

Gene Intersection Method for Understanding Complex Disease

Example One: Tourette's Related Conditions

L. Van Warren

Genes Shared in Tourette Related Disorders

	Anxiety	Bipolar Disorders	Tourette	OCD	ADHD	Depressive Disorders	Tics	Phobic Disorders	Learning Disorders	Language Disorders	Sleep Disorders
Anxiety	90	9	5	4	1	3	3	6	3	1	0
Bipolar	9	63	3	4	3	8	1	0	4	1	0
Tourette Syndrome	5	3	20	4	6	2	6	2	1	0	0
OCD	4	4	4	12	3	2	3	2	2	0	0
ADHD	1	3	6	3	20	2	2	2	3	0	0
Depressive Disorders	3	8	2	2	2	14	1	2	0	0	0
Tics	3	1	6	3	2	1	10	2	0	0	0
Phobic Disorders	6	0	2	2	2	2	2	8	0	0	0
Learning Disorders	3	4	1	2	3	0	0	0	29	1	0
Language Problems	1	1	0	0	0	0	0	0	1	10	0
Sleep Disorders	0	0	0	0	0	0	0	0	0	0	5

© 2006 L. Van Warren . Warren Design Vision . All Rights Reserved

Introduction

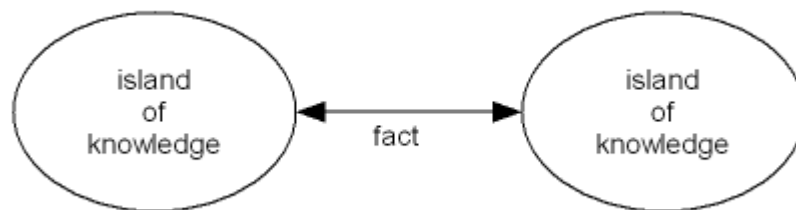
Diabetes, cancer and many psychiatric disorders are complex diseases. They operate in stark contrast to conditions of a single factor. For example, in sickle cell anemia, a single mutation in the hemoglobin coding gene causes suffering for millions. A complex example is the spectrum of Tourettes-related disorders. Patients suffering from one aspect of the disorder, often suffer from one or more companion disorders, at a frequency shown in the figure above. The mechanism for these remains a mystery.

In medieval times, physical illness was explained in vague and unsatisfactory terms. Metabolism and bodily function were seen to depend on vague or elemental influences such as earth, wind, fire, and water. A complete dissection of the living organism, and its companion sciences required several hundred years and was often carried on in secret. This exploration is just beginning at the genomic level. The result is an explosion in factual knowledge. This knowledge often contradicts intuitive explanations and traditions.

As recently as a century ago, conditions such as appendicitis were most often fatal, not for their complexity, but because of ignorance of cause and treatment. Anesthesia, sterile-field surgery, antibiotics and imaging techniques enhanced survival immensely. So too was the case for diabetics. In modern practice candidate diabetics can be given a glucose tolerance test. In this test, a measured amount of glucose is administered, and the patient's blood glucose is measured over a period of hours. Those who are positive for diabetes or hypoglycemia show characteristic deviations from a normal response curve. They are then treated with insulin and now with continuous monitoring solutions that restore a normal glucose response.

For the genes that cause mental illness however, we still live in a dark age. Although some specific conditions are well-addressed with medication, some are not. Selective serotonin reuptake inhibitors have improved the lives of many with depressive and anxiety disorders, yet there remains no direct measurement technique for intrasynaptic serotonin levels. Further, complex pathways mediate serotonin and other critical neurotransmitters. Unlike the glucose tolerance test, we have no method of direct inspection or continuous monitoring to keep patient's within a robust and normal range.

The advent of bioinformatics and the information-age is pulling back some of the curtains however. Key facts that link previously-separated islands of knowledge open new diagnostic and treatment opportunities..



For the last decade, the ubiquity of information is changing the practices we use to gain understanding.¹ The shotgun sequencing method that enabled the human genome project was initially criticized as a "fishing expedition".² Combinatorial chemistry techniques have suffered similar disdain. We live in the first generation of "hot and cold running knowledge". Those of us in this transition are rapidly adopting new thinking practices that exploit the explosion of genomic knowledge. The decisive moment comes at the instant that knowledge produces benefit for the individual. Knowledge, like all information, is functionally useless until the moment that it is applied.

In previous work, a theory of knowledge mapping was developed to visualize complex relationships in biological systems.³ This article describes a computational algorithm for finding genes common to related conditions. It is called, "Gene Intersection Method" or GIM for short. GIM can be used to design gene chips for clinical laboratory assessment. GIM prioritizes results such that genes of greatest influence are ranked highest.

