individualization of drug therapy: history, present state, and opportunities for the future. *Clin. Pharmacol. Ther.* 92, 458–466. doi: 10.1038/clpt.2012.113
- Liou, S. Y., Stringer, F., and Hirayama, M. (2012). The impact of pharmacogenomics research on drug development. *Drug Metab. Pharmacokinet.* 27, 2–8. doi: 10.2133/dmpk.DMPK-11-RV-093
- Lum, P. Y., Singh, G., Lehman, A., Ishkanov, T., Vajdem-Johansson, M., Alagappan, M., et al. (2013). Extracting insights from the shape of complex data using topology. *Sci. Rep.* 3:1236. doi: 10.1038/srep01236
- Luttrell, L. M., and Kenakin, T. P. (2011). Refining efficacy: allosterism and bias in G protein-coupled receptor signaling. *Methods Mol. Biol.* 756, 3–35. doi: 10.1007/978-1-61779-160-4\_1
- Martin, B., Brennenman, R., Golden, E., Walent, T., Becker, K. G., Prabhu, V. V., et al. (2009). Growth factor signals in neural cells: coherent patterns of interaction control multiple levels of molecular and phenotypic responses. *J. Biol. Chem.* 284, 2493–2511. doi: 10.1074/jbc.M804545200
- Martin, B., Chadwick, W., Yi, T., Park, S. S., Lu, D., Ni, B., et al. (2012a). VENNTURE—a novel Venn diagram investigational tool for multiple pharmacological dataset analysis. *PLoS ONE* 7:e36911. doi: 10.1371/journal.pone.0036911
- Martin, B., Chadwick, W., Cong, W. N., Pantaleo, N., Daimon, C. M., Golden, E. J., et al. (2012b). Euglycemic agent-mediated hypothalamic transcriptomic manipulation in the N171-82Q model of Huntington disease is related to their physiological efficacy. *J. Biol. Chem.* 287, 31766–31782. doi: 10.1074/jbc.M112.387316
- Mattson, M. P., Maudsley, S., and Martin, B. (2004). BDNF and 5-HT: a dynamic duo in age-related neuronal plasticity and neurodegenerative disorders. *Trends Neurosci.* 27, 589–594. doi: 10.1016/j.tins.2004.08.001
- Maudsley, S., Martin, B., and Egan, J. M. (2011a). To be or not to be—obese. *Endocrinology* 152, 3592–3596. doi: 10.1210/en.2011-1615
- Maudsley, S., Chadwick, W., Wang, L., Zhou, Y., Martin, B., and Park, S. S. (2011b). Bioinformatic approaches to metabolic pathways analysis. *Methods Mol. Biol.* 756, 99–130. doi: 10.1007/978-1-61779-160-4\_5
- Maudsley, S., Martin, B., and Luttrell, L. M. (2005). The origins of diversity and specificity in G protein-coupled receptor signaling. *J. Pharmacol. Exp. Therap.* 314, 485–494. doi: 10.1124/jpet.105.083121
- Maudsley, S., and Mattson, M. P. (2006). Protein twists and turns in Alzheimer disease. *Nat. Med.* 12, 392–393. doi: 10.1038/nm0406-392
- Maudsley, S., Pierce, K. L., Zamah, A. M., Miller, W. E., Ahn, S., Daaka, Y., et al. (2000). The beta(2)-adrenergic receptor mediates extracellular signal-regulated kinase activation via assembly of a multi-receptor complex with the epidermal growth factor receptor. *J. Biol. Chem.* 275, 9572–9580. doi: 10.1074/jbc.275.13.9572
- Mootha, V. K., Lindgren, C. M., Eriksson, K. F., Subramanian, A.,

- Sihag, S., Lehar, J., et al. (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* 34, 267–273. doi: 10.1038/ng1180
- Nagahara, A. H., and Tuszyński, M. H. (2011). Potential therapeutic uses of BDNF in neurological and psychiatric disorders. *Nat. Rev. Drug Dis.* 10, 209–219. doi: 10.1038/nrd3366
- Nicolau, M., Levine, A. J., and Carlsson, G. (2011). Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival. *Proc. Natl. Acad. Sci. U.S.A.* 108, 7265–7270. doi: 10.1073/pnas.1102826108
- Oddo, S., Caccamo, A., Shepherd, J. D., Murphy, M. P., Golde, T. E., Kaye, R., et al. (2003). Triple-transgenic model of Alzheimer's disease with plaques and tangles: intracellular Abeta and synaptic dysfunction. *Neuron* 39, 409–421. doi: 10.1016/S0896-6273(03)00434-3
- Orlov, N., Shamir, L., Macura, T., Johnston, J., Eckley, D. M., and Goldberg, I. G. (2008). WND-CHARM: multi-purpose image classification using compound image transforms. *Pattern Recognit. Lett.* 29, 1684–1693. doi: 10.1016/j.patrec.2008.04.013
- Rakhmanov, E. A., Saff, E. B., and Zhou, Y. M. (1995). "Electrons on the sphere," in *Computational Methods and Function Theory*, eds R. M. Ali, S. Ruscheweyh, and E. B. Saff (River Edge, NJ: World Scientific Publishing), 111–127.
- Ratti, C., Wang, Y., Ishii, H., Piper, B., and Frenchman, D. (2004). Tangible User Interfaces (TUIs): a novel paradigm for GIS. *Trans. GIS* 8, 407–421. doi: 10.1111/j.1467-9671.2004.00193.x
- Ruf, F., Fink, M. Y., and Sealfon, S. C. (2003). Structure of the GnRH receptor-stimulated signaling network: insights from genomics. *Front. Neuroendocrinol.* 24, 181–199. doi: 10.1016/S0091-3022(03)00027-X
- Saff, E. B., and Kujalaars, A. B. (1997). Distributing many points on the sphere. *Math. Intell.* 19, 5–11. doi: 10.1007/BF03024331
- Samwald, M., Jentzsch, A., Bouton, C., Kallese, C. S., Willighagen, E., Hajagos, J., et al. (2011). Linked open drug data for pharmaceutical research and development. *J. Cheminform.* 3:19. doi: 10.1186/1758-2946-3-19
- Schmalzigaug, R., Rodriguez, R. M., Phillips, L. E., Davidson, C. E., Wetsel, W. C., and Premont, R. T. (2009). Anxiety-like behaviors in mice lacking GIT2. *Neurosci. Lett.* 451, 156–161. doi: 10.1016/j.neulet.2008.12.034
- Shadbolt, N., Hall, W., and Berners-Lee, T. (2006). The semantic web revisited. *Intell. Syst. IEEE* 21, 96–101. doi: 10.1109/MIS.2006.62
- Smith, R. L. (1984). Efficient Monte Carlo procedures for generating points uniformly distributed over bounded regions. *Oper. Res.* 32, 1296–1308. doi: 10.1287/opre.32.6.1296
- Swayne, D. F., Cook, D., and Buja, A. (1998). Xgobi: interactive dynamic data visualization in the X window system. *J. Comput. Graph. Stat.* 7, 113–130.
- Theus, M. (2002). Interactive data visualization using Mondrian. *J. Stat. Softw.* 7, 1–9.
- Thomsen, K. (2007). *Generalized Spiral Points: A Slight Adjustment*. Ramsen: Sci Math.
- Wang, X., Liao, S., Nelson, E. R., Schmalzigaug, R., Spurney, R. F., Guilak, F., et al. (2012). The cytoskeletal regulatory scaffold protein GIT2 modulates mesenchymal stem cell differentiation and osteoblastogenesis. *Biochem. Biophys. Res. Commun.* 425, 407–412. doi: 10.1016/j.bbrc.2012.07.111
- Wang, Y., Xiao, J., Suzek, T. O., Zhang, J., Wang, J., Zhou, Z., et al. (2011). PubChem's Bioassay Database. *Nucleic Acids Res.* 40, D400–D412. doi: 10.1093/nar/gkr1132
- Wold, S., Esbensen, K., and Geladi, P. (1987). Principal components analysis. *Chemometr. Intell. Lab. Syst.* 2, 37–52. doi: 10.1016/0169-7439(87)80084-9
- Wurtele, E. S., Li, J., Diao, L., Zhang, H., Foster, C. M., Fatland, B., et al. (2003). MetNet: software to build and model the biogenetic lattice of Arabidopsis. *Comp. Funct. Genomics* 4, 239–245. doi: 10.1002/cfg.285
- Zeller, T., Blankenberg, S., and Diemert, P. (2012). Genomewide association studies in cardiovascular disease—an update 2011. *Clin. Chem.* 58, 92–103. doi: 10.1373/clinchem.2011.170431

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 08 February 2013; paper pending published: 19 March 2013; accepted: 01 July 2013; published online: 22 July 2013.

Citation: Martin B, Chen H, Daimon CM, Chadwick W, Siddiqui S and Maudsley S (2013) Plurigon: three dimensional visualization and classification of high-dimensionality data. *Front. Physiol.* 4:190. doi: 10.3389/fphys.2013.00190

This article was submitted to *Frontiers in Systems Biology*, a specialty of *Frontiers in Physiology*.

Copyright © 2013 Martin, Chen, Daimon, Chadwick, Siddiqui and Maudsley. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.