In: Enterprise Biology Software, Version 15.0 © 2015 Robert P. Bolender

# Enterprise Biology Software: XVI. Research (2015)

ROBERT P. BOLENDER

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

http://enterprisebiology.com

## SUMMARY

The human brain supports a curious mission statement. In addition to maintaining a healthy state, it must also produce and accommodate a large number of different disorders. How does the brain manage this seemingly contradictory task? Attempting to answer such a question will involve understanding how biology manages the disease process at a more basic level. To do this, we can use mathematical markers to phenotype twenty-one disorders of the brain, mix them all together, and then use hierarchical cluster analysis in an attempt to workout the relationships of markers to disorders to symptoms. We begin by creating a database of duplicate markers (those that occur in more than one disorder) derived from the Internet Brain Volume Database (IBVD) of Kennedy et al. (2012). Using this derived database as our parallel complexity, we proceed by unfolding the twenty-one disorders stepwise until only pairs of disorders remain. This type of analysis reveals that disorders of the brain display distinct quantitative relationships and patterns based on well-defined modules. By knowing the quantitative relationship of one disorder to another, we can add in a symptoms layer to see how they compare. Furthermore, we can use the duplicates database to tease out the specific parts of the brain that contribute most often to the disease process. Such an analysis puts the hippocampus at the top of the list, followed by the amygdala, caudate, and putamen. In summary, the report suggests that the complexity of the brain in health and disease derives – at least in part – from the arrangement and rearrangement of highly conserved modules – all of which can be captured as mathematical markers. The current software package includes a copy of the duplicates database along with instructions for its use.

## INTRODUCTION

Complexity theory contributes to the ongoing process of rethinking the way we approach and study biology as a science. It provides a seemingly endless collection of lenses through which we can view published data in innovative ways. In this report, we will look at disorders of the brain not as separate topics, but rather as a single problem. Such an approach provides new insights into the way the brain manages the disease process. As a follow up, we will apply a similar analysis to the symptoms-based approach of clinical medicine.

As usual, we will proceed by subcontracting the hardest parts of the undertaking to biology, which will show us how to assemble and interpret the complexity of disorders by way of a database designed to serve as its parallel complexity.

Living systems remain notoriously difficult to study and understand because they form hierarchical structures, wherein complexities embed in complexities to form n-dimensional networks of interconnected parts. To the observer, such arrangements often appear as impenetrable tangles. By phenotyp-

ing disorders of the brain with mathematical markers, however, one quickly discovers that a given marker may or may not be unique to a given disorder. At first, this creates confusion. Fortunately, storing unique and duplicate markers in separate databases eliminates such confusion. A database of unique markers is well suited to the tasks of diagnosis and prediction (Bolender, 2014), whereas one containing duplicate markers supports the task of unfolding the disease process.

The principal goal of the current report is to identify generalizations that we can apply to disorders of the human brain. To this end, we will assemble a database of duplicate markers - derived from the Internet Brain Volume Database (IBVD) of Kennedy et al., (2012) – and use it to look for patterns by applying cluster analysis. The advantage of this approach is that we can use it to create patterns from different types of information, which, in turn, we can stack vertically to explore the relationship of one complexity to another. In time, such stacks should allow us to work our way back from the adult structures of the brain to their origins in the genome.

## METHODS AND RESULTS

The software package for 2015/2016 includes new databases and software tools for finding patterns related to disorders of the brain – using mathematical markers derived from published data.

### Enterprise Biology Software Package

The software includes eight screens offering ready access to programs, databases, and documents (Figure 1).

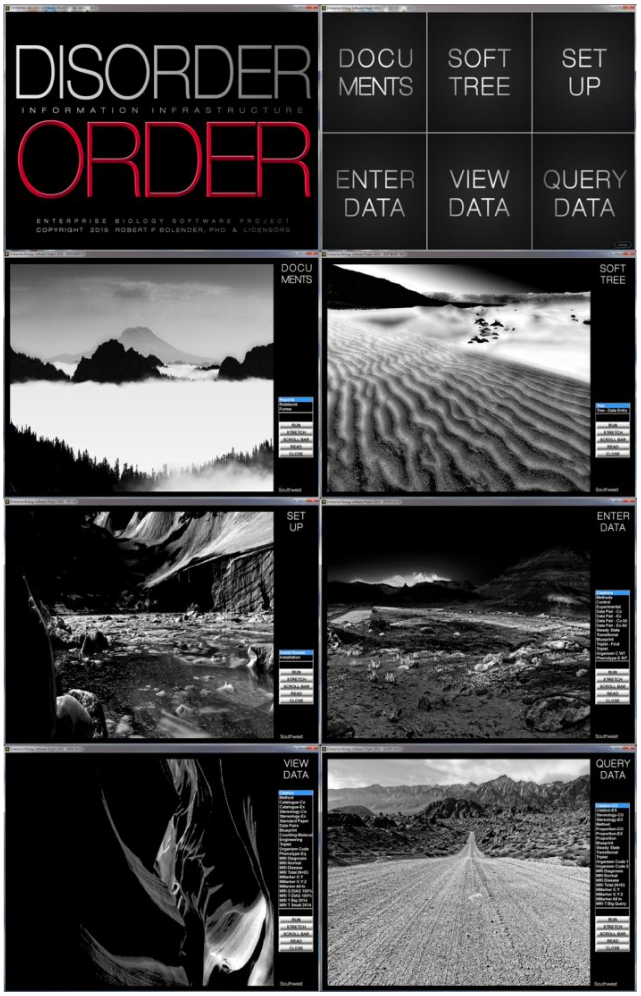



**Figure 1. Enterprise Biology Software Package – 2014/15. The 4GB package contains 354 files stored in 9 folders.**

### Assembling a Database of Duplicate Markers

The report uses mathematical markers to phenotype twenty-one disorders of the human brain with the goal of identifying fundamentals of the disease process. To this end, a database of triplet markers (AX:BY:CZ) characterizing these disorders was assembled and filtered to provide the attributes of a parallel complexity appropriate to the task (mri_t_dups_2015 .accdb).

Starting with a table of triplet markers containing control and experimental data sets, duplicates that existed between controls and experiments were removed first at the level of individual papers (local) and then at the level of the entire data set (global).

In the final set of experimental markers, only those that existed as duplicates in more than one paper (≥2) were retained. Moreover, for each paper, a given marker for a given disorder appeared only once in the database. The resulting duplicates database became the parallel complexity that provided the data used to unfold the disease process graphically.

## Graphical Analysis - Formatting Data

Running a graphic analysis involves a multistep process. First, we select the data of interest from the duplicates database (Microsoft Access) and then transfer them to an Excel worksheet. At this point, the data consist of two columns in a table – one populated with mathematical markers and the other with the names of the disorders. For example:

brain1cerebellum0.1cerebrum0.8 autism.

Before we can operate on these markers graphically in Mathematica (Wolfram), however, we need to add formatting, as shown below:

"brain1cerebellum0.1cerebrum0.8" -> "autism", .

This is accomplished by first adding a new column to the table for each formatting character set (|"|"->"|",|) and then using the resulting table to create formatted strings – using the concatenation option in Excel. In turn, the concatenated column is copied to an empty column and pasted using the Paste Values option, identified as 123. The resulting formatted data set is now ready to be pasted into a Mathematica worksheet. The software package includes the Access database of the formatted strings used to generate the images of the cluster analysis (mri_markers_plot.accdb).

## Communities of Disorders

The human brain supports a wide range of disorders, some of which appear in Figure 2. To explore these disorders, we will use the affinities that exist between markers and disorders to pry apart the nested complexities. This creates an opportunity to view the inner workings of the disease process.

Starting with the collection of mathematical markers associated with the 21 disorders illustrated in Figure 2, we can use the CommunityGraphPlot of Mathematica to identify the relationships of disorders as clusters (Figures 3-6). This allows us to start with a single, large complexity (21 disorders) and unfold it progressively into smaller sets of complexity. Along the way, patterns appear that reveal elements of the strategy biology uses to assemble disorders.

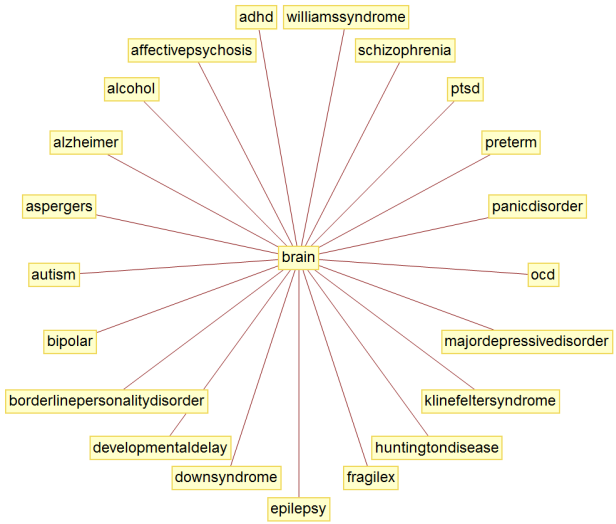



**Figure 2 The database of shared markers includes a collection of mathematical markers quantifying several disorders of the brain.**

**Step 1:** A CommunityGraphPlot – applied to the entire database of duplicate markers (Figure 2) - produces five distinct clusters (Figure 3), each of which represents a subcomplexity. Notice that extensive connections remain between the clusters of dots (mathematical markers and disorders), indicating widespread sharing of markers across disorders.

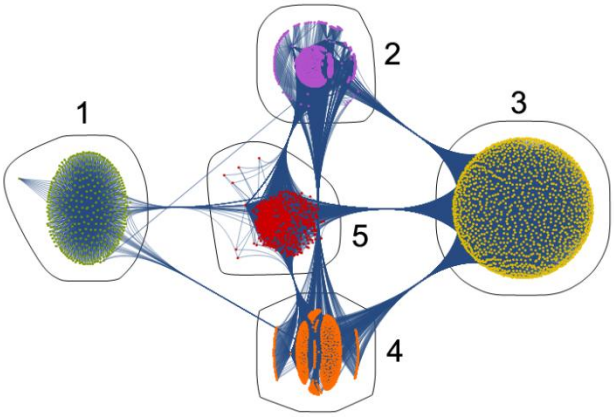
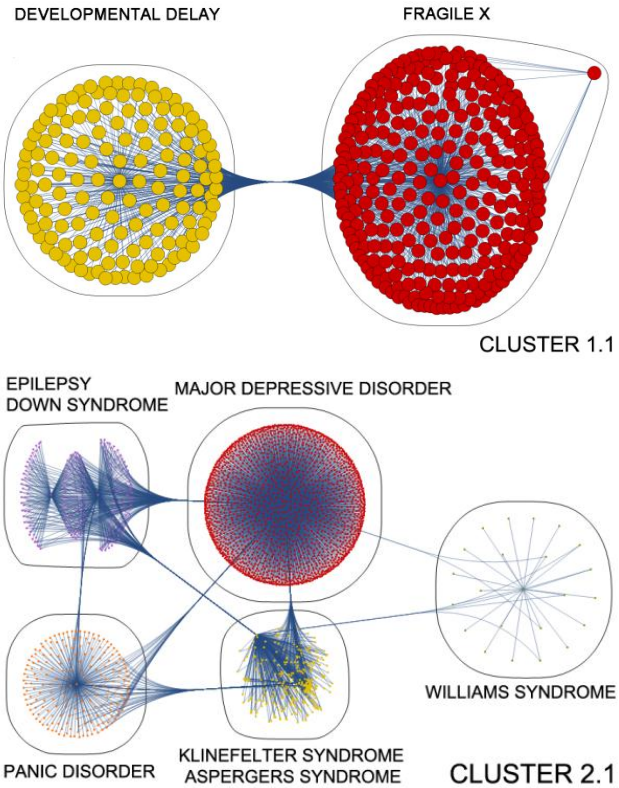



**Figure 3 Step 1: The collection of shared mathematical markers from the human brain (see Figure 2) distribute – as communities - into five distinct clusters. Note that each dot pinpoints a mathematical marker or a disorder and that the dark blue lines identify the connections between the markers and disorders.**

**Step 2:** Next, we can plot each of the five clusters identified in Figure 3 separately with the CommunityGraphPlot (Figure 4). Clusters 1, 2, 4, and 5 in the plot of Step 1 now display clusters of their own. Note that cluster 3 contains a single disorder (Alzheimer).



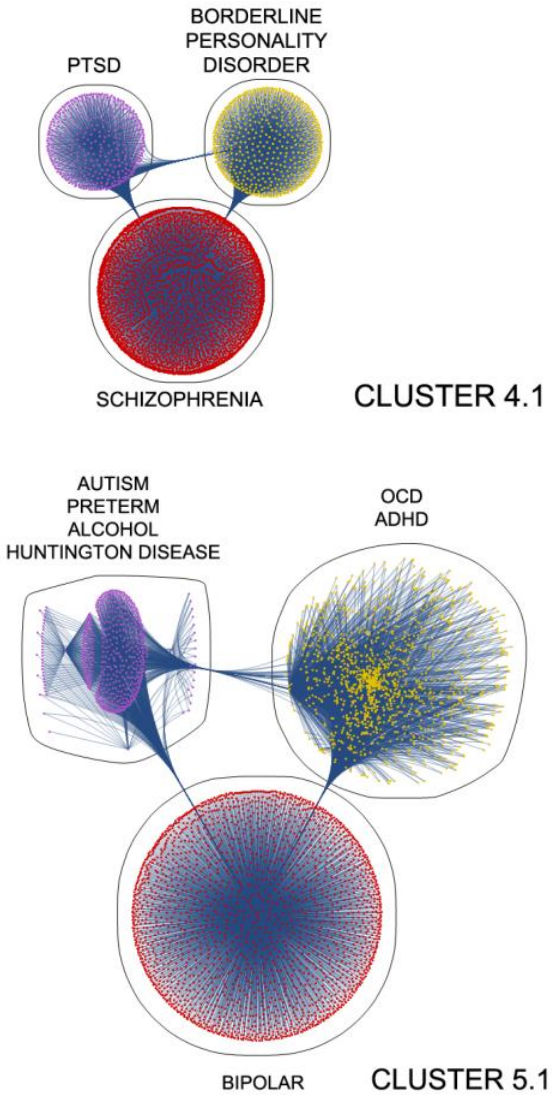
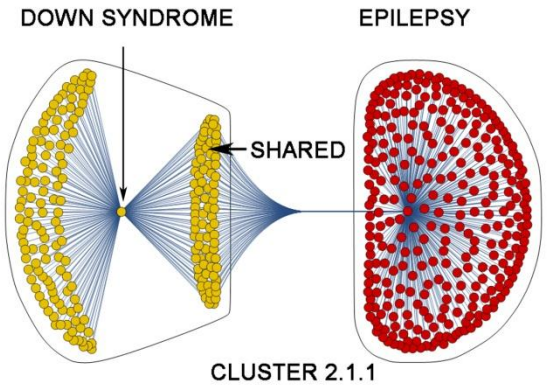
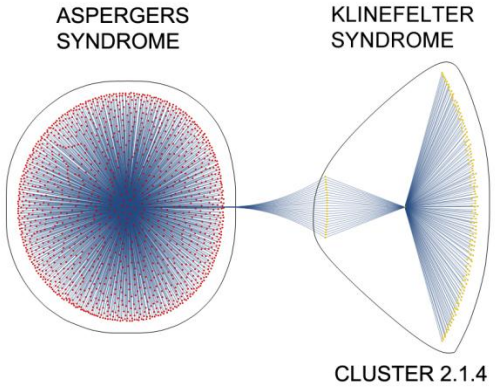
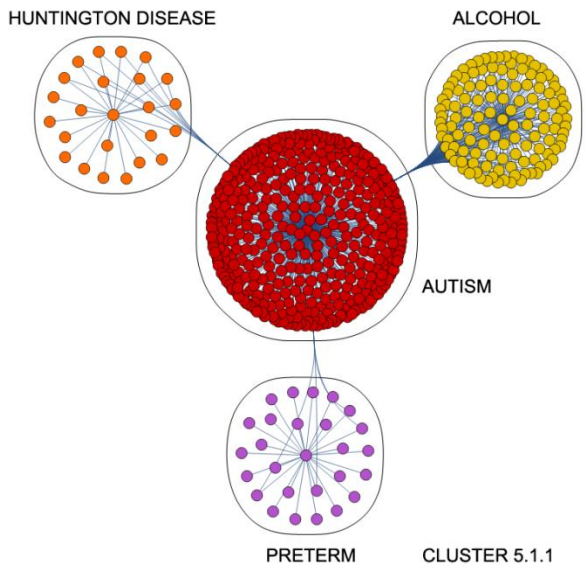




**Figure 4 Step 2: CommunityGraphPlots illustrate the sharing of markers (modules) between disorders. Each dot represents a mathematical marker or a disorder; the number of connections (blue lines) indicates the extent of the sharing.**

**Step 3:** Since clusters 2.1, 4.1, and 5.1 of Figure 4 still show multiple complexities, we can continue to separate the disorders with the CommunityGraphPlot (Figure 5). Now, for the most part, the clusters include only two disorders. At this point, the shared markers become readily apparent as belonging to an intervening spindle shaped structure. Also, notice that some of the disorders can now be identified as a single dot.



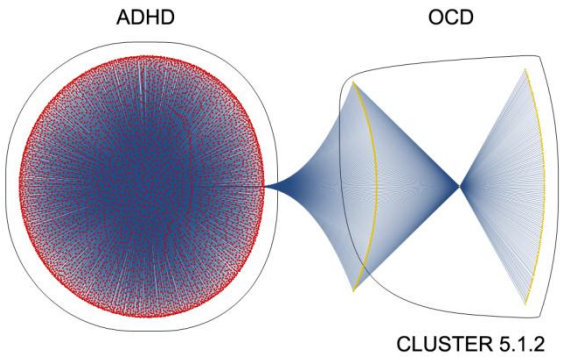
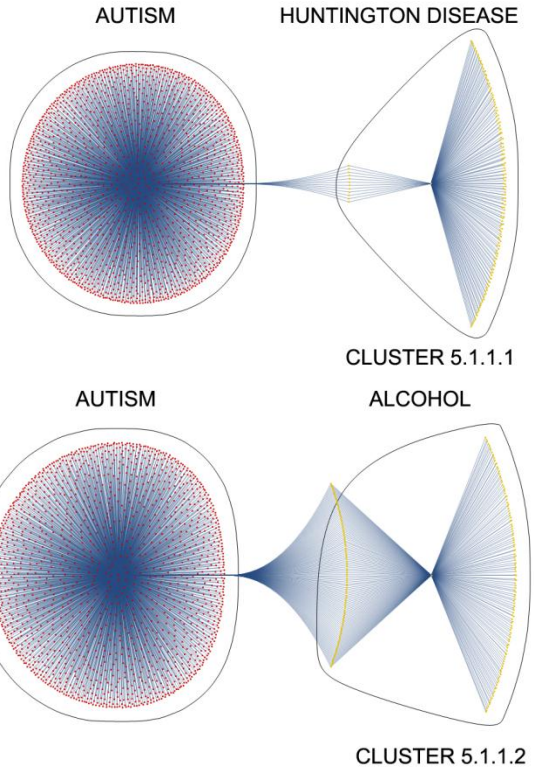




**Figure 5 Step 3: Clusters containing multiple disorders in Step 2 are unfolded into clusters containing a single disorder – except 5.1.1.**

**Step 4:** In this the final step, cluster 5.1.1 is resolved into pairs of clusters (Figure 6).



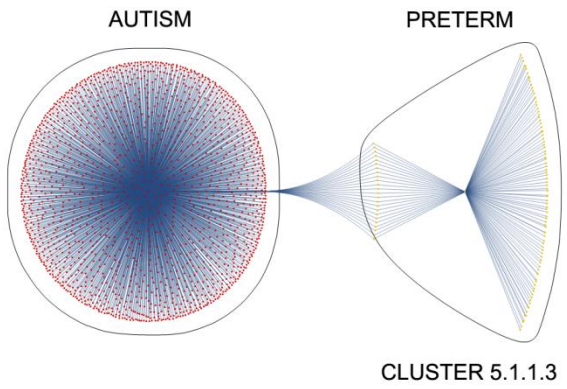



**Figure 6 Step 4: Cluster 5.1.1 is resolved into three sets of clusters relating the markers of autism to Huntington disease, alcohol, and preterm.**

**Summary:** Figure 7 summarizes the original set of clusters identified by the first CommunityGraphPlot (Figure 3). The results indicate that the phenotypes of the disorders in a given cluster have more in common with each other than with any of the remaining disorders – even when substantial connection exist between all clusters. Such information suggests that disorders sharing similar properties (read markers) may also share similar origins, treatments, and solutions. Keep in mind, however, that this analysis relies on a data set defined by the IBVD.

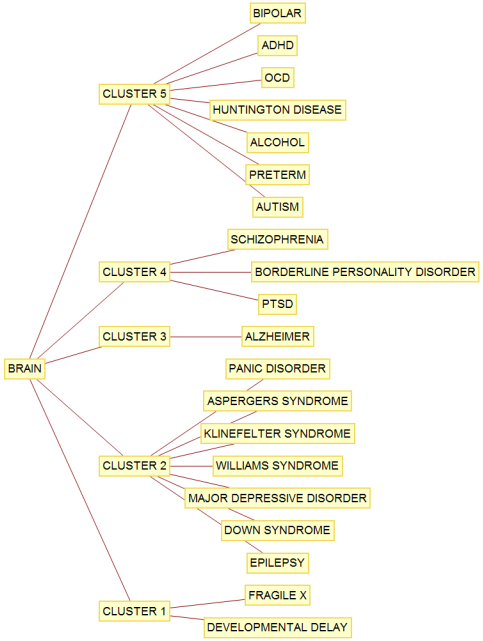



**Figure 7 This graph summarizes the modular organization of disorders in the human brain. It is the result of applying cluster analysis to duplicate mathematical markers hierarchically (Figures 3-6).**

Since mathematical markers identify patterns known to be highly conserved in both normal and abnormal settings (Bolender, 2011-2014), finding a similar pattern here for the disorders of the brain comes as no surprise. The results of the analysis remind us that phenotypes consist of specific collections of modules, many of which can exist in more than one disorder. Remember that in addition to this collection of shared markers, another level of complexity exists in that the disorders also carry individual sets of unique markers (Bolender, 2014).

In disorders of the brain, sharing identical markers (modules) appears to be a universal property of the disease process. This may signal the existence of a common origin for some disorders, followed by specialization to become a specific disorder. The five clusters shown in Figure 3, for example, may be operating under such a principle. If one identifies a pattern of markers common to the disorders of a given group, one also knows the names of the parts and the extent to which they are in play. Such information may prove helpful in a variety of clinical settings.

Consider, for example, the relationship of ADHD to OCD (Figure 8). Notice that more than half of the markers in the OCD cluster are being shared with those in the ADHD cluster. Does this mean that OCD shares a similar history – at some level – with ADHD? Supposedly, these phenotypes are showing us the downstream products of gene expression. If we work out the patterns for both disorders by stacking data sets over time vertically, then at some point they will begin to overlap. The branching point tells us when, where, and maybe how the disorders diverge to become specific disorders. Alternatively, of course, AHDH and OCD may simply be two variations of the same disorder. The point to be made is that mathematical markers behave like genes in that they allow us to work out the origins of disorders - objectively.

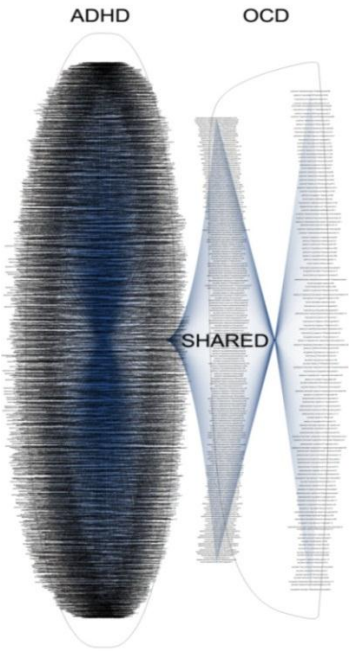



**Figure 8 The clusters formed by ADHD and OCD show a strong sharing of markers. Notice in this figure that mathematical markers – identified as alpha-numeric strings - replace the dots shown in Cluster 5.1.2 of Figure 5.**

## Schizophrenia – A Tree with Many Branches

When we use the duplicates database to look at disorders of the brain through the lens of schizophrenia, the result is Table 1. Notice that most of the disorders listed share a surprisingly large number of markers with schizophrenia. This suggests that schizophrenia, whose structure includes a well-defined set of markers – may represent the sum of many disorders or it may serve as a principle source thereof. In any case, schizophrenia appears to be the most invasive disorder.

Consider our options going forward. If the expression of a disorder requires a cadre of key markers, we can identify them with a duplicates database. In turn, we can combine this information with that of other databases to track the precursors back to the genome where sequences can be located and repaired. By identifying markers shared by many disorders, we increase the likelihood of solving not just one disorder at a time, but several.

**Table 1 The duplicates database contains mathematical markers (A), shared with either schizophrenia (B) or with other disorders (C). Notice that most disorders share a high percentage of their markers with schizophrenia (D).**

| MATHEMATICAL MARKERS (MM) SHARED WITH SCHIZOPHRENIA (B/A %) | | | | |
|---|---|---|---|---|
| | A | B | C | D |
| DISORDER (D) | MM(D) | MM(D)=MM(S) | MM(D)&MM(S)=OTHER | B/A (%) |
| | | | | |
| AFFECTIVEPSYCHOSIS | 10 | 9 | 1 | 90.0% |
| DOWNSYNDROME | 134 | 118 | 16 | 88.1% |
| BORDERLINEPERSONALITY | 625 | 487 | 625 | 77.9% |
| PANICDISORDER | 210 | 163 | 47 | 77.6% |
| BIPOLAR | 3339 | 2571 | 768 | 77.0% |
| EPILEPSY | 172 | 130 | 42 | 75.6% |
| ADHD | 1577 | 1177 | 400 | 74.6% |
| MAJORDEPRESSIVEDISORDER | 2112 | 1553 | 559 | 73.5% |
| PTSD | 426 | 309 | 117 | 72.5% |
| AUTISM | 507 | 360 | 147 | 71.0% |
| ALZHEIMER | 1967 | 1385 | 582 | 70.4% |
| ASPERGERSSYNDROME | 424 | 296 | 128 | 69.8% |
| KLINEFELTERSYNDROME | 46 | 30 | 16 | 65.2% |
| ALCOHOL | 124 | 76 | 48 | 61.3% |
| HUNTINGTONDISEASE | 23 | 11 | 12 | 47.8% |
| OCD | 54 | 22 | 32 | 40.7% |
| WILLIAMSSYNDROME | 22 | 8 | 14 | 36.4% |
| VELOCARDIOFACIAL | 489 | 132 | 357 | 27.0% |
| PRETERM | 26 | 2 | 24 | 7.7% |
| SCHIZOPHRENIA (S) | 3578 | | | |

As an example, Figure 9 illustrates the sharing arrangement between schizophrenia and Down's syndrome. It shares 88.1% of its 132 markers with schizophrenia.

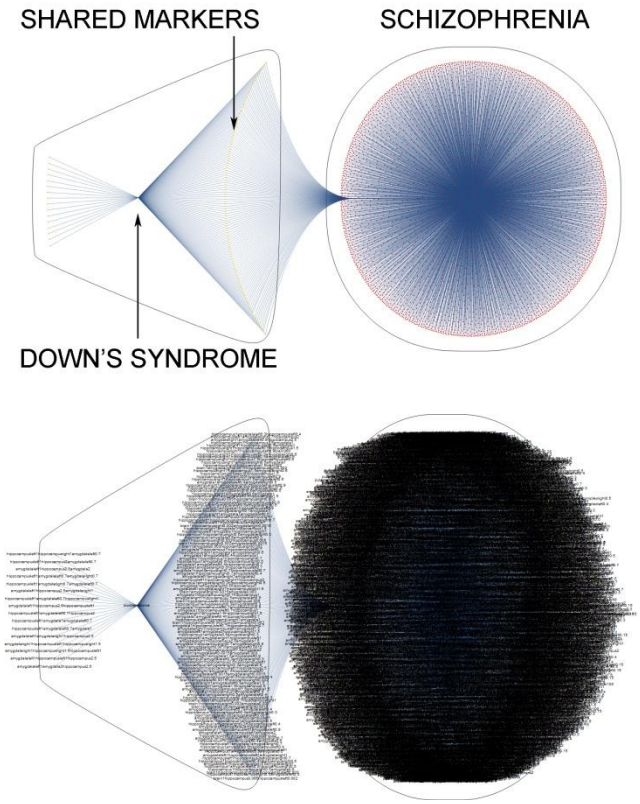



**Figure 9 Down's syndrome shares all but 16 of its mathematical markers with schizophrenia. The remaining 16 markers are shared with other disorders. Increase the magnification as needed.**

## Sharing Markers and Symptoms

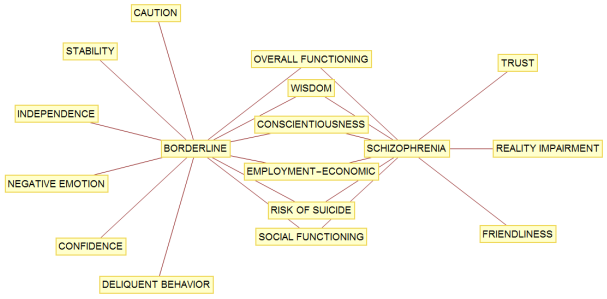
A phenotype can expresses itself structurally, functionally, and behaviorally. Since both mathematical markers and symptoms serve to identify the same disorders, we can look at the relationship of one to the other by stacking the individual plots vertically.

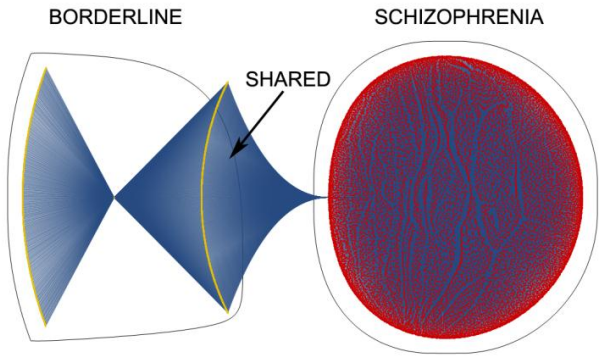
We begin with a table of symptoms for a panel of twelve disorders (Table 2). The first thing to notice is that the disorders have so many symptoms in common that making the correct diagnosis is downright challenging. Nonetheless, we have enough preliminary information to compare symptoms to mathematical markers.

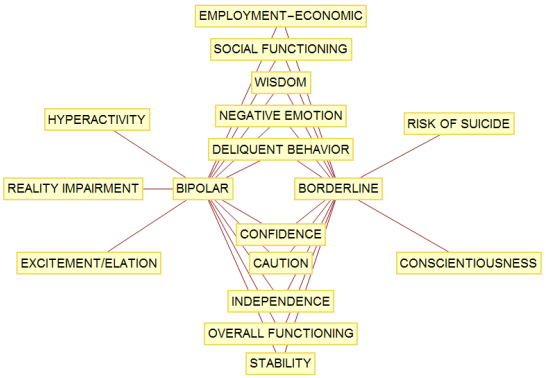
**Table 2 The table identifies symptoms for various disorders as impairments. Given the subjective nature of identifying impairments and the fact that a given impairment applies to many different disorders, making a differential diagnosis requires a major effort (Adapted from Internet Mental Health © 1995-2015 Phillip W. Long, M.D.).**

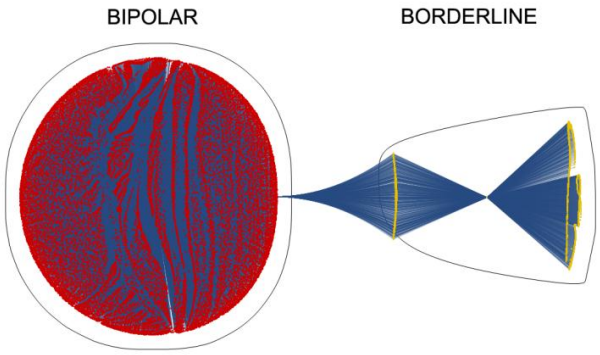
| IMPAIRED BY DISORDER (Symptoms) | ADHD | ALCOHOL | ALZHEIMER | ASPERGERS | AUTISM | BIPOLAR | BORDERLINE PD | MAJOR DD | OCD | PANIC DISORDER | PTSD | SCHIZOPHRENIA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PHYSICAL HEALTH | | X | | | | | | X | X | X | X | |
| SOCIAL FUNCTIONING | X | X | X | X | X | X | X | X | | | | X |
| EMPLOYMENT-ECONOMIC | X | X | X | X | X | X | X | X | X | X | X | X |
| DELIQUENT BEHAVIOR | | X | | | | X | X | | | | | |
| SUBSTANCE ABUSE | | X | | | | | | | | | | |
| PHOBIA/PANIC/ OBSESSION | | | | | | | | | X | X | X | |
| NEGATIVE EMOTION | | X | X | | | X | X | X | X | | X | |
| RISK OF SUICIDE | | X | | | | | X | X | | X | | X |
| EXCITEMENT/ ELATION | | | | | | X | | | | | | |
| HYPERACTIVITY | X | | | | | X | | | | | | |
| REALITY | | X | | | | X | | X | | | | X |
| TRUST | | | | | | | | | | | | X |
| FRIENDLINESS | | | | X | X | | | | | | | X |
| JUSTICE | | X | X | | | | | | | | | |
| WISDOM | X | X | X | X | X | X | X | X | | | | X |
| CAUTION | X | | | | | X | X | | | | | |
| STABILITY | | | X | | | X | X | | | | | |
| CONSCIENTIOUSNESS | | | | | | | X | | X | | | X |
| CONFIDENCE | | | | | | X | X | X | | | | |
| INDEPENDENCE | | | X | | | X | X | | | X | | |
| OVERALL FUNCTIONING | X | | | X | X | X | X | X | X | X | X | X |

Figure 10 plots the symptoms above and the markers below for pairs of disorders that were identified as sharing the same markers. This reveals the extent to which similar symptoms share similar markers. Notice that bipolar and borderline disorders share 10 of the 15 symptoms, but far fewer elsewhere.



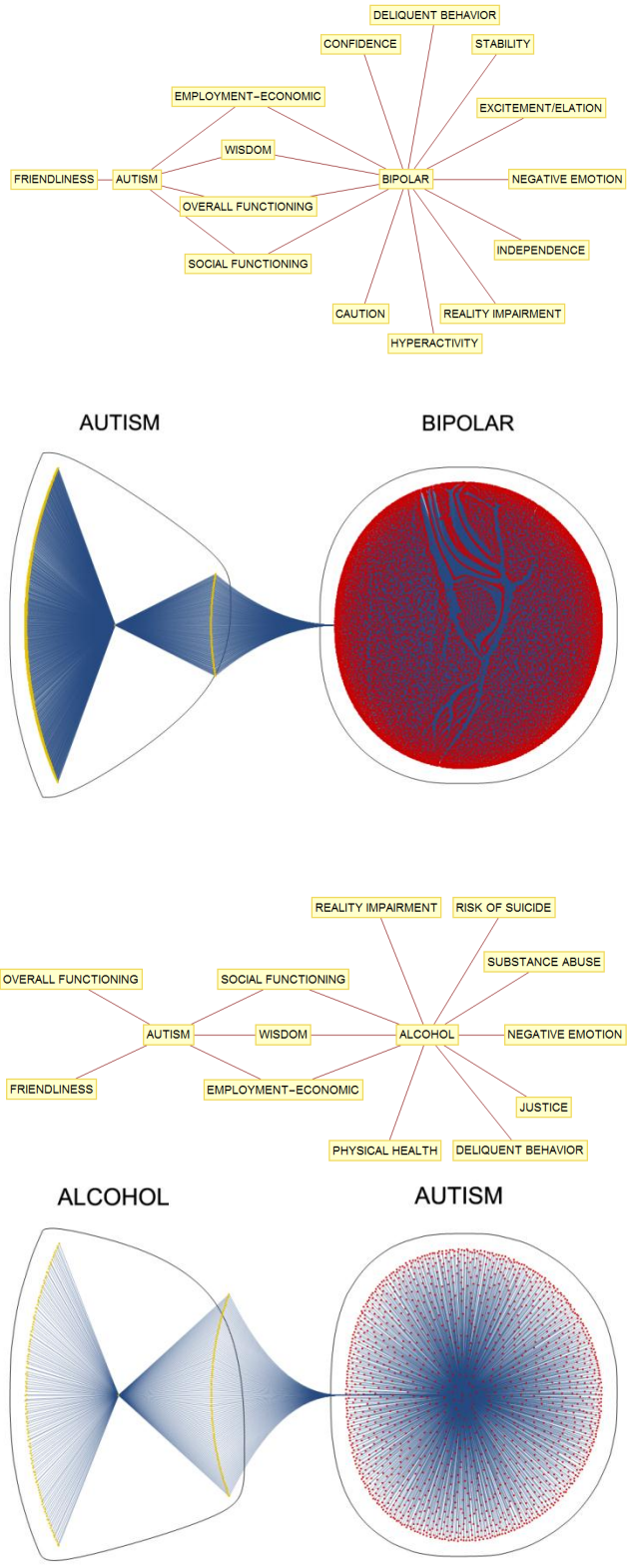



**Figure 10 Linear (above) and community (below) plots illustrate the extent to which different disorders share similar symptoms and markers.**

## Frequency of Shared Markers

The duplicates database suggests that some markers appear to be more highly conserved in that they exist with a greater frequency – typically in association with a wide range of published studies. This may be the consequence of the contents of the IBVD or it may point to a biological generalization. Highly conserved markers, for example, suggest the presence of common triggers needed to initiate or maintain a disorder. If true, then we would expect to find the same cast of markers occurring in multiple disorders. Of course, a simpler explanation might be that diseases are like recipes wherein some ingredients simply appear more often than others do. Although we can extract the frequency distributions of markers from the IBVD database, explaining why they exist continues to remain an open question – at least for now.

**The Method:** Starting with the database of duplicate markers, we can collect markers according to their frequency of occurrence, ranging from >2 (read 3 or more copies) to >11 (read 12 or more copies). To focus on just the markers in play, we can apply a filter to the database that limits each disorder to a single copy of a given marker.

**Displaying the Results:** In Figures 11 to 17, **t**he CommunityGraphPlots illustrate the relationship of markers to disorders (above) and parts to disorders (below). The figure legends identify the parts in play, which were derived from the mathematical markers (above).

**Duplicates:** The clusters, which form around disorders, diminish in number from four to one as the number of duplicates increases from more than 3 copies to more than 12 (Figures 11-17).

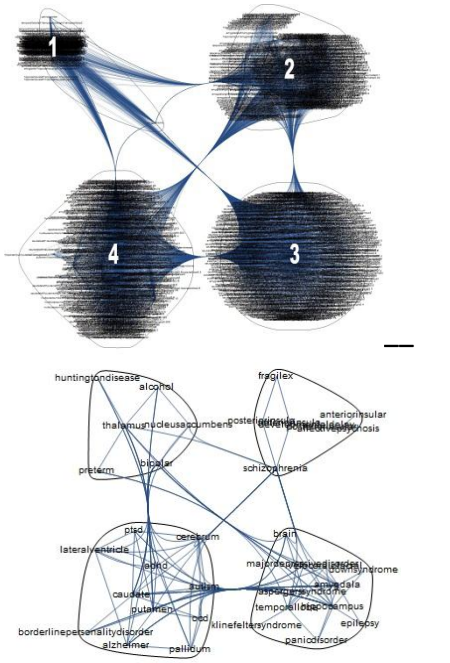



**Figure 11 Duplicates >3. A CommunityGraphPlot identifies mathematical markers and disorders in a community (cluster) more similar to each other than to those in another community. Above: In a clockwise direction, markers and disorders cluster around the brain-thalamus (1), anterior and posterior insula (2), nucleusaccumbens-cerebrum-pallidum-caudate-putamen-lateral ventricle (3), and hippocampus-amygdala-temporal lobe (4). Below: Although plots of markers vs. parts display somewhat different patterns, the names of the parts are easier to read.**

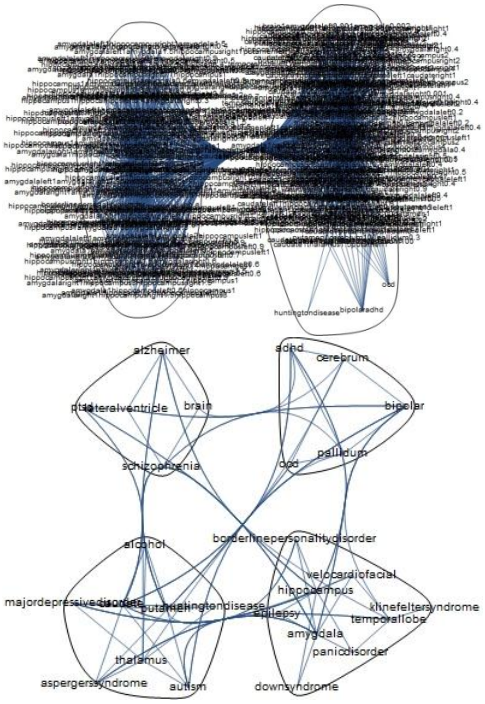



**Figure 12 Duplicates >5. The disorders cluster around the thalamus-nucleusaccumbens, anterior and posterior insula, brain-amygdala-cerebrum-caudate-putamen-pallidum.**

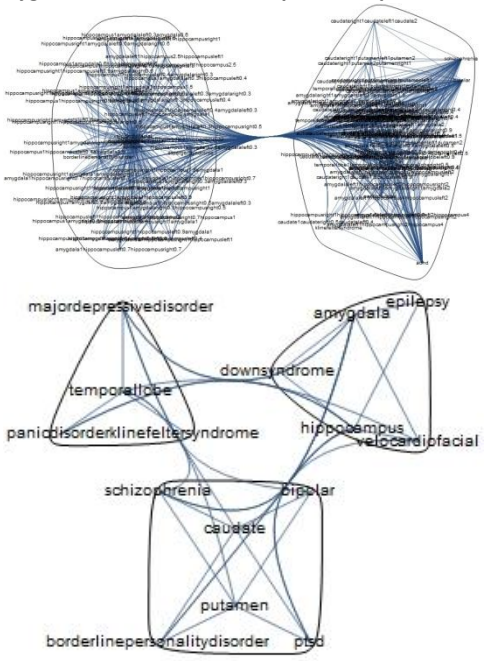



**Figure 13 Duplicates >6. The disorders cluster around the lateral ventricle – brain, cerebrum-pallidum, hippocampus- temporal lobe, and putamen-caudate-thalamus.**

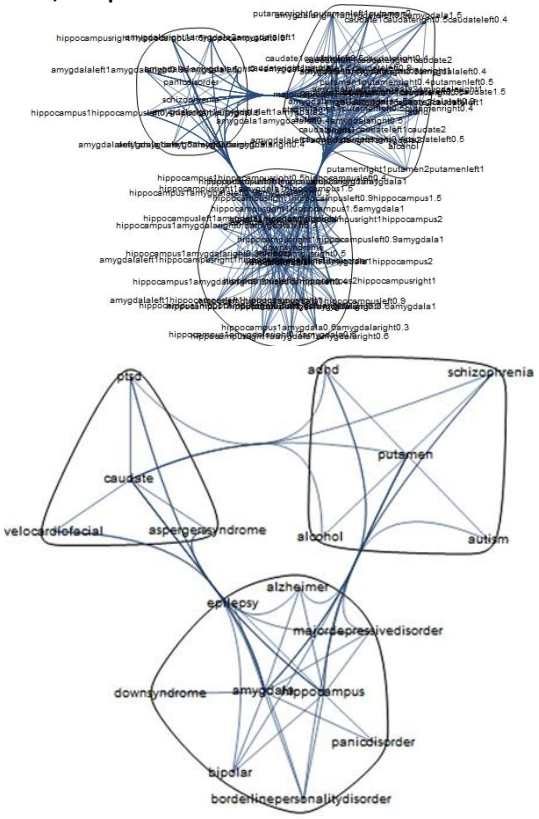


**Figure 14 Duplicates >7. The disorders cluster around the putamen, caudate, and hippocampus-amygdala.**

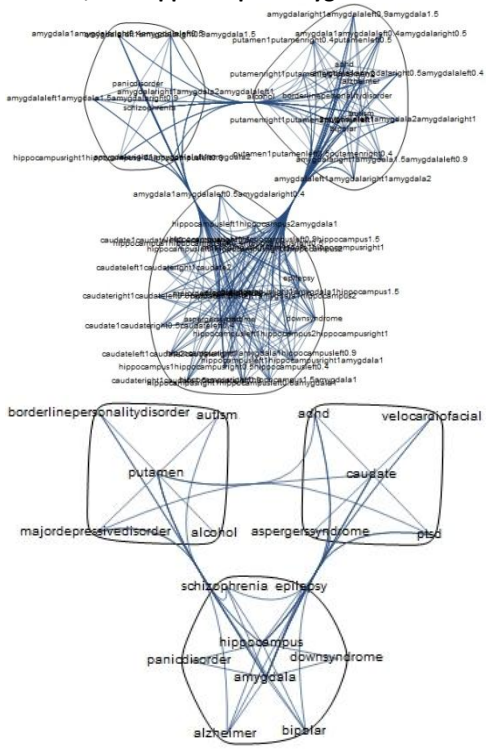



**Figure 15 Duplicates >8. The disorders cluster around the putamen, caudate, and hippocampus-amygdala.**

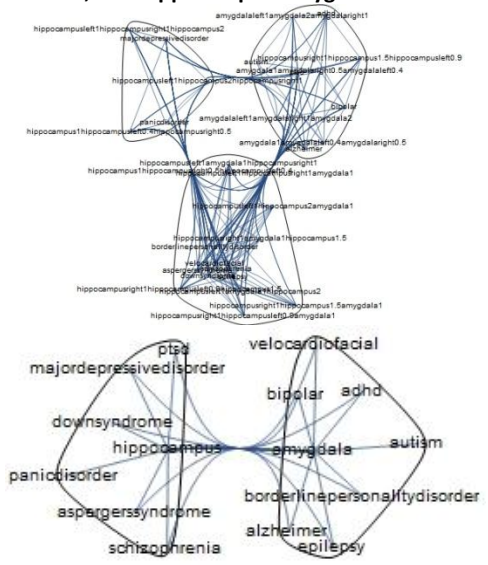
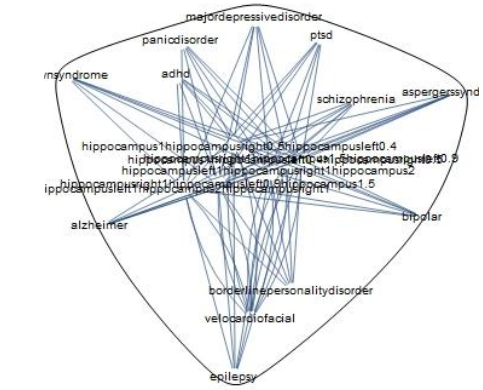



**Figure 16 Duplicates >9. The disorders cluster around the hippocampus and amygdala.**



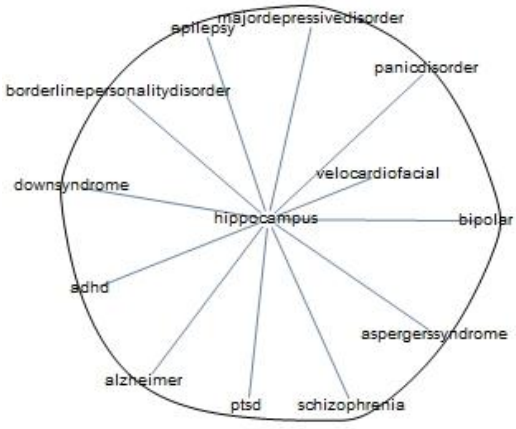




**Figure 17 Duplicates >11. The disorders cluster around the hippocampus. It is the one part of the brain involved most often in the disease process.**

**The Major Players:** Table 3 summarizes the findings of Figures 11-17. Notice in the table that most of the disorders depend importantly on abnormalities in just five parts – the hippocampus, amygdala, putamen, caudate, and temporal lobe. The table also reveals an extensive sharing of parts by the disorders. Taken together, Figures 11-17 and Table 3 continue to suggest a modular origin for disorders of the human brain.

**Table 3 The table shows the relationship of parts to disorders arranged according to the number of duplicate markers. These data suggest that a relatively small number of parts play a disproportionately large role in the disease process.**

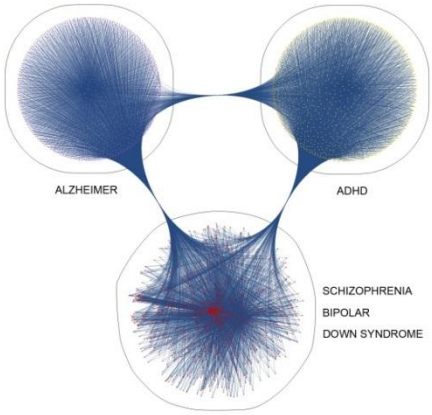
| | >11 | >9 | | >8 | | | >7 | | | >6 | | | >5 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | HIPPOCAMPUS | HIPPOCAMPUS | AMYGDALA | HIPP-AMYG | PUTAMEN | CAUDATE | HIPP-AMYG | PUTAMEN | CAUDATE | HIPP-AMYG | PUTA-CAUD | TEMPLOBE | HIPP-AMYG-TEMPLO | PALLI-CEREBR | BRAIN-LATVEN | PUTA-CAUD-THAL |
| adhd | X | | X | | | X | | X | | | | | | X | | |
| alcohol | | | | | X | | | X | | | | | | | | X |
| alzheimer | X | | X | X | | | X | | | | | | | | X | |
| aspergers-syndrome | X | X | | | | X | | | X | | | | | | | X |
| autism | | | X | | X | | | X | | | | | | | | X |
| bipolar | X | | X | X | | | X | | | | X | | | X | | |
| borderline-personality-disorder | X | | X | | X | | X | | | | X | | X | | | |
| down-syndrome | X | X | | X | | | X | | | X | | | X | | | |
| epilepsy | X | | X | X | | | X | | | X | | | X | | | |
| huntington-disease | | | | | | | | | | | | | | | | X |
| klinefelter-syndrome | | | | | | | | | | | | X | X | | | |
| major-depressive-disorder | X | X | | | X | | X | | | | | X | | | | X |
| ocd | | | | | | | | | | | | | | X | | |
| panic-disorder | X | X | | X | | | X | | | | | X | X | | | |
| ptsd | X | X | | | | X | | | X | | X | | | | X | |
| schizophrenia | X | X | | X | | | | X | | X | X | | | | X | |
| velocardiofacial-syndrome | X | | X | | | X | | | X | | | | X | | | |

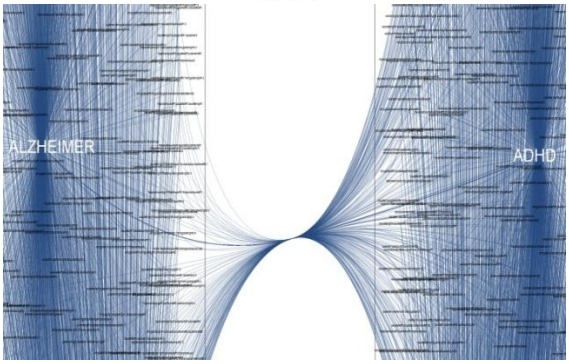
**Designing Disorders:** When biology constructs a disorder, it changes a surprisingly large number of parts and connections. Using mathematical markers, we can identify many of these changes within and across disorders.

Consider ADHD as an example. In Table 4, we can identify four disorders (alzheimer, bipolar, down syndrome, and schizophrenia) that share similar parts (hippocampus, amygdala, caudate, and putamen) with ADHD. These five disorders, which share 10,031 duplicate markers, are illustrated in Figures 18 to 20. Notice the extensive connectivity displayed by disorders sharing many of the same markers. This is exactly what we would expect to see in our data if biology is using a modular approach to building disorders.

**Table 4 Different disorders often share the same parts.**

| | >11 | >9 | | >8 | | | >7 | | | >6 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | HIPPOCAMPUS | HIPPOCAMPUS | AMYGDALA | HIPP-AMYG | PUTAMEN | CAUDATE | HIPP-AMYG | PUTAMEN | CAUDATE | HIPP-AMYG | PUTA-CAUD | TEMPLOBE |
| adhd | X | | X | | | X | | X | | | | |
| alzheimer | X | | X | X | | | X | | | | | |
| bipolar | X | | X | X | | | X | | | | X | |
| down-syndrome | X | X | | X | | | X | | | X | | |
| schizophrenia | X | X | | X | | | | X | | X | X | |



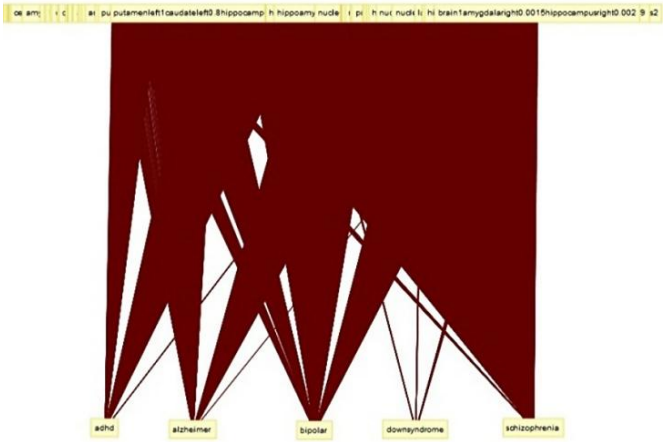




**Figure 18 TOP & MIDDLE: The disorders listed in Table 4 use many of the same abnormal markers (modules). BOTTOM: Mathematical markers containing four different parts (hippo-**

**campus, amygdala, caudate and putamen) from the five different disorders of Table 4 show massive connectivity and widespread sharing.**

Figure 19 selects just three mathematical markers from the data set of Figure 18 to illustrate the details of the sharing and connectivity. Figure 20 illustrates the connectivity of a larger set of markers and disorders.

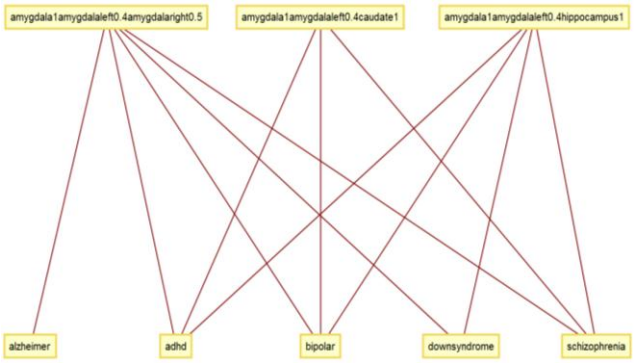



**Figure 19 The five disorders of Table 4 are plotted against just three markers to illustrate the complexity of sharing. Notice that the same, highly specific markers appear in what we identify as distinctly different disorders.**

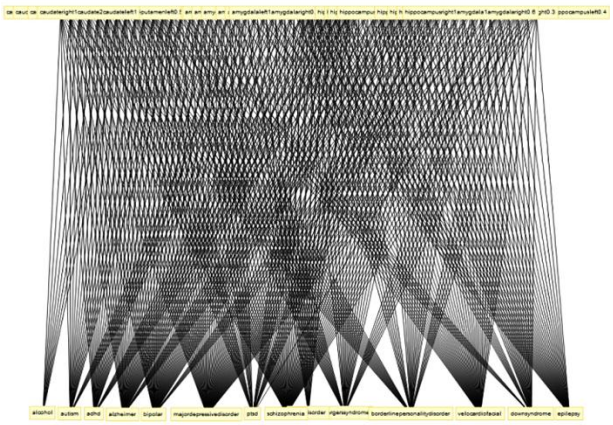


**Figure 20 By expanding the number of disorders and markers in play, we can begin to appreciate the complexity of the disease process.**

**Duplicate Databases:** The software package includes copies of two Access databases – one for sorting and filtering duplicate markers (MRI_T_Dups_2015.accdb) and another for selecting scripts to plot patterns with Mathematica (MRI_Markers_Plot.accdb).

Figure 21 displays a portion of the duplicates database table to show that the same mathematical marker (Field1) occurs as a duplicate locally in the same citation (Field2; c, e) and globally in multiple citations (Field1 and Field2). Reproducibility, an often rare commodity in the literature, becomes commonplace when we express MRI data as mathematical markers – even when we limit the number of markers to two per paper (one each for control and experimental time points).

MRI_T_BIG_DUPES-F1-C-E-REMOVE-INDIVIDUAL-DUPS | Find duplicates-FIELD1AND3 | Find duplicates for MRI_T_BIG_DUPES-F1-C-E-REMOVE-INDIVIDUAL-DUP

| Field1 | Field2 | Field3 | Field4 | Field5 | Field6 | Field7 |
|---|---|---|---|---|---|---|
| amygdala1amygdalaleft0.4amygdalaleft0.4 | 655 e | | amygdala | amygdalaleft | amygdalaleft | 2677 |
| amygdala1amygdalaleft0.4amygdalaleft0.4 | 655 c | | amygdala | amygdalaleft | amygdalaleft | 2595 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 42 e | | amygdala | amygdalaleft | amygdalaright | 6.9 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 42 c | | amygdala | amygdalaleft | amygdalaright | 7.5 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 50 e | | amygdala | amygdalaleft | amygdalaright | 4.52 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 50 c | | amygdala | amygdalaleft | amygdalaright | 4.9 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 78 e | | amygdala | amygdalaleft | amygdalaright | 4.26 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 78 c | | amygdala | amygdalaleft | amygdalaright | 4.73 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 126 e | | amygdala | amygdalaleft | amygdalaright | 2.59 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 126 c | | amygdala | amygdalaleft | amygdalaright | 2.41 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 178 e | | amygdala | amygdalaleft | amygdalaright | 3.7 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 178 c | | amygdala | amygdalaleft | amygdalaright | 4.18 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 308 e | | amygdala | amygdalaleft | amygdalaright | 4.09 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 308 c | | amygdala | amygdalaleft | amygdalaright | 3.9 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 314 e | | amygdala | amygdalaleft | amygdalaright | 4.15 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 314 e | | amygdala | amygdalaleft | amygdalaright | 4.17 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 482 e | | amygdala | amygdalaleft | amygdalaright | 3.02 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 482 c | | amygdala | amygdalaleft | amygdalaright | 3.16 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 484 e | | amygdala | amygdalaleft | amygdalaright | 5.19 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 484 c | | amygdala | amygdalaleft | amygdalaright | 4.55 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 486 e | | amygdala | amygdalaleft | amygdalaright | 3.66 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 486 c | | amygdala | amygdalaleft | amygdalaright | 4.32 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 618 e | | amygdala | amygdalaleft | amygdalaright | 3441.57 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 618 c | | amygdala | amygdalaleft | amygdalaright | 3222.55 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 621 e | | amygdala | amygdalaleft | amygdalaright | 4.56 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 621 c | | amygdala | amygdalaleft | amygdalaright | 4.6 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 626 e | | amygdala | amygdalaleft | amygdalaright | 3.71 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 626 c | | amygdala | amygdalaleft | amygdalaright | 3.97 |
| amygdala1amygdalaleft0.4amygdalaright0.5 | 635 e | | amygdala | amygdalaleft | amygdalaright | 3890 |

**Figure 21. Field1 of the MRI_T_Dups_2015 database illustrates the local and global persistence of the same marker in control (c) and experimental (e) settings, as evidenced by the same marker associated with one or several citation numbers (Field2).**

Figure 22 demonstrates the persistence of markers across disorders of the brain. It illustrates the scripts (Field3) that were used to generate the plots with Mathematica.

E-NODUPS-DISEASE-MRI_T_BIG-173787-MM-TOTAL-SHARED-MARKERS | Find duplicates for E-NODUPS-DISEASE-MRI_T_BIG-173787-MM-TOTAL-S

| Field1 | Field2 | Field3 |
|---|---|---|
| hippocampus1putamen1cerebralcortex70 | ocd | "hippocampus1putamen1cerebralcortex70"->"ocd", |
| hippocampus1putamen1cerebralcortex70 | schizophrenia | "hippocampus1putamen1cerebralcortex70"->"schizophrenia", |
| hippocampus1putamen1cerebrum100 | adhd | "hippocampus1putamen1cerebrum100"->"adhd", |
| hippocampus1putamen1cerebrum100 | ocd | "hippocampus1putamen1cerebrum100"->"ocd", |
| hippocampus1putamen1cerebrum100 | schizophrenia | "hippocampus1putamen1cerebrum100"->"schizophrenia", |
| hippocampus1putamen1cerebrum150 | adhd | "hippocampus1putamen1cerebrum150"->"adhd", |
| hippocampus1putamen1cerebrum150 | bipolar | "hippocampus1putamen1cerebrum150"->"bipolar", |
| hippocampus1putamen1cerebrum150 | schizophrenia | "hippocampus1putamen1cerebrum150"->"schizophrenia", |
| hippocampus1putamen1hippocampusleft0.4 | adhd | "hippocampus1putamen1hippocampusleft0.4"->"adhd", |
| hippocampus1putamen1hippocampusleft0.4 | alzheimer | "hippocampus1putamen1hippocampusleft0.4"->"alzheimer", |
| hippocampus1putamen1hippocampusleft0.4 | bipolar | "hippocampus1putamen1hippocampusleft0.4"->"bipolar", |
| hippocampus1putamen1hippocampusleft0.4 | borderlinepersonality | "hippocampus1putamen1hippocampusleft0.4"->"borderlinepersonalit |
| hippocampus1putamen1hippocampusleft0.4 | majordepressivedisor | "hippocampus1putamen1hippocampusleft0.4"->"majordepressivedisc |
| hippocampus1putamen1hippocampusright0.5 | adhd | "hippocampus1putamen1hippocampusright0.5"->"adhd", |
| hippocampus1putamen1hippocampusright0.5 | alzheimer | "hippocampus1putamen1hippocampusright0.5"->"alzheimer", |
| hippocampus1putamen1hippocampusright0.5 | bipolar | "hippocampus1putamen1hippocampusright0.5"->"bipolar", |
| hippocampus1putamen1hippocampusright0.5 | borderlinepersonality | "hippocampus1putamen1hippocampusright0.5"->"borderlinepersonal |
| hippocampus1putamen1hippocampusright0.5 | majordepressivedisor | "hippocampus1putamen1hippocampusright0.5"->"majordepressivedis |
| hippocampus1putamen1hippocampusright0.5 | schizophrenia | "hippocampus1putamen1hippocampusright0.5"->"schizophrenia", |
| hippocampus1putamen1nucleusaccumbens0.15 | adhd | "hippocampus1putamen1nucleusaccumbens0.15"->"adhd", |
| hippocampus1putamen1nucleusaccumbens0.15 | bipolar | "hippocampus1putamen1nucleusaccumbens0.15"->"bipolar", |
| hippocampus1putamen1nucleusaccumbens0.15 | majordepressivedisor | "hippocampus1putamen1nucleusaccumbens0.15"->"majordepressivec |
| hippocampus1putamen1nucleusaccumbens0.15 | schizophrenia | "hippocampus1putamen1nucleusaccumbens0.15"->"schizophrenia", |
| hippocampus1putamen1pallidum0.4 | adhd | "hippocampus1putamen1pallidum0.4"->"adhd", |
| hippocampus1putamen1pallidum0.4 | alzheimer | "hippocampus1putamen1pallidum0.4"->"alzheimer", |
| hippocampus1putamen1pallidum0.4 | bipolar | "hippocampus1putamen1pallidum0.4"->"bipolar", |
| hippocampus1putamen1pallidum0.4 | schizophrenia | "hippocampus1putamen1pallidum0.4"->"schizophrenia", |

**Figure 22. Field3 of the MRI_Markers_Plot database illustrates the scripts used by Mathematica to produce CommunityGraphPlots. Once selected, a set of scripts can be copied and pasted into a Mathematical template wherein patterns can be analyzed and viewed graphically.**

# DISCUSSION

Biology is in the business of creating two types of brains, one we identify as normal and the other abnormal. When we populate a parallel complexity with duplicate markers coming from 21 disorders, we begin to understand what it takes to be an abnormal brain. Using this composite brain as our starting point, we can ask - and answer – a variety of questions related to the disease process.

The vehicle for our questioning becomes the triplet mathematical marker, a quantitative unit of complexity consisting of three named parts accompanied by the ratio of their volumes. Millions of these alphanumeric strings begin to define the phenotype of an abnormal brain as a complexity. A remarkable, but little appreciated property of this phenotype is that everything connects mathematically within and often between disorders. As such, we have at our disposal the means to pursue and understand the strategy biology uses to create these disorders.

**Communities and Modules:** Using the duplicates database, we wanted to work out the relationship of disorders to mathematical markers. By analyzing the entire set of markers with cluster analysis, communities of markers and disorders appeared according to the strength of their affinities. The resulting phenotypic patterns were both striking and widespread, extending all the way down to the final pairs of disorders (Figures 3-6).

Notice the simplicity of the approach. We began with 21 identifiable disorders and then homogenized the data set by forming mathematical markers - such that biology's rules remained encapsulated in the markers. Starting with this "homogenate", we used these rules to separate out the disorders and to identify those disorders that were most closely related.

By allowing the disorders to interact freely with the markers, we can see that biology uses a highly conservative strategy when activating the disease process. It constructs modules of disorder and then reuses them repeatedly in new configurations to generate additional disorders with new properties. Such is the pattern displayed repeatedly in the figures. This proclivity for sharing and mixing modules appears central to biological systems, which are fundamentally modular in design. Recall that the design of organs, cells, organelles, molecules, and genes all follow a similar strategy.

**Duplicate Markers:** The extensive sharing of markers becomes a clinically relevant piece of information because it tells us where synergies might exist when we are looking for root causes or for treatment protocols that might be equally effective in closely related disorders.

**Schizophrenia:** Given the results presented in Table 1 for schizophrenia and the widespread sharing of markers seen throughout the report, it appears that many disorders of the brain are reading from the same playbook. This effectively reduces the study of these disorders to a giant mathematical puzzle, the solution to which becomes a function of the number of pieces available.

**Markers and Symptoms:** We wanted to see if a clear link existed between mathematical markers and symptoms. Such was not the case. The relationships identified in Table 2 limited our ability to identify distinct patterns because the disorders shared many of the same symptoms. Moreover, it seems more likely that symptoms represent emergent properties coming from many interacting factors, not just from markers.

**Major Players:** Key players represent those parts of the brain that occur in disorders with the greatest frequency. Using the SQL scripts available in the Access database, markers selected according to their frequency of occurrence were plotted. As the number of duplicate copies increased from 3+ to 12+, the plots showed how the disorders, markers, and parts cluster. When the number of duplicate copies became $\geq$12, the hippocampus emerged as the key player, appearing in 12 of the 21 disorders (Figure 17). In addition to the abnormal hippocampus, 6 of these 12 disorders showed an abnormal amygdala, 3 an abnormal caudate, and 2 an abnormal putamen (Table 4). Taken together, the exercise identifies the

hippocampus as playing a role in more than half of the disorders, with lesser roles going to a relatively small number of parts.

The major players identified as parts above combine to form the markers listed in Table 5. These markers, which occur in at least 9 different disorders, represent modules being used and reused by biology as basic building blocks of the disease process.

**Table 5 Each marker listed occurs in at least nine different disorders. Markers having the same Y and Z components, but in reverse order, were deleted.**

| MARKERS SHARED BY >8 DISORDERS |
|---|
| amygdala1amygdalaleft0.4amygdalaright0.5 |
| amygdala1amygdalaright0.4amygdalaleft0.5 |
| amygdalaleft1amygdala1.5amygdalaright0.9 |
| amygdalaleft1amygdala2amygdalaright1 |
| amygdalaleft1amygdalaright0.9amygdala1.5 |
| amygdalaright1amygdala2amygdalaleft1 |
| amygdalaright1amygdalaleft0.9amygdala1.5 |
| caudate1caudateleft0.4caudateright0.5 |
| caudateleft1caudate2caudateright1 |
| caudateright1caudateleft0.9caudate1.5 |
| hippocampus1hippocampusleft0.4hippocampusright0.5 |
| hippocampusleft1amygdala1hippocampus2 |
| hippocampusleft1amygdala1hippocampusright1 |
| hippocampusleft1hippocampus2hippocampusright1 |
| hippocampusleft1hippocampusright1amygdala1 |
| hippocampusright1amygdala1hippocampus1.5 |
| hippocampusright1amygdala1hippocampusleft0.9 |
| hippocampusright1hippocampus1.5hippocampusleft0.9 |
| hippocampusright1hippocampusleft0.9amygdala1 |
| putamen1putamenright0.4putamenleft0.5 |
| putamenright1putamenleft1putamen2 |

## Concluding Comments

In this report, we focused our attention on a parallel complexity consisting of just duplicate markers to study not just a single disorder, but rather a cohort thereof. Such an approach allowed us to hunt for common denominators of the disease process, which we found to include a modular format, shared markers, and a playing field consisting of many interconnected parts.

Taken together, these results suggest that a strategy based on studying the abnormal brain rather than focusing on individual disorders may prove to be a more effective way of advancing our understanding of the disease process. Moreover, this approach raises an important question as to our perception of these disorders. What we choose to label as disorders of the brain, biology may see as a natural consequence of obeying its cardinal rule of adapting and evolving.

Biology routinely nudges us - as a species - toward the edge of chaos because its charge is to push the envelope of possibilities - relentlessly. Such appears to be the nature of survival. Given the advances in molecular biology, however, we are rapidly becoming the primary driving force of our own evolution. In effect, we are putting our future literally into our own hands. One cannot even begin to imagine the unintended consequences of this new reality. Since we can readily make changes to our DNA, we will be changing the initial conditions without a parallel ability to detect and evaluate the phenotypic consequences. If we decide to play this complexity game without knowing the rules, then the best we can do is to remain hopeful that this will not lead to worst-case scenarios. Herein we find a compelling argument for developing literature databases similar to the IBVD – sooner than later.

## REFERENCES

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

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.