History Repeats Itself

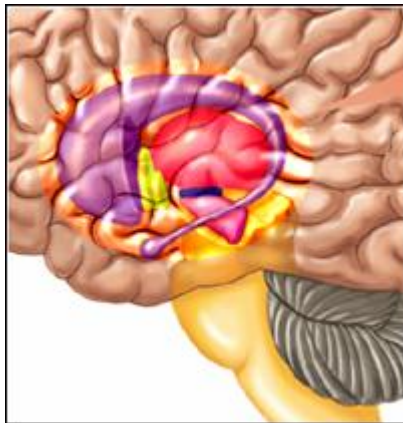
In popular culture, disorders of mind have suffered an obstruction in understanding similar to those of somatic illness of times past. Neuron biology and imaging techniques such as functional MRI, PET-scanning, and neural circuit mapping are changed this rapidly and profoundly.⁴ Brain pathologies at several decades of scale are being understood from a functional point of view. Molecular biologists, psychologists, and computer scientists are working from the bottom up enabling a new age of genomic understanding.

Disorders of Mind vs. Body

The terms, "mental illness", and "psychiatric disorder" conjure images of hopelessness and incurability. Gene-based assessment offers the possibility of better diagnosis and treatments.⁵ Treatments based on these have the potential to correct neurochemical disequilibrium of genetic origin. The first application of the gene intersection method will be for disorders of the mind. In this, they demonstrate significant opportunity.

Morphology vs. Biochemistry

Consider two views of the functional brain. One is the macroscopic architecture, the gross anatomy of neural circuitry.⁶ Obsessive-compulsive disorder (OCD) is often referenced as a "basal ganglia disorder".⁷ The basal ganglia are involved in a constant neuronal feedback loops that enable initiation of new tasks and maintenance of current tasks, for example holding a glass of wine while talking.⁶



Basal Ganglia – Macroscopically⁸

The other view is from the nanoscopic level of synaptic transmission via neurotransmitters and receptors. Complex gene-expression networks modulate an elegant balance involving multiple players on multiple tiers of regulation. In the small, tissue specific receptors and multiple neurotransmitters are regulated in magnificent orchestration. Fixing one's gaze or holding a glass are amazing feats of biological control requiring a hierarchy of orchestrated players.

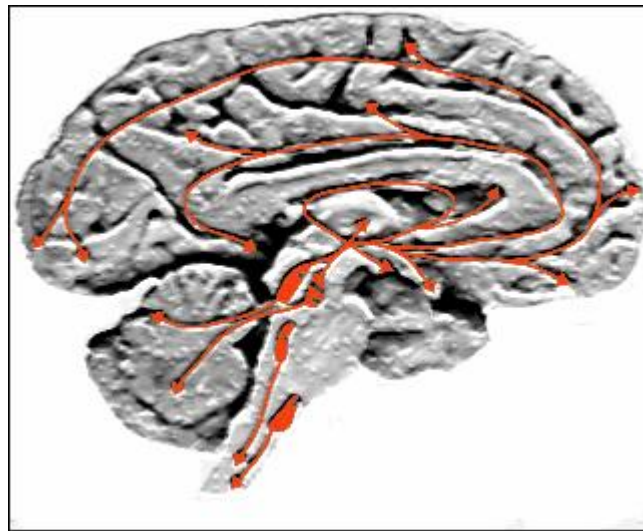
Small molecule neurotransmitters			Neuropeptide neurotransmitters
Type	Neurotransmitter	Postsynaptic effect	
Amino acids	Acetylcholine	Excitatory	Corticotropin releasing hormone
	Gamma aminobutyric acid (GABA)	Inhibitory	Corticotropin (ACTH)
	Glycine	Inhibitory	Beta-endorphin
	Glutamate	Excitatory	Substance P
	Aspartate	Excitatory	Neurotensin
Biogenic amines	Dopamine	Excitatory	Somatostatin
	Noradrenaline	Excitatory	Bradykinin
	Serotonin	Excitatory	Vasopressin
	Histamine	Excitatory	Angiotensin II

- after ⁹.

When players in this neurochemical orchestra are absent, broken, or out of tune, the result can be chaos for equilibrium of the mind. Attempts to change the level of one receptor or neurotransmitter affect a complex and nonlinear regulatory network. Because of the complexity of this network, it is necessary to determine all receptor and ligand players, not just for the idealized person, but for the afflicted individual. It is essential to measure what the regulator network is supposed to be doing versus what the network is actually doing. Adding to the mix is the dynamic and chaotic behavior of biological systems, especially the mind. Did you get enough sleep last night?

So we have two views. The first view is a top down architectural view whose exploration begins macroscopically. The second is a bottom up biochemical building block view. The root of the latter view is the gene expression level.

With both views, the “highway map” of neurochemical communication is being understood, enabling better assessment of disorders of mind.



Serotonin Highway – Simplified Macroscopic View¹⁰

Discussion of the First Example

At the clinical level, several disorders of mind appear to be functionally related. These disorders are Tourette’s syndrome, depression, obsessive-compulsive disorder (OCD), attention deficit-hyperactivity disorder (ADHD), tics, and sleep disorders.¹¹ Recent advances in gene ontology enable discovering the genes common to related conditions. This set of disorders seemed a good test case for gene intersection method.

The technique discussed here, Gene Intersection Method (GIM), explores genes common to related conditions. The method provides a mathematical definition of related conditions. It enables identification of disorders that are related by involvement of a common gene. GIM speeds the definition and construction of diagnostic and monitoring tests. The presumption is that accurate genetic diagnosis will result in treatments that are more effective. Eventually real-time monitoring will be enabled knowing **what** to monitor for specifically afflicted individuals, or improving quality of life for normal individuals.

Leckman and Cohen's excellent treatise on Tourette's Syndrome opens with the following definition: ¹¹

Tourette's syndrome is a developmental, neuropsychiatric disorder defined by persistent, motor and vocal tics, [...] obsessions, compulsions, and attentional difficulties. Often based upon a multigenerational, genetic predisposition...

They then go on to list eleven related conditions that are explored below using the gene intersection method and filtering the results to include only those entries for which gene description and locus is known.

Let us begin by using GIM to find genes common to pairs of disorders.

First Results of Gene Intersection (GIM) Method

Genes Common To Pairs of Tourette's Related Disorders									
ADHD	TOURETTES				126452	609678	191290	104210	126455
ADHD	BIPOLAR				126452		191290		
ADHD	ANXIETY				126452				
ADHD	DEPRESSION						191290		
ADHD	LEARNING-DIS							104210	162200
ADHD	OCD					609678			
ADHD	TICS					609678			
ANXIETY	TOURETTES				126452			126450	143100
ANXIETY	BIPOLAR	116790	182138	309550	126452	138040	122560	113505	607478
ANXIETY	OCD	116790	182138					113505	
ANXIETY	LEARNING-DIS	116790		309550					261600
ANXIETY	DEPRESSION		182138						
BIPOLAR	TOURETTES				126452		191290		606941
BIPOLAR	LEARNING-DIS	116790		309550					
BIPOLAR	OCD	116790	182138					113505	
BIPOLAR	DEPRESSION		182138				191290		300207
BIPOLAR	TICS								314850
DEPRESSION	TOURETTES						191290		
DEPRESSION	OCD		182138						
LEARNING-DIS	OCD	116790							
LEARNING-DIS	TOURETTES							104210	
OCD	TOURETTES					609678			604569
OCD	TICS					609678			
TICS	TOURETTES					609678			van-wdv 06
Gene Key	Name								Locus
116790	CATECHOL-O-METHYLTRANSFERASE								22q11.2
309550	FRAGILE X SYNDROME								Xq27.3
138040	GLUCOCORTICOID RECEPTOR								5q31
122560	CORTICOTROPIN-RELEASING HORMONE								8q13
607478	TRYPTOPHAN HYDROXYLASE								12q21.1
182138	SEROTONIN TRANSPORTER								17q11.1-q12
126452	DOPAMINE RECEPTOR D4								11p15.5
113505	BRAIN-DERIVED NEUROTROPHIC FACTOR								11p13
609678	SLITRK1, INTEGRAL MEMBRANE PROTEIN								13q31
191290	TYROSINE HYDROXYLASE								11p15.5
126455	DOPAMINE TRANSPORTER								5p15.3
104210	ALPHA-2A-ADRENERGIC RECEPTOR								10q24-q26
162200	NEUROFIBROMIN GENE								2p22-p21, 17q11.2
261600	PHENYLKETONURIA								12q24.1
143100	HUNTINGTON CHOREA								4p16.3
314850	NEUROACANTHOCYTOSIS, MCLEOD TYPE								Xp21.2-p21.1
604569	CONTACTIN-ASSOCIATED PROTEIN 2								7q35-q36
300207	G PROTEIN-COUPLED RECEPTOR 50								Xq28
606941	DISRUPTED IN BIPOLAR DISORDER 1								11q23
126450	DOPAMINE RECEPTOR D2								11q23

Compounded GIM: Genes Common to Triplets of Related Disorders

The method can be compounded and used to find genes shared by three conditions and so on. These provide interesting and pivotal genes to focus on. Note that fewer genes are shared among progressively longer lists of disorders. List compression occurs as non-common genes are eliminated. This offsets combinatorial complexity and leaves us focusing on the genes of most concern first. We will boil down Tourette's related conditions to a few special genes listed below:

Genes Common To Triplets of Tourette's Related Disorders						
ANXIETY	BIPOLAR	OCD	116790	182138		113505
ANXIETY	BIPOLAR	LEARNING-DIS	116790			309550
ANXIETY	LEARNING-DIS	OCD	116790			
BIPOLAR	LEARNING-DIS	OCD	116790			
ANXIETY	BIPOLAR	DEPRESSION		182138		
ANXIETY	DEPRESSION	OCD		182138		
BIPOLAR	DEPRESSION	OCD		182138		
ADHD	BIPOLAR	TOURETTES			126452	191290
ADHD	ANXIETY	BIPOLAR			126452	
ADHD	ANXIETY	TOURETTES			126452	
ANXIETY	BIPOLAR	TOURETTES			126452	
ADHD	OCD	TICS				609678
ADHD	OCD	TOURETTES				609678
ADHD	TICS	TOURETTES				609678
OCD	TICS	TOURETTES				609678
ADHD	BIPOLAR	DEPRESSION				191290
ADHD	DEPRESSION	TOURETTES				191290
BIPOLAR	DEPRESSION	TOURETTES				191290
ADHD	LEARNING-DIS	TOURETTES		van-wdv 06		104210
Gene Key	Name		Locus			
116790	CATECHOL-O-METHYLTRANSFERASE		22q11.2			
182138	SEROTONIN TRANSPORTER		17q11.1-q12			
126452	DOPAMINE RECEPTOR D4		11p15.5			
609678	SLITRK1, INTEGRAL MEMBRANE PROTEIN		13q31			
191290	TYROSINE HYDROXYLASE		11p15.5			
113505	BRAIN-DERIVED NEUROTROPHIC FACTOR		11p13			
309550	FRAGILE X SYNDROME		Xq27.3			

The following table is the four way case. This small table is the most important one for now.

Genes Common To Quadruplets of Disorders				Name	Locus
ADHD	ANXIETY	BIPOLAR	TOURETTES	126452	DOPAMINE RECEPTOR D4 11p15.5
ADHD	BIPOLAR	DEPRESSION	TOURETTES	191290	TYROSINE HYDROXYLASE 11p15.5
ADHD	OCD	TICS	TOURETTES	609678	SLITRK1, INTEGRAL MEMBRANE PROTEIN 13q31
ANXIETY	BIPOLAR	DEPRESSION	OCD	182138	SEROTONIN TRANSPORTER 17q11.1-q12
ANXIETY	BIPOLAR	LEARNING-DIS	OCD	116790	CATECHOL-O-METHYLTRANSFERASE 22q11.2

A Brief Digression on Equivalence Relationships

Set theory¹² teaches three kinds of relationships these table entries can have, easily remembered using the mnemonic **SRT**. **S** stands for symmetric. A relationship is symmetric if:

$$(a \text{ isRelatedTo } b) \text{ implies } (b \text{ isRelatedTo } A)$$

All the relationships we are considering here are symmetric. If a gene is common to bipolar disorder and OCD, then the reverse is true; It is common to OCD and bipolar disorder. This distinction may sound ridiculous at first but it is mathematically necessary. The **R** stands for reflexive. A relationship is reflexive if it is true for itself. In the first table, there are 90 genes found for anxiety. The reflexive case is just the number of genes common to anxiety and anxiety. This number is a measure of the complexity of the disorder and research progress in the disorder. This number increases over time until all the genes relating to that disorder have been discovered and verified. All numbers along the principal diagonal of the table are reflexive entries. So we see that 90 genes have been discovered that relate to anxiety disorders.

The **T** in **SRT** stands for transitive. A relationship is transitive if:

$$(a \text{ isRelatedTo } b) \text{ AND } (b \text{ isRelatedTo } c) \text{ implies } (a \text{ isRelatedTo } c)$$

Incidentally, the relationship "equals" satisfies all three of SRT.

It is informative to consider everyday relationships which do, or do not possess these properties, but I digress. The intent of the graphic is this; Anxiety disorders share **nine** "genes" with bipolar disorders. That is, of the 90 genes associated with anxiety and the 63 genes associated with bipolar disorder, there are nine genes in common between the two conditions. As a matter of parity $(90 + 63) - 9 = 144$, genes are not shared, a measure of the relative complexity of the two illnesses and the effort that has been expended to address them. We would like to understand the neurochemical pathways that the genes share. Doing so gives us the value of understanding two disorders for the price of one. It allows us to prioritize our concerns.

A Strength of Genetic Intersection Method

A strength of GIM is the identification of hotspots shared among disorders. