Enterprise Biology Software Project

PO Box 292
Medina, Washington 98039-0292

Enterprisebiology.com
playingcomplexitygames.com

ANNUAL REPORT

2016

Since biology as a science continues to struggle with the problem of reproducibility, many now believe that every effort should be made to increase the precision of our estimates. Although a worthwhile goal, others might argue that precision in science is a poor substitute for accuracy. To avoid taking sides, the report uses a reproducibility test to discover where the precision and accuracy of biology can exist together in our published data. The advantage of such an approach to problem solving is that it requires little more than figuring out where and how to look.

Enterprise Biology Software: XVII. Research (2016)

ROBERT P. BOLENDER

Enterprise Biology Software Project, P.O. Box 292, Medina, WA 98039-0292, USA
(enterprisebiology.com)

SUMMARY

Our current test for reproducibility in the life sciences relies largely on local precision, namely duplicating the results of a given experiment. In contrast, biology seems to need two rules to fulfill its reproducibility requirements. The size of a given part is under local rules (precision), whereas the relationship of one part to another follows global rules (accuracy). In this report, we will attempt to tap into these rules of biology with the goal of constructing a reproducibility test based on accuracy. Since any given species includes many similar copies, we already know that biology is very good at playing the accuracy game globally. This simplifies our task considerably by reducing it to an exercise in finding these global patterns in the biology literature. Moreover, such patterns would offer empirical evidence for the existence of design principles basic to reproducibility. To assemble a reproducibility test, published data from more than a thousand papers were translated into triplets and expressed as mathematical markers and connection ratios. When mathematical markers were plotted against their connection ratios and viewed graphically, global patterns consistent with both accuracy and precision appeared in surprisingly large numbers. These patterns were found in the MRI data (volumes) of a single species (human) and in stereological data (numbers and surfaces) encompassing many species. In short, all the data sets examined demonstrated the ability to pass a reproducibility test based on accuracy. Such results suggest that the biology literature contains data far more accurate than our traditional precision-centric tests are telling us. The test databases also detected the presence of global patterns within and between organelles, cells, species, exposures, and disorders. Although such an outcome was unanticipated, this is exactly the overarching pattern we should be finding throughout the biology literature. In addition to many animal species sharing considerable portions of their genomes, it would now appear that DNA sharing applies not only to genes coding for proteins, but also to the yet unidentified sequences or other devices coding for the design principles responsible for phenotypic patterns. If this proves to be the case, then prediction to and from DNA becomes a reasonable goal. Since only about 3% of our DNA is allocated to genes coding for proteins, it seems likely that the larger part of the DNA story (97%) will be about the roles being played by these intergenic sequences in designing and controlling phenotypes. The report also considers how this expanded view of DNA might influence the way we study the human brain. By identifying extensive populations of abnormal phenotypic patterns in brains displaying disorders, literature databases might speed the task of figuring out what parts of the DNA need repairing – genetic, intergenic, or both. Here our goal would be to create a quantitative link between the phenotype and its DNA, wherein extensive data interactions become an essential part of the problem-solving process. A final note. The reports and resources previously distributed on DVD to contributing authors will now be published online at playingcomplexitygames.com.

INTRODUCTION

Reproducibility is defined as an ability to duplicate the results of an experiment either by the same researcher or by an independent one. Defined as such, it appears to be more a measure of precision (repeating the same result), than of accuracy (correctness). Herein lies a problem. To be valid, a measurement system requires both precision and accuracy.

Reproducibility continues to be the Achille's heel of our scientific community (Collins and Tabak (2014), Begley and Ioannidis (2015), Freedman et. al (2015), Roth and Cox (2015), Engber (2016)). *Nature* (Baker, 2016), for example, recently used a questionnaire to highlight the consensus view that a reproducibility crisis currently exists in science. After reading *Nature's* questionnaire and viewing the recommendations of the *Biophysical Journal* (2015), however, one is led to believe that improving reproducibility will result in a higher quality of science. Since many of the improvements being suggested deal largely with precision, such a prognosis might be overly optimistic. Recall that a precise estimate can be either correct or incorrect. This is an indisputable fact.

Given this widely understood limitation of precision, why not base our definition of reproducibility on accuracy? But how can we expect to demonstrate accuracy with biological data when we're currently having such a hard time with precision? All we need are a few basic questions to get us started. Does reproducibility exist in biology? Yes, biology reproduces – very effectively – vast numbers of animals that we identify as belonging to distinct species. Moreover, reproducibility is fundamental to all parts of biology because to have emergent properties (e.g., life, thoughts, success, survival) it must connect these parts to form highly specific, predictable, and adaptable patterns. If true, then one can argue that biology must be taking the high road – the accuracy approach to reproducibility. Any proof? Yes. Is biology suffering from a reproducibility crisis? No.

Is it possible to assemble a robust test for reproducibility - based on a strategy similar to the one biology uses? What – exactly - is biology's strategy? When faced with a complex problem, biology can bring at least three key players into the game - parts, connections, and complexity. Can we do the same? What would it take?

To be convincing, we would need to make our reproducibility test much harder to pass than the one described in *Nature* (Baker, 2016). Instead of limiting the definition of reproducibility to that of a single study, we would have to expand it to include all applicable studies stored in our biology literature databases. Notice that by moving the definition from local to global, we automatically shift the focus of the test from precision to accuracy. Moreover, by replacing the simple variable (data point) with a complex one (triplet mathematical marker), six variables – not just one – must be duplicated not just once, but three or more times across the literature. Notice the strategy in play. By making the test impossible to pass unless biology allows it, we leave ourselves little choice but to accede to biology's definition of reproducibility, which includes accuracy. Besides, deferring to biology tends to inspire confidence. It already knows how to solve the most difficult problems.

METHODS AND RESULTS

The Enterprise Biology Package

In collaboration with the community, the Enterprise Biology Software Project transforms published research data into patterns capable of

addressing and solving a wide range of complex problems in biology (enterprisebiology.com). A yearly progress report includes databases, files, and directions for repeating the current results (Figure 1). This year, a worked example of the reproducibility test described herein is being distributed online (playingcomplexitygames.com). It introduces the reader to the process of solving complex problems with databases.

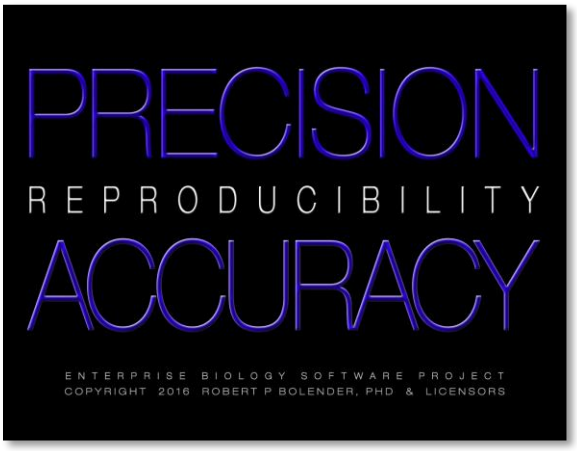

Figure 1 The package includes the yearly report accompanied by worked examples, databases, et cetera.

The Reproducibility Test

The reproducibility test requires a new data type called the connection ratio. Defined as an alphanumeric string, it consists of a triplet mathematical marker (AX:BY:CZ) with the names of its parts (A, B, C) all changed to the same name (Part): (PartX:PartY:PartZ). As such, only the values of the ratio (X:Y:Z) are in play. The renaming operation is done in Excel and can be duplicated by downloading one of the files described in Appendix I.

Table 1 summarizes the volume, number, and surface databases used in the report. The volume (V) data were derived from the Internet Brain Volume Database (IBVD) (Kennedy et.al 2012), whereas the number (N) and surface area (S) data came from the stereology literature database (Bolender, 2001-2016). Notice that the MRI data represented the most efficient source of global information (41.6%), whereas cell counts (N) the least (1.1%). At 5.7%, surface areas had an efficiency rating roughly five times greater than the one for number – even with its smaller sample size.

Table 1 The table summarizes the sample sizes used for the reproducibility tests. Note that the volume-based plot displayed in Figure 3 used only 2,500 of the 155,891 duplicate markers (≥3). Abbreviations include MRI (magnetic resonance imaging), LM (light microscopy), and TEM (transmission electron microscopy).

Data Types →	**Volume**	**Number**	**Surface**
Data Sources→	**MRI**	**LM**	**TEM**
Mathematical Markers	374,906	108,824	21,402
Duplicates (≥3)	155,891	1,236	1,221
Efficiency (Reproducibility)	41.6%	1.1%	5.7%

Global Data ⇒ Reproducibility ⇒ Accuracy

When working under the aegis of complexity theory, the focus invariably shifts from local to global. Both mathematical markers and connection ratios signal the presence of global data when the same string occurs repeatedly. For our purposes here, global data will serve as the measure of reproducibility.

Design of the Test: The reproducibility test uses the CommunityGraphPlot of Mathematica 11 (Wolfram) to display the relationship of mathematical markers to connection ratios (Figure 2). The results are expressed as a collection of units, wherein a measure of accuracy (a mathematical marker) is duplicated as a measure of reproducibility. In each unit, the number of duplicate markers is equal to the number of lines connecting a mathematical marker to its connection ratio. The strength of the test depends on the number of connecting lines (forming blue spindles) and the total number of units in

play. The number of different mathematical markers associated with a given connection ratio is a measure of the preference given to that ratio by biology. To pass the test, reproducibility must exist within and across many units.

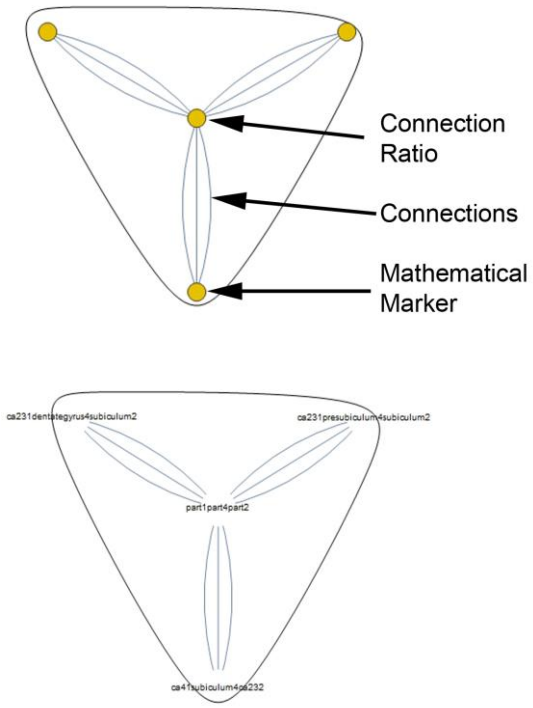

Figure 2 Reproducibility can be expressed quantitatively as a unit of complexity, one that displays a highly specific pattern. By plotting mathematical markers (peripheral dots) against their connection ratios (central dot), we can test for the presence of reproducibility by counting the number of lines connecting the dots (3 in this case). Moreover, the plot detects the number of different mathematical markers using the same connection ratio (3 identified). Since only one copy of a given marker is taken from a publication (or experiment), 3 or more (≥3) connections signal the presence of reproducibility – at the global level.

Volume: Since biological complexity remains intact in living subjects, the test was applied first to MRI data coming from patients. The test data were derived the Internet Brain Volume Database (IBVD) (Kennedy, et al, 2012).

Reproducibility Test (Volume): The reproducibility test generates what amounts to a fact pattern – a complex assembly of facts from which to draw conclusions. Recall that our primary goal here is not to duplicate the results of a given experiment, but rather to detect global patterns that will tell us something about the basic principles in play. Reproducibility is being treated as such a principle.

The MRI data set passed the reproducibility test easily. Notice in Figure 3 that the overall pattern consists of many, individual sets of facts (units), which in combination provide evidence for the presence of widespread reproducibility in the data set derived from the IBVD. The figure also shows that a single connection ratio provides the mathematical hub for several different mathematical markers and that biology favors a relatively small set of connection ratios, as shown by the frequencies of the lines in the blue spindles.

From where does the reproducibility come? It comes from biology's ability to maintain the ratios of its parts with a high degree of accuracy. Mathematical markers detect this accuracy by finding multiple copies of the same patterns distributed across the literature. Since >40% of the data set is global and reproducible (Table 1), the argument for the existence of accuracy becomes a compelling one.

The key points to take from Figure 3 include:

1) A given connection ratio (central point) can accommodate several mathematical markers (peripheral points) carrying similar or different scripts.
2) The plot serves as a reproducibility test based on specific, quantitative patterns.
3) Such a test can be applied to most – if not all – data types reported in the biology literature.

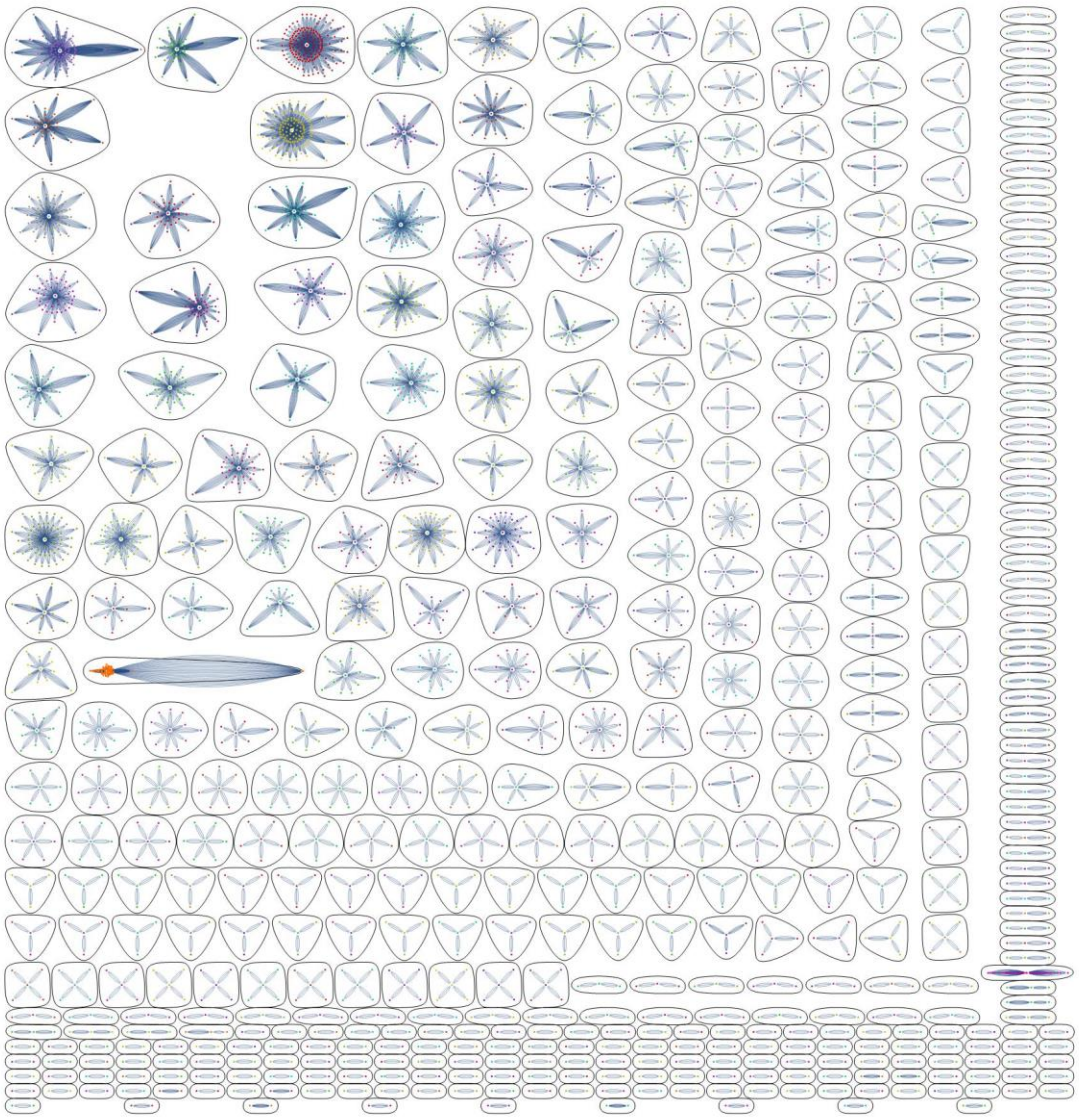

Figure 3 The MRI volume data of patients produce a wealth of global data as indicated by the many units and the high concentrations of connections (seen as dense blue spindles) linking mathematical markers to their connection ratios. The original data set was filtered to to select only those connections that occurred 3 or more times. Note that the figure uses just 2,500 of the 155,891 duplicate markers (≥3) – Table 1. Since the presence of global data demonstrates the presence of reproducibility, the MRI data passed the reproducibility test - very convincingly.

Number: Since the previous section with volumes showed that plotting connection ratios against their mathematical markers gives an effective test for reproducibility, can it also be applied to stereological estimates for cell numbers and membrane surface areas? Estimates for cell number based on the fractionator (Gunderson et al., 1988) can be expected to give reliable estimates because the method is volume independent, which means that it is free of the volume related biases associated with estimates for stereological densities. Moreover, forming ratios of numerical densities also minimizes the volume related biases by cancelling out the reference volumes.

Reproducibility Test (Number): When applied to the number data in the stereology literature

database, the test produced a pattern like the one displayed previously by volume (Figure 3). Mathematical markers once again congregated around a central connection ratio and displayed the multiple copies consistent with global data and reproducibility (Figure 4). In effect, the cell counts passed the reproducibility test. Notice, however, that relatively few mathematical markers (peripheral dots) have more than three duplicate copies.

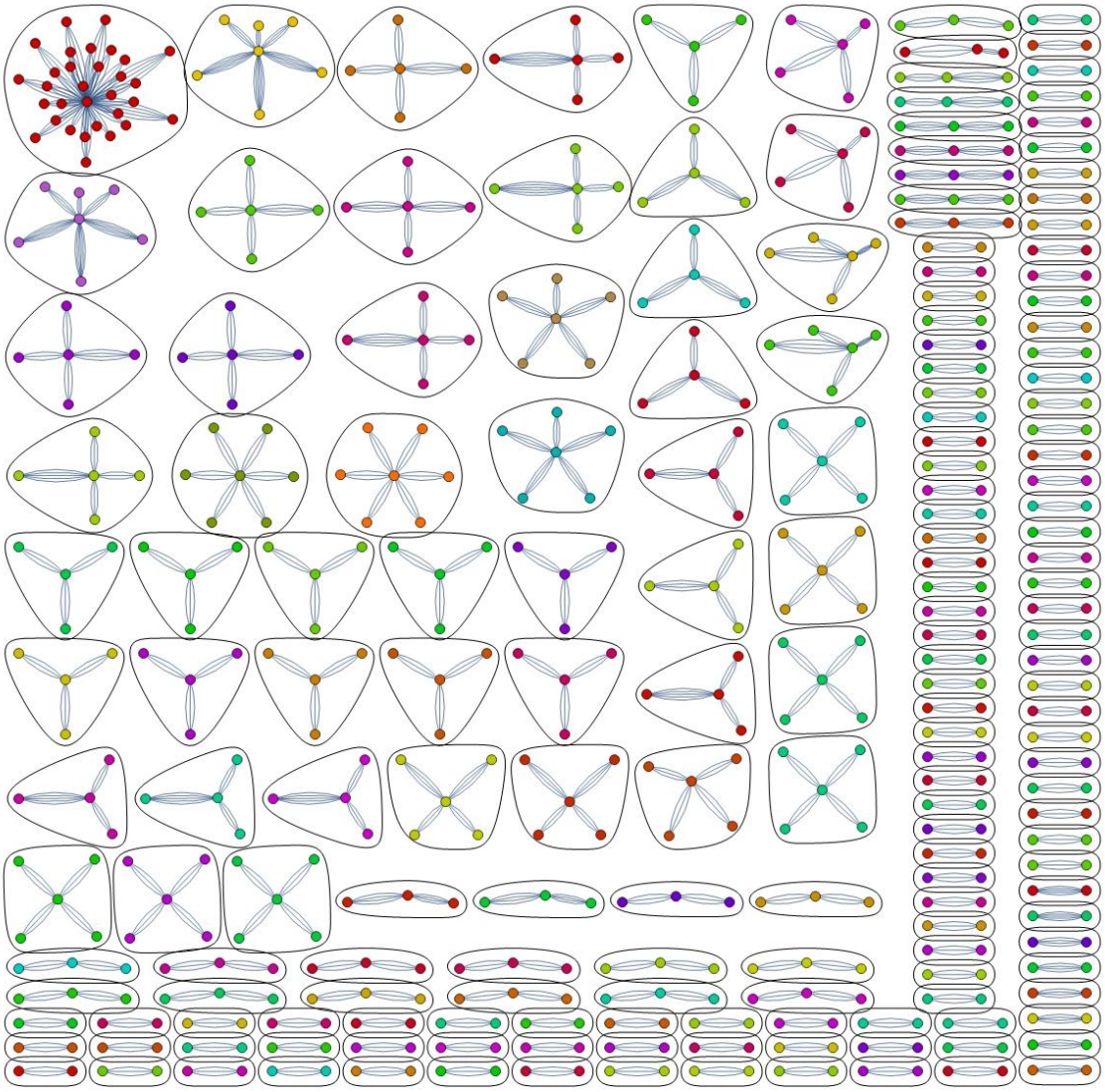

Figure 4 When the reproducibility test was applied to the the data set consisting of cell counts (numbers), the pattern consistent with the presence of global data and reproducibility appeared. Most of the connections, however, contained the minimum number of connections allowed – just 3. Curiously, only about 1% of the mathematical markers in the data set qualified for the test (Table 1) – the lowest score to date.

Surface: Surface areas estimated as surface densities (S/V) with stereological methods are volume dependent and as such subject to volume related distortions. Forming ratios, however, effectively minimizes such distortions mathematically by cancelling out the offending reference volumes. Consequently, surface areas should also provide global data comparable to that of the volume independent estimates for numbers (cell counts; Figure 4). We can test this assumption by running the reproducibility test.

Reproducibility Test for Surface Areas: When plotted as mathematical markers vs. connection ratios, the surface area data also passed the reproducibility test (Figure 5) – even more convincingly than the one for numbers (Figure 4). Notice that many more connections exist between the markers and the connection ratios for surfaces than for numbers (compare Figures 4 and 5). This occurred even though the number of mathematical markers available for the surface area test was only 20% of that used for the numbers and even less than that (6%) for volumes (Table 1).

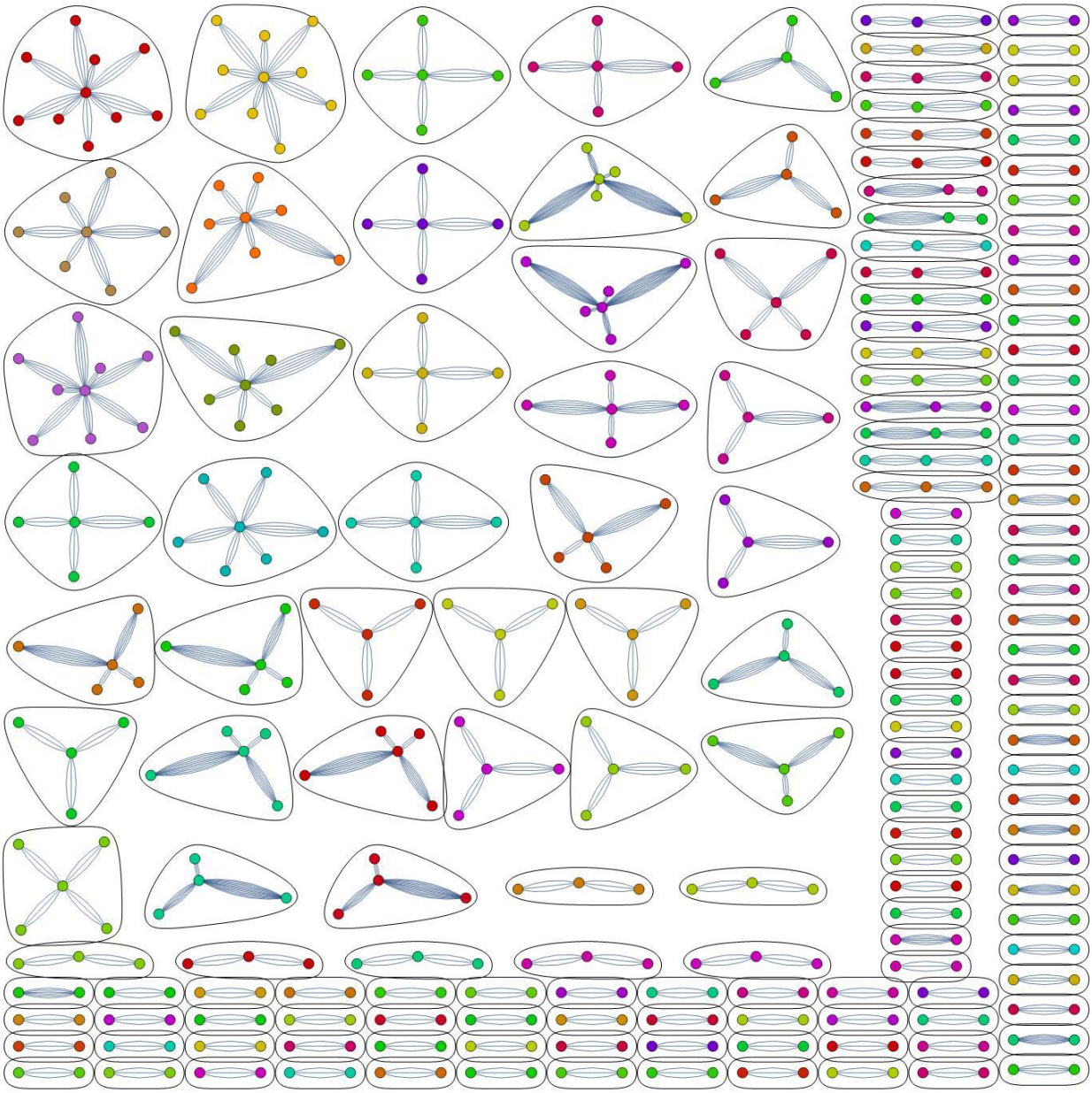

Figure 5 The surface area data (cell organelles) also passed the test, displaying many of the dark blue spindles seen earlier for the MRI volume data. In spite of its relatively smaller size, the surface areas displayed a reproducibility efficiency five times greater than that seen for the numbers.

The remainder of the report illustrates that the reproducibility test can assume many forms. If, for example, we plot the mathematical markers for cell organelles against their citation numbers, we get a global view of data connectivity – as it exists in the stereology literature database (Figure 6). In the figure, each box contains the citation number of a paper, as indexed in the database (now available online). The presence of the same mathematical marker in multiple publications demonstrates the global nature of the stereology literature. In effect, accuracy and reproducibility characterize both biology and the biology literature.

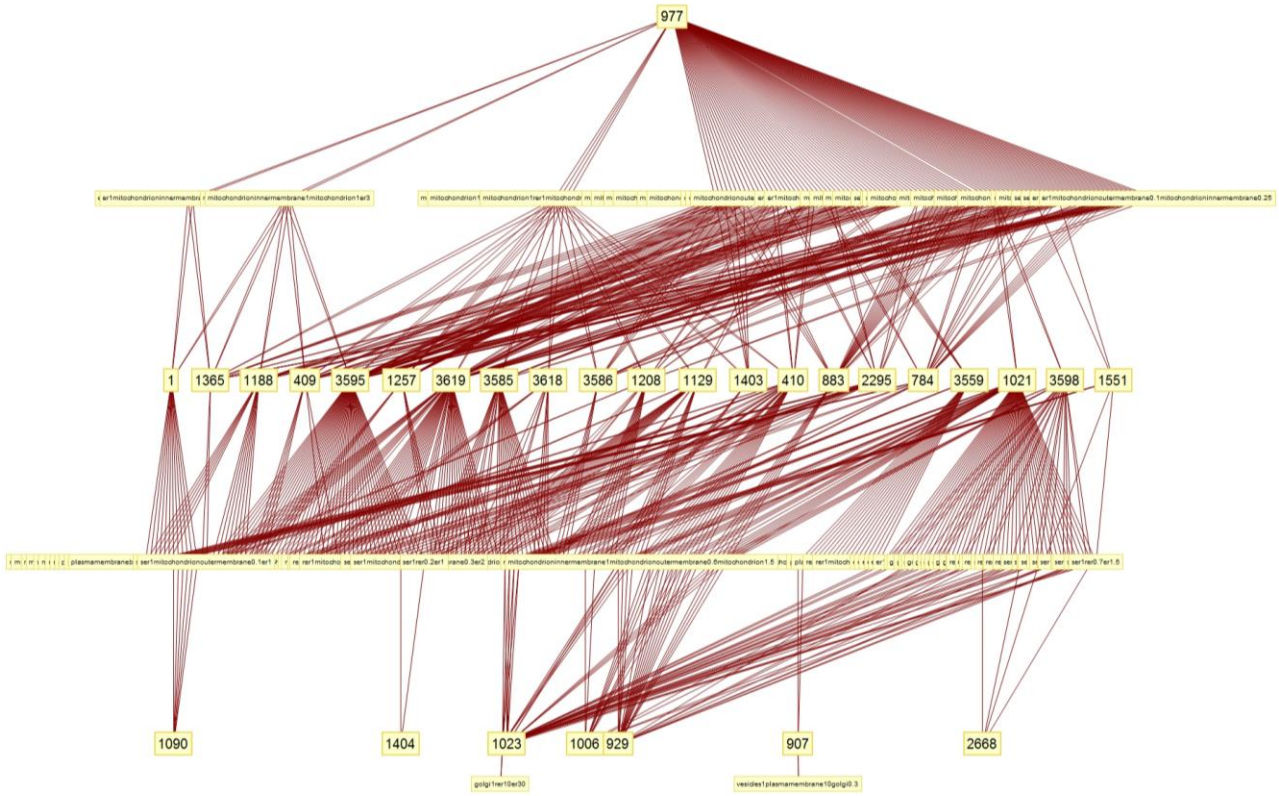

Figure 6 When surface areas of organelles coming from a wide variety of cell types and species (frogs to humans) are expressed as mathematical markers, related to their citation numbers (boxed), and plotted, global patterns in the stereology literature become apparent. Such patterns indicate that reproducibility and accuracy are widespread.

The Role of Connectivity

Why is connectivity so important to biology? It connects its parts by rule to form complexities, as a way of optimizing those outcomes that ensure its success and survival. Our current strategy consists of tapping into these biological complexities by connecting snippets of these rules into informative patterns. We can do this because recent advances in technology allow us to assemble and interpret large data sets and patterns.

Figures 3, 4, and 5 indicate that biology routinely favors some connection ratios over others, at least for volumes, numbers, and surfaces (see also Bolender, 2006, 2010). Recall that these figures indicate that a given connection ratio (central point) can accommodate both similar (multiple blue lines) and different mathematical markers (multiple peripheral points).

An immediate consequence of comparing connection ratios to mathematical markers is that it

raises fundamental questions about the relationship of parts to connections. Although we can be reasonably confident that DNA codes for both parts and connections, we don't know if they are being controlled together or separately. If, for example, biology is controlling its parts and connections separately, then disorders – or for that matter any type of phenotypic change – is subject to at least two levels of oversight. This means that disorders of the brain, for example, could be explained by abnormal parts, abnormal connections, or some combination of the two. Consequently, knowing what's in play becomes critical to interpreting results and pursuing solutions.

Seeing the Big Picture: Since both mathematical markers and their connection ratios represent universal data types, we can use them to summarize large and otherwise heterogeneous data sets. Figure 7, for example, shows that the combined data set for volumes, numbers, and surfaces – expressed as connection ratios – are highly interconnected. Such a pattern suggests that different data types are sharing similar rules of organization.

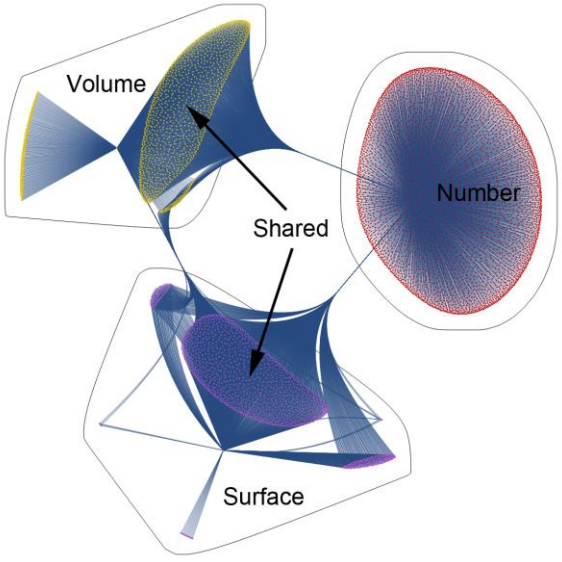

Figure 7 Although the volumes, surfaces, and numbers of parts represent distinctly different data types in terms of what they measure, their connection ratios show extensive connectivity. In effect, different data types can share the same rules.

If, instead, we plot the connection ratios of Figure 7 against species, the three data types (V, S, N) continue to display extensive connectivity (Figure 8). Note that the animals in the figure are grouped according to the similarity of their connections – an analytical feature of the CommunityGraphPlot (Mathematica 11).

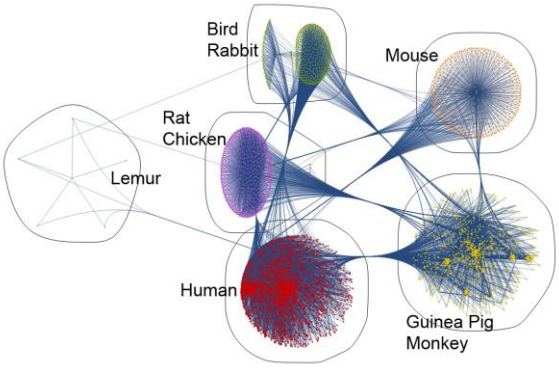

Figure 8 Even after the connection ratios of the volume, surface, and number databases were combined and plotted against species, the connectivity remained. Recall that these data came from parts ranging in size from organs to organelles, responding to a wide range of experimental conditions, and carrying different methodological biases. In effect, connectivity appears to be a major unifying principle in living systems.

If, for example, we look just at the surface areas of cell organelles distributed across the animal kingdom, we can see the extent to which connectivity operates as a fundamental principle of biological systems (Figure 9). It appears that all the species included in the plot are playing by the same set of connectivity rules.

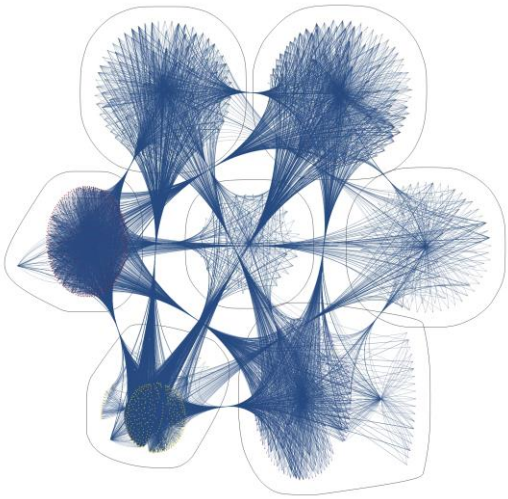

Figure 9 When organelle surface areas are plotted against connection ratios, they form seven clusters - all of which are highly interconnected.

We can also look at data from a single paper and compare the connectivity of mathematical markers to those of connection ratios. This tells us something about the relationship of one cell to another in same tissue. Consider, for example, hepatocytes, fat storing cells, endothelial cells, and Kupffer cells of the rat liver. When we plot the mathematical markers for the surface areas of cell organelles against the cell types, we find that all the cells share some of the same markers (Figure 10).

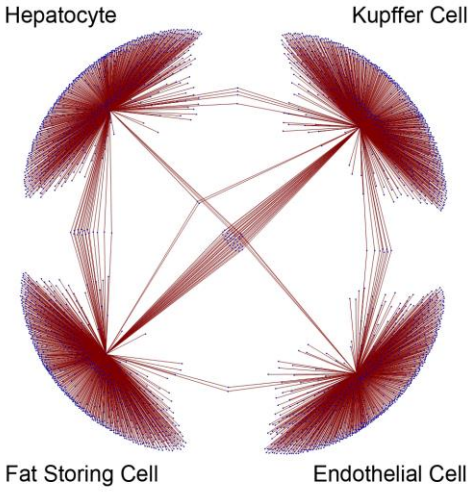

Figure 10 In the rat liver (Adapted from Blouin et al., 1977 [977]), hepatocytes, fat-storing cells, Kupffer cells, and endothelial cells display both unique and shared mathematical markers. Notice how each cell connects to the other three.

If, however, we replace the mathematical markers of Figure 10 with their equivalent connection ratios (Figure 11), the number of cell to cell connections increases dramatically. The figure suggests that the design of a given cell appears to include two distinct subpopulations of connection ratios – one shared and the other not shared. In effect, the genetic programming of cell organelles can be detected phenotypically by the presence (or absence) of specific mathematical markers and connection ratios. Such information may prove helpful in understanding the fundamentals of differentiation or in explaining how a cell can change its phenotype.

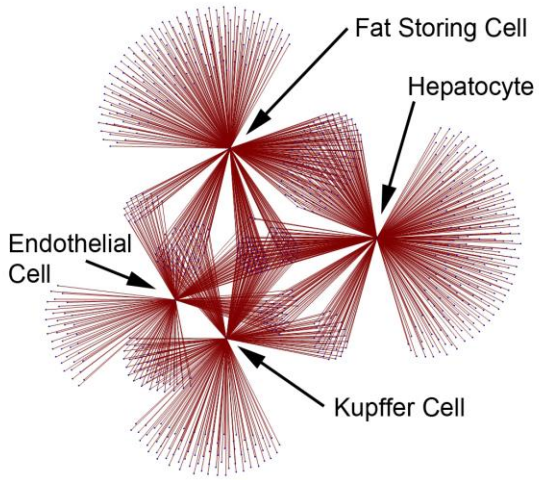

Figure 11 When the mathematical markers of Figure 10 are replaced by their connection ratios, connectivity between the cells increases markedly. Notice how tightly the cells are interconnected phenotypically.

When we replace the dots shown in Figure 11 with the names of the cells and connection ratios (Figure 12), specific connections can be identified.

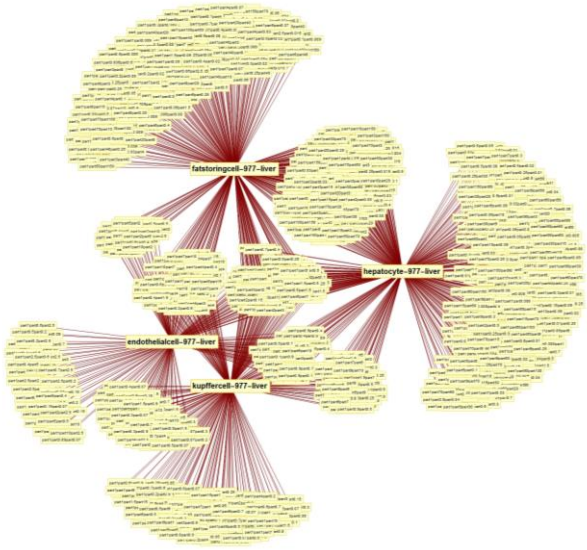

Figure 12 Names and locations of the cells and connections summarize the complexity of the relationship of one cell type to another in the rat liver. Enlarge image as needed.

To view cell to cell connectivity at a global level, we can plot the organelle surfaces of all the cell types in the stereology database against their mathematical markers (Figure 13). Once again, we find a similar pattern of shared markers. Since the differentiated cells of a given animal derive from the same source (zygote) and different animals share many of the same genes, this is exactly the pattern we would expect to see.

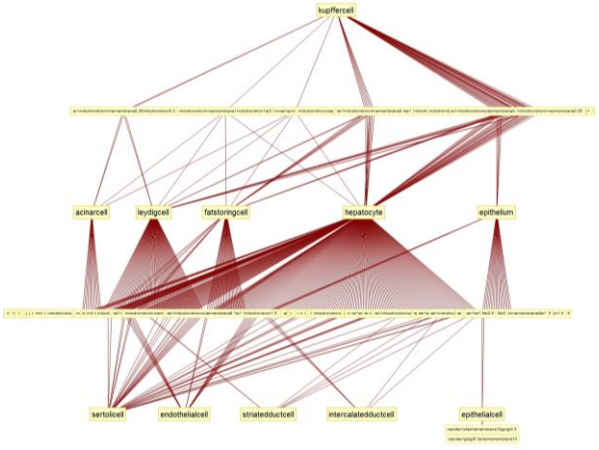

Figure 13 Extensive sharing of mathematical markers – derived from the surface areas of cell organelles - occurs within and across animal species.

If we take the same set of mathematical markers shown in Figure 13 but substitute the citation numbers for the cell types, we get a global view of cell to cell connectivity as it exists in the biology literature (Figure 14).

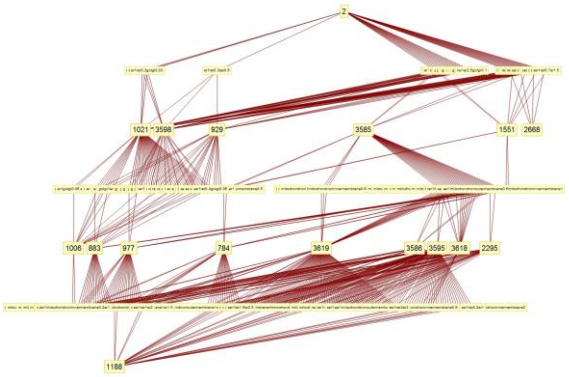

Figure 14 One way of testing for reproducibility consists of plotting the mathematical markers of cells against their citation numbers. The plot demonstrates the global nature (read reproducibility) of published data. Recall that the mathematical markers being used here include alphanumeric strings containing six variables – making it a tough test to pass.

Finally, if we replace the mathematical markers of Figure 14 with their connection ratios (Figure 15), the resulting view of the literature reveals an underlying global pattern of remarkable connectivity and reproducibility.

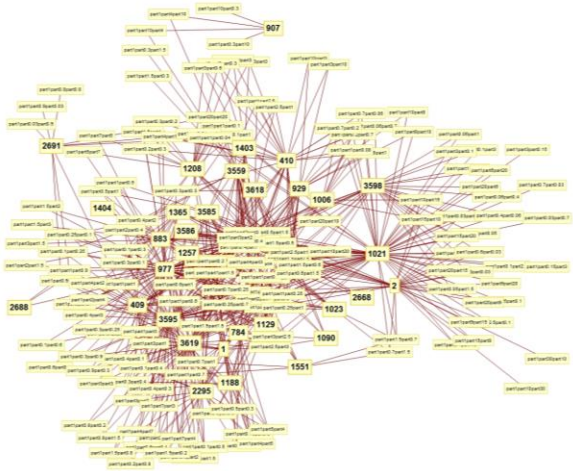

Figure 15 The plot displays the relationship between citations and connection ratios, which were derived from the surface areas of cell oranelles. It illustrates the presence of widespread connectivity within the stereology literature.

In summary, we can draw the following conclusions. By equating the existence of global data to that of reproducibility, it now appears that the biomedical literature contains far more reproducible results than we previously thought. The point to take from these examples is that we can readily find reproducibility throughout the literature, once we know where and how to look for it.

The Highly Adaptive Brain

Although earlier reports have explored the relationship of mathematical markers to disorders of the brain (Bolender, 2011-2015), our purpose here is to upgrade the topic by adding the new results coming from the connections ratios.

In spite of knowing many of the detailed mechanisms by which biological parts are produced, we still know surprisingly little about the quantitative relationship of parts to their connections. It now appears that a deeper understanding of this relationship will be fundamental to our understanding of biology and of the disease process.

Volumes: When, for example, we plot connection ratios against 24 disorders of the brain, the picture that emerged from the MRI data set (IBVD) was one of intense connectivity (Figure 16). The plot demonstrats the extent to which different disorders share a common connectivity platform.

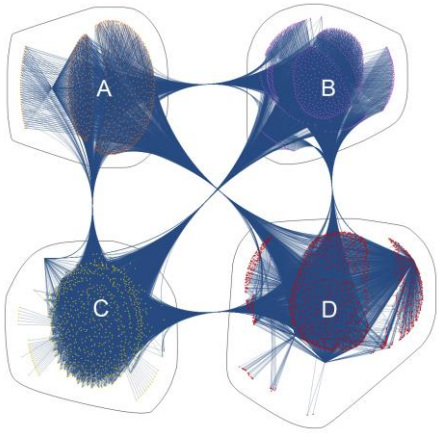

Figure 16 Connection ratios derived from a panel of 24 disorders of the human brain display four clusters all of which are interconnected. This MRI data set – derived from the IBVD - indicates that disorders share many of the same connection ratios. The clusters group disorders according to their affinities: A (Alzheimer, Williams-syndrome), B (adhd, Aspergers-syndrome, Huntington-disease), C (alcohol, autism, borderline-personality-disorder, Down-syndrome, dyslexia, epilepsy, fragile x, Klinefelter-syndrome, major-depressive-disorder, ocd, ptsd), and D (bipolar, panic-disorder, schizophrenia, velocardiofacial).

We can begin to understand the complexity of the disorders puzzle by plotting the connection ratios of normal patients (C) against those presenting with disorders (E). The results in Figure 17 show two sets of unique connection ratios (C, E) separated by a shared set (E=C). The extent of the sharing seen at the global level was unexpected because duplicate copies of the mathematical markers (E=E and C=E) were removed from the individual papers – before their connection ratios were plotted. Since a shared category exists (C=E), it appears that a given connection ratio can support either a normal or abnormal complement of parts. However, the key point to take from the figure is that distinct populations of connection ratios exist for both normal and abnormal brains (Figure 17). From this, it appears that connectivity is either contributing to the disease process or being produced by it. Whatever the case may be, the answer presumably originates somewhere in our DNA. But, where?

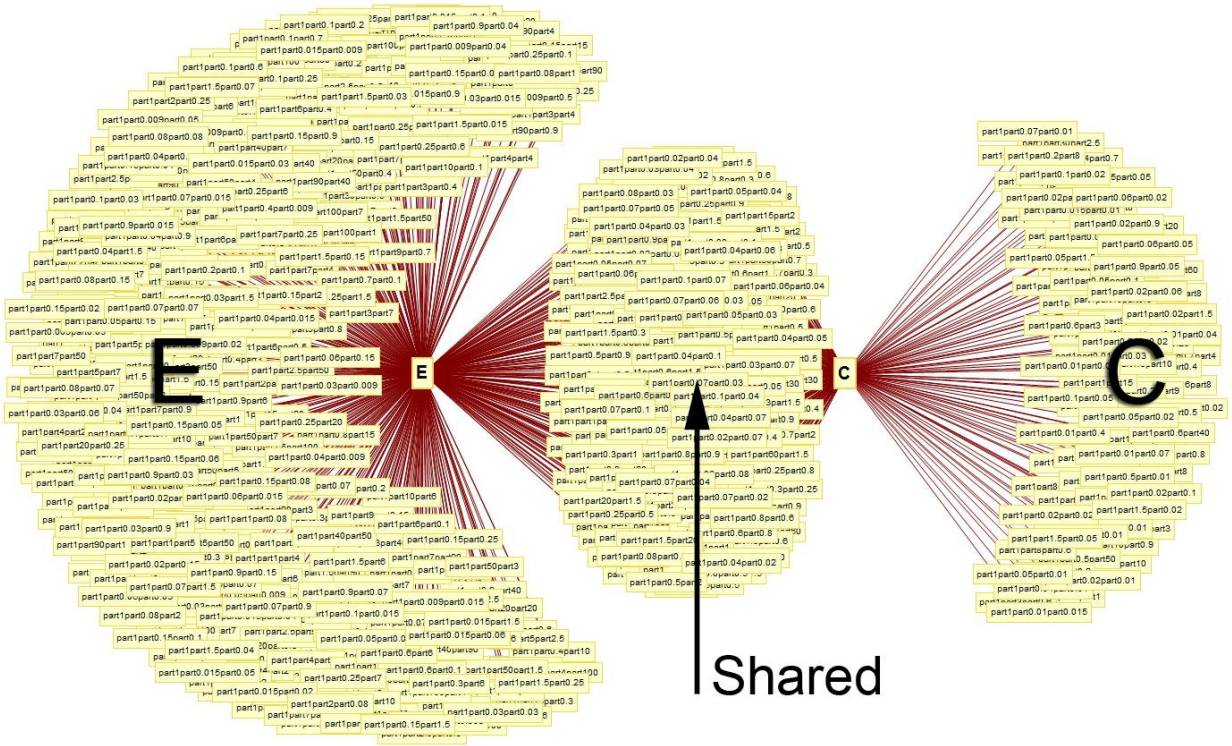

Figure 17 When the MRI data of the IBVD are translated into connection ratios for control (C) and experimental (E) patients and plotted, they form three distinct groups of connection ratios – abnormal (E), shared (E=C), and normal (C). Notice that once again we find a large number of connection ratios in play.

Next, we can burrow into the data set shown in Figure 17 to tease out some of the details associated with the connectivity patterns. If, for example, we take two MRI publications on schizophrenia (citations 652 and 5165) and plot the parts estimated therein against their citation numbers, we find that the papers have only a single part in common - the temporal lobe (Figure 18). Given such a result, it would seem logical to conclude that the data of the two papers have very little in common. In fact, this turned out not to be the case.

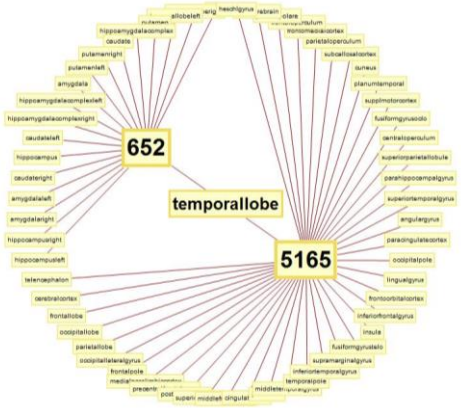

Figure 18 When the names of the parts published in two different papers on schizophrenia are plotted against their citation numbers, only only 1 of the 61 parts is shared by both papers (temporal lobe). Original data attributed to Swayze et al., 1992 [652] and Goldstein et al., 1999 [5165].

When we convert the volume data of these two papers (652 and 5165) into connection ratios and plot them against their citation numbers, notice what happens. Now, we find widespread connectivity. Almost all the connection ratios (97%) in publication 652 duplicate those in 5165 (Figure 19). Such a result suggests that schizophrenia is redefining the connectivity of brain parts in a way apparently unrelated to the identities of the individual parts. In effect, the fundamental cause of schizophrenia might be explained entirely or in part by the presence of abnormal connectivity ratios.

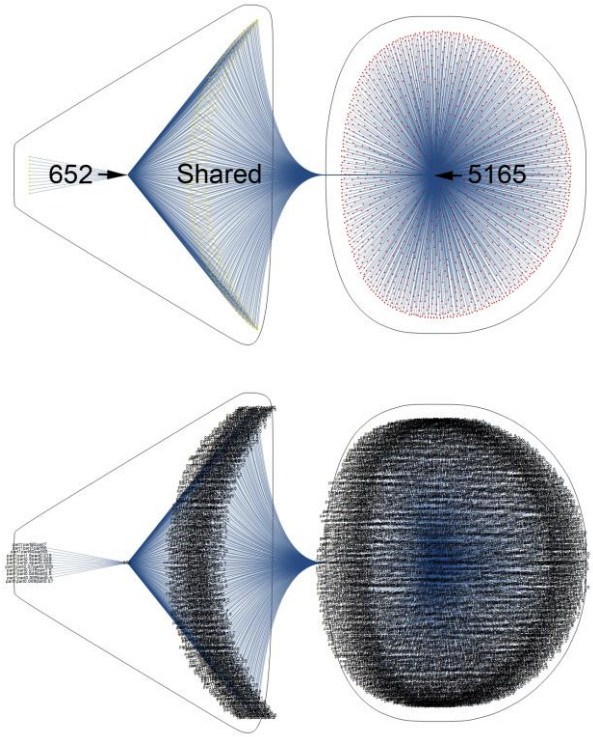

Figure 19 Publications 652 and 5165 share almost identical connection ratios, even when the parts of the two data sets are almost entirely different. Such a finding suggests that schizophrenia is instigating major changes in connectivity patterns throughout the human brain. If true, then schizophrenia may be interfering with the design rules at a deeper level, perhaps resetting a major portion of the biology blueprint.

This raises the possibility yet again that schizophrenia – as well as other disorders of the brain – have their roots in both parts and connections. If this is the case, then attempting to solve disorders such as schizophrenia using just the data of parts may prove to be unproductive. Once again, we seem to be remined that complex problems require equally complex solutions.

Numbers: Cell counts (numbers) associated with schizophrenia displayed a pattern similar to the one just described for volumes. The plot of parts vs. citation numbers also included a single match, which, in this case, was the medial dorsal nucleus (Figure 20). If, as before, we replace the names of the parts with their connection ratios, a similar burst of connectivity results (Figure 21).

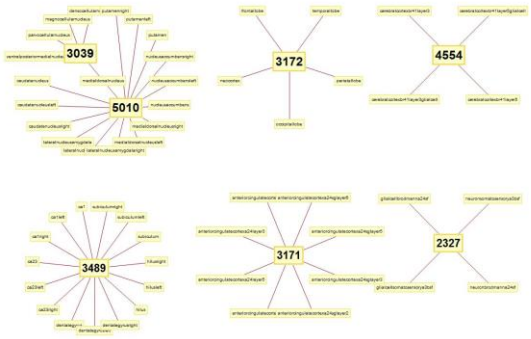

Figure 20 For those papers characterizing schizophrenia with cell counts (numbers) and mathematical markers, only two (3039 and 5010) shared the same part (medial dorsal nucleus).

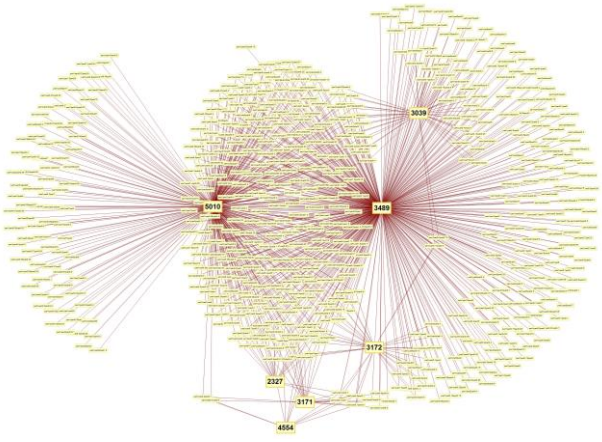

Figure 21 When the citation numbers of the papers listed in Figure 20 were plotted against their connection ratios, a pattern of widespread connectivity appeared.

The pattern persists for species. When we look for patterns in cell counts (control and experimental) across different animals, only humans and rats share the same mathematical markers (Figure 22). In contrast, Figure 23 shows that connection ratios cast a far wider net.

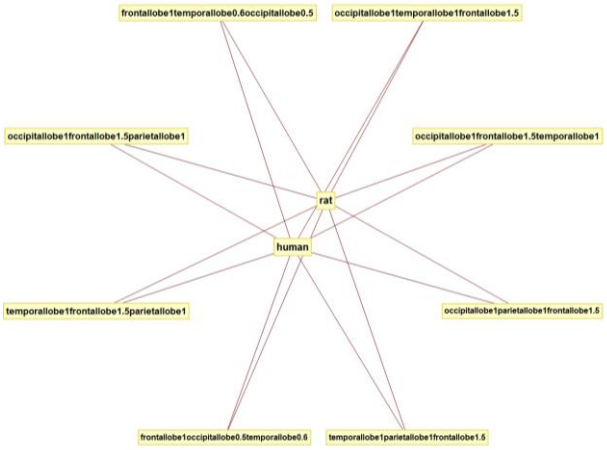

Figure 22 When mathematical markers derived from cell counts are plotted against species, only humans and rats shared the same markers.

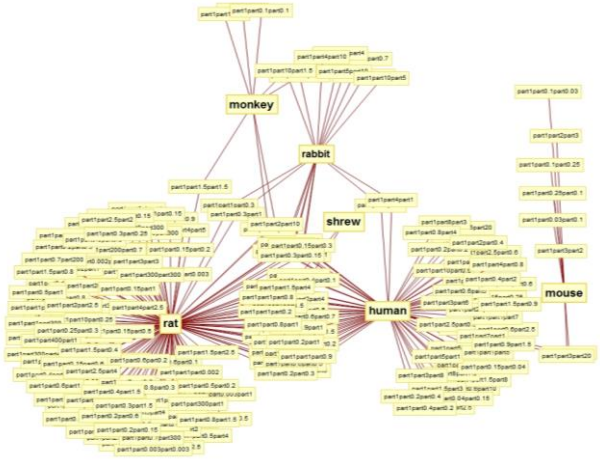

Figure 23 When, however, the connection ratios replace the mathematical markers, many more animals became connected by sharing similar patterns. Note that of the total number of connection ratios (108,824), 1,236 occurred at least three times. The distribution included: human (396), monkey (38), mouse (24), rabbit (72), rat (704), and shrew (2).

DISCUSSION

Finding a solution to the reproducibility problem required the use of patterns. In complex settings, such as biology, only patterns seem capable of accessing the deeper levels of information needed to address fundamental questions. By managing reproducibility effectively, we can now enjoy the advantage of having both accuracy and precision in play. The test, which found reproducibility to exist across all the data sets considered (Figures 1-23), demonstrated that the biology literature can be a rich source of reliable information.

This global reach of patterns stands in sharp contrast to our current approach to reproducibility in the life sciences, which consists largely of duplicating the data points of a single experiment. Although increasing precision can improve reproducibility locally, advanced players are likely to prefer accuracy because of its relationship to first principles. Besides, finding an accuracy-based solution to the reproducibility problem required little more than noticing that biology expresses its reproducibility principle as ratios of its parts. Of course, the genius of biology is clearly in play in that it uses these ratios to let accuracy and precision converge – producing the best possible solution.

Arguably, playing the complexity game with biology requires rethinking almost everything we do. Not taking the leap into complexity, however, appears to be even more daunting. Our chronic inability to solve the reproducibility problem serves as a stark reminder of the limitations being imposed on our science by a grievously outdated theory structure. To make matters worse, our statistical allies are now taking us to task over the inadequacies of our research data (Ioannidis, 2005; Van Regenmortel, 2004), our funding is on a downward spiral (NIH Re-

search Funding Trends, Nature Cell Biology Editorial, 2012), we continue to be cut off from our published data by paywalls (Murphy, 2016), and our results are chronically open to criticism (The Economist, 2016). The recent article in *Nature* (Baker, 2016) aptly calls our current state a crisis, but our problems go far deeper than just precision. We find ourselves in trouble because we are failing to run our science in a way consistent with the principles of the living systems we are trying to understand.

If we look at reproducibility from biology's perspective, its solution to the problem illustrates how it deals with complexity. To maintain its status as a living system, it must enforce strict rules of reproducibility to insure the existence of its emergent properties (e.g., life, cognition, survival). At the same time, however, it must allow its parts to change in response to those forces within and beyond its control. In effect, using ratios to maintain order instead of focusing on just individual parts represents a clever solution, one that obviously works quite well.

Reproducibility Repackaged

Our current willingness to base our definition of reproducibility on precision (Baker, 2016), comes with more than a few drawbacks. We can flesh out some of these negatives by asking simple, but thought-provoking questions.

Consider the standard definition of reproducibility given at the outset: "Reproducibility is defined as an ability to duplicate the results of an experiment either by the same researcher or by an independent one." If we express this definition as an expression, we would expect it to contain a variable that acts like a constant. To wit:

$$P_{i\,(experiment\ 1)} = P_{i\,(experiment\ 2)} \ldots = P_{i\,(experiment\ n),} \quad (1)$$

where P is a data point characterizing some part i.

Let's look at a published data set to see if equation 1 gives us what we want. If, for example, we select estimates for the amygdala from the IBVD and plot them, we find a widely-dispersed cloud of data points (Figure 24), illustrating the variation we have come to expect from biology. If we duplicate any one of these points – in keeping with the logic of equation 1 - will we satisfy the reproducibility requirement? In keeping with the definition above, yes.

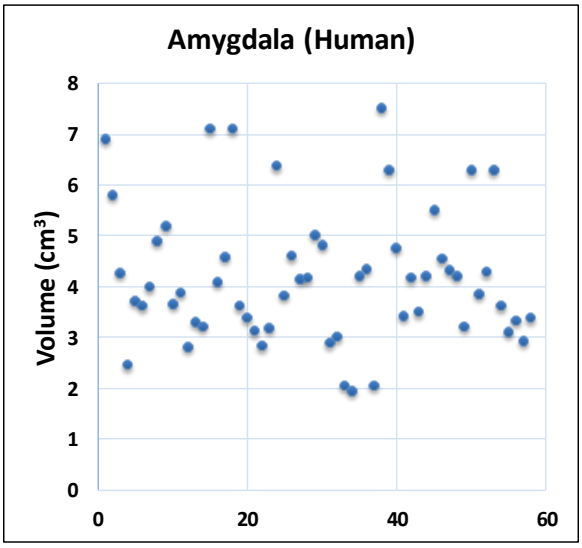

Figure 24 A plot of 58 estimates for the volume of the amygdala produces a cloud of data points (IBVD). Note that each data point represents the average of several patients.

Herein lies the problem. Every point we choose to duplicate represents a local solution, which satisfies the precision part of the reproducibility definition, but ignores the more important role of accuracy (correctness). Such an arrangement triggers dicey questions. Does this mean that by duplicating an experimental outcome we can consider the outcome to be correct – even if it is incorrect (precise, but inaccurate)? If an experiment cannot be duplicated, does it mean that the original results were wrong? Can one answer such contradictory questions convincingly? Most likely, no.

Since estimates for the amygdala produced a cloud of scattered data points (Figure 24), does this mean that an accurate estimate is impossible because of biological variation and experimental biases? Such an irritating question becomes inescapable when we use the precision argument to define reproducibility. To avoid having to answer such unanswerable questions, the report substituted a definition for reproducibility based on accuracy. This put the onus of having to defend the results on biology and its logical surrogate - the biology literature (Figures 3-5).

The strategy behind taking an accuracy-based approach to reproducibility, of course, requires asking seemingly unanswerable questions with the knowledge that biology already knows the answers. How can this be demonstrated? Consider the following example. If the amygdala is subject to biology's accuracy rule, is it possible to show that the 58 points displayed in Figure 24 were generated accurately – despite their scattered distribution? The answer, of course, is yes - if we know how biology applies its accuracy rule.

Let's start with the target analogy often used to explain accuracy and precision. Recall that accuracy can be likened to a set of data points (hits) all of which cluster in the center of the bullseye. Clearly, this is not the case in Figure 24. Accordingly, most reasonable people would conclude that biology ignores accuracy when it produces an amygdala. Biology, however, would strongly disagree. If we look at the set of points shown in Figure 24 through biology's eyes, the amygdala would appear as a single point (Figure 25). Why? Because biology puts its accuracy where it works best for biology.

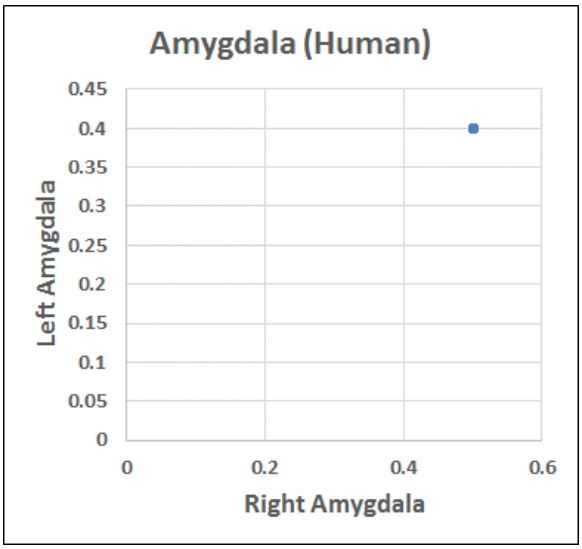

Figure 25 When the left and right sides of the amygdala are expressed as 58 ratios, the scatter plot shown in Figure 24 is replaced by a single point (a decimal ratio value). Notice that such a point becomes a measure of both accuracy and precision. Adapted from Bolender, 2012; see reference for details).

What does this mean? Biology defines accuracy in the amygdala not in terms of a single volume, but rather as the ratio of two volumes - left to right (0.4:0.5). Biology knows how to have it all (Figure 25) – accuracy (all 58 pairs of points are the same) and precision (all the points are superimposed).

Reproducibility Tested

The reproducibility test provides two new capabilities. It gives us a single test that can be applied to a wide range of data types (Figures 3-5) and it supplies the global data needed to optimize both precision and accuracy (Figures 1-25). By equating reproducibility to global data, we can define the phenotype as a continuous, quantitative platform extending from MRI (volumes) to light microscopy (cell counts) to electron microscopy (surface areas) to molecular assays (optical densities). The design of such a platform - defensible by quantitative arguments - simplifies our task of making the transition in

our approach to biology from simple to complex.

The reproducibility test was designed to make it extremely difficult to pass. Triplet markers consisting of six variables had to appear as duplicates in at least three different papers - or separate studies - to qualify. Nonetheless, the databases demonstrated repeatedly that the biology literature can meet and exceed this minimum requirement routinely (Figures 3-5, Figure 26). Consider, for example, the results shown in Figure 26. It would be impossible for this image to appear unless biology and the biology literature were on the same page.

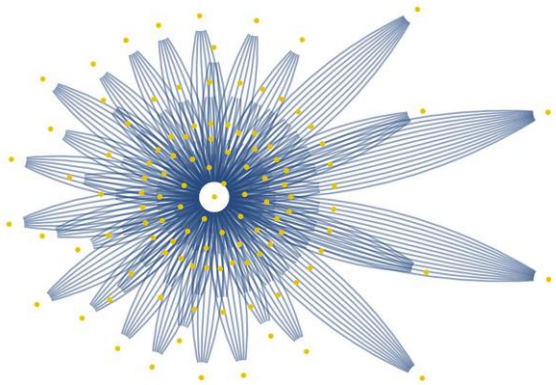

Figure 26 The pattern illustrates biology's rule-based approach to managing reproducibility. When MRI data are plotted, mathematical markers form such rosettes (reproducibility units) centered on unique connection ratios. Since many repeats (blue lines) can be seen for the peripheral points (mathematical markers), we can now imagine that both biology and the biology literature are remarkably good at using complex patterns to oversee accuracy. See Figure 2 for labels.

The big win to come from the reproducibility test, however, was the appearance of global patterns for membrane surface areas - estimated stereologically with electron microscopy (Figures 9-15). Since biological membranes supply the structural platform for constitutive marker enzymes, they should allow us to complete a key part of the biology puzzle - a quantitative phenotype extending all the way from organisms to molecules.

Such a phenotype will have diagnostic and predictive properties and can be stored as a relational database with unlimited entry and exit points.

Summary: Reproducibility Test

- *The reproducibility test consists of plotting mathematical markers against their connection ratios.*
- *It tests for the presence of global data, which serves as a measure of reproducibility.*
- *Reproducibility comes from global patterns that come from biological accuracy.*
- *It is assumed that global data are present when three or more duplicate copies (mathematical markers) exist for a given connection ratio.*
- *The reproducibility test identifies MRI data collected from patients as its gold standard.*

Connectivity Generates Patterns

Global connectivity can be shown to exist in the biology literature when data are expressed as mathematical markers and connection ratios. In living systems, it appears that change, diagnosis-prediction, disorders of the brain, and now reproducibility all derive their global properties largely from the connectivity of parts (Bolender, 2012-2015).

The connection ratios, which are defined herein as alphanumeric strings (partX:partY:partZ), represent a new data type. They form distinct patterns, exist either as unique or shared strings, and define quantitatively a fundamental property of biology.

Moreover, a given connection ratio embedded in a mathematical marker can be repopulated with different parts or alternate between steady and transitional states during the phases of a change (Bolender, 2016; Figure 1.11). In other words, biology can hold connections constant and change the parts or hold parts constant and change the connections. Detecting such events would appear to be essential to understanding the complexity of a biological

change. As we discovered earlier with ratios, valences, and mathematical markers, connection ratios offer deeper insights into the inner workings of biology.

We are beginning to understand that some of the best questions are those that seem at first impossible to answer. If, for example, connection ratios identify rule-based patterns common to living systems, where do such rules come from and how does biology know when and where to apply them? What factors, for example, influence biology to select normal or abnormal connection ratios (Figures 16-21)? To answer such questions, two possibilities come to mind. Some design principles might be explained by self-assembly, whereas others may be the result of algorithms scripted in DNA. If such scripts exist, how does one find and interpret them? When we change the coding of our DNA, how will this affect our global patterns?

Summary: Connectivity

- *Connections represent one of two major components fundamental to biology as a complex system – parts (amount, composition) and connections (ratios).*
- *Connectivity can be identified as the major source of global data, reproducibility, and accuracy in living systems.*
- *Connection ratios represent a new data type, largely unexplored.*
- *Connection data account for a substantial portion of the quantitative information coming to us from biology.*

Disorders Reorder

Disorders of the brain can be studied using parts (volume, surface, length, and number), connections (connection ratios), and a combination of the two (mathematical markers). Problems arise, however, in that each option, which defines a distinct data platform, can lead to different outcomes with different interpretations.

Moreover, even bigger problems arise when we collect data from one platform and then try to interpret it on another. In fact, our current practice of routinely switching between platforms helps to explain why the biology literature can be so contradictory. Consider, for example, our standard experimental method. We simplify biology by selecting a few parts, quantify them, and then try to explain how and why they behave as they do in a complex biological setting - knowing little about the behavior of the complexity. A change, when taken out of context, is interpreted out of context. It becomes disconnected from the reality of biology.

If, instead, we start with simple data, translate them into complex patterns, and then use the patterns to solve a complex problem, we usually get what we want (e.g., a reproducibility test based on accuracy). Since this approach to problem solving generates new discovery platforms, we can use them to explore new solutions to other problems.

Consider disorders of the human brain. After extracting connection ratios from their mathematical markers and then plotting all 24 disorders as a single group, widespread connectivity appeared within and across MRI publications (Figure 16). When, in turn, control data were added to the data set of Figure 16, three unique clusters of data appeared - abnormal (E), shared (E=C), and normal (C) (Figure 17). This result shows that disorders of the brain define a distinct population of abnormal connection ratios, which in some way are the result of an unknown process capable of modifying ratios. The presence of such a large population of abnormal ratios (E) highlights the magnitude of the problem. However, we are presented with yet another layer of complexity by the shared cluster, where C=E. The same connection ratio can be populated with either normal or abnormal parts. Such a population becomes a ready source of false positives.

A curious pattern appeared repeatedly in Figures 16 to 21. The disorders seem to depend

more on the connectivity of the parts (ratios) than on the parts themselves (numerical values). This suggests that solving the disorders puzzle – either individually or as a group – will require a far better understanding of the biological principles underlying connectivity.

Summary: Disorders of the Brain

- *Disorders of the brain define a unique set of connection ratios.*
- *Connection ratios can define the relationship of one disorder to another quantitatively.*
- *Connection ratios can become false positives.*

Concluding Comments

If we define reproducibility as an ability to duplicate the results of an experiment, but have little success in doing so, then the simplest solution is to change the definition. Such was the strategy pursued in this report.

Reproducibility was redefined as the ability of biology (or an observer) to duplicate complex patterns globally. In turn, the report determined the effectiveness of this revised definition by testing three different data sets for the presence of global data (Figures 1-23). The results provided compelling evidence that the biology literature contains vast quantities of global data quite capable of passing reproducibility tests. Why? By synchronizing the biology literature with biology, we were able to connect the dots: global data…reproducibility…accuracy.

With the reproducibility piece of the biology puzzle in our portfolio, we find ourselves one step closer to the goal of building a single network of information stretching seamlessly from phenotypes to DNA. By extending the reach of our quantitative phenotype to membrane organelles (surfaces), we now have a platform from which to attempt a jump from cell organelles to molecules. Since biological membranes carry protein molecules coded for in the genome, such molecules become a convenient link to the databases of molecular biology. By translating molecular data into complex data types, we can begin to explore the relationship of patterns in the phenotype to those of DNA and RNA. With diagnosis, prediction, and reproducibility in play, we now know where and how to look.

In a world defined by information, a science defines itself by the way it makes its information available to those responsible for producing it.

REFERENCES

Baker, M. (2016) Is There A Reproducibility Crisis? Nature 533, 452–454 (26 May 2016) doi:10.1038/533452a; Nature's Questionnaire (2016) : http://www.nature.com/polopoly_fs /7.36741%21/ file/Reproduciblility%20Questionnaire.doc

Begley, C.G., and Ioannidis,J.P. (2015) Reproducibility in Science: Improving the Standard for Basic and Preclinical Research. CircRes 116: 116126.doi: 10.1161/CIRCRESAHA. 114.303819 PMID: 25552691

Blouin A., Bolender R.P., Weibel E.R. Distribution of organelles and membranes between hepatocytes and nonhepatocytes in the rat liver parenchyma. A stereological study. J Cell Biol. 1977 Feb;72(2):441–455.

Bolender, R. P. 2006 Enterprise Biology Software VII. Research (2006) In: Enterprise Biology Software, Version 6.0 © 2006 Robert P. Bolender

Bolender, R. P. 2010 Enterprise Biology Software XI. Research (2010) In: Enterprise Biology

Software, Version 10.0 © 2010 Robert P. Bolender

Bolender, R. P. 2011 Enterprise Biology Software XII. Research (2011) In: Enterprise Biology Software, Version 11.0 © 2011 Robert P. Bolender

Bolender, R. P. 2012 Enterprise Biology Software XIII. Research (2012) In: Enterprise Biology Software, Version 12.0 © 2012 Robert P. Bolender

Bolender, R. P. 2013 Enterprise Biology Software XIV. Research (2013) In: Enterprise Biology Software, Version 13.0 © 2013 Robert P. Bolender

Bolender, R. P. 2014 Enterprise Biology Software XV. Research (2014) In: Enterprise Biology Software, Version 13.0 © 2013 Robert P. Bolender

Bolender, R. P. 2015 Enterprise Biology Software XV. Research (2015) In: Enterprise Biology Software, Version 14.0 © 2015 Robert P. Bolender

Bolender, R.P. (2016) Playing the Complexity Game with Biology. © 2016 Enterprise Biology Software Project, PO Box 292 Medina, WA 98039-0292

Collins, F.S. and Tabak L.A. (2014) NIH plans to enhance reproducibility. Nature 505:612–613.PMID: 24482835

Engber, D. (2016) Cancer Research Is Broken - There's a replication crisis in biomedicine—and no one even knows how deep it runs. Future Tense, April 19 2016 FROM SLATE, NEW AMERICA, AND ASU

Freedman, L.P., Cockburn, I.M., and Simcoe, T.S. (2015) The Economics of Reproducibility in Preclinical Research. PLOS Biology|DOI:10.1371/journal.pbio.1002165

Goldstein J.M., Goodman J.M., Seidman L.J., Kennedy D.N., Makris N., Lee H., Tourville J., Caviness V.S. Jr, Faraone S.V., and M.T. Tsuang. (1999) Cortical abnormalities in schizophrenia identified by structural magnetic resonance imaging.

Gundersen H. J. G., Bagger P., Bendtsen T. F., Evans S. M., Korbo L., Marcussen N., Moller A., Nielsen K., Nyengaard J. R., Pakkenberg B., Sorensen A., Vesterby, and M. J. West. (1988) The new stereological tools: disector, fractionator, nucleator and point sampled intercepts and their use in pathological research and diagnosis. Acta Pathol. Microbiol. Immunol. Scand. 96: 857-881.

Ioannidis, J.P.A. 2005 Why most published research findings are false. PLoS Med 2: e124.

Kennedy D. N., Hodge S. M., Gao Y., Frazier J. A., and C. Haselgrove. (2012) The internet brain volume database: a public resource for storage and retrieval of volumetric data. Neuroinformatics. Apr; 10(2):129-40.

Murphy, K. (2016) Should All Research Papers Be Free? The New York Times, March 12, 2016

Nature Editorial (Online): Science funding: championing research in tough times. Nature Cell Biology 14, 439 (2012) doi:10.1038/ncb2499

NIH Research Funding Trends: http://faseb.org/Science-Policy-and-Advocacy/Federal-Funding-Data/NIH-Research-Funding-Trends.aspx

Roth, K.A. and Cox, A.E. (2015) Science Isn't Science If It Isn't Reproducible. AmJPathol 185:2–3.doi:10. 1016/j.ajpath.2014.11.001 PMID:25529794

Swayze V.W., Andreasen N.C., Alliger R.J., Yuh W.T., and J.C. Ehrhardt. (1992) Subcortical and temporal structures in affective disorder and

schizophrenia: a magnetic resonance imaging study. Biol Psychiatry. 31(3):221-40.

The Economist (2016) Incentive malus: why bad science persists. September 22, 2016 http://www.economist.com/news/science-and-technology/21707513-poor-scientific-methods-may-be-hereditary-incentive-ma-lus?frsc=dg%7Cd

Van Regenmortel. M.H.V. (2004) Reductionism and complexity in molecular biology. EMBO 5(11): 1016–1020. doi:10.1038/sj.em-bor.7400284 PMCID: PMC1299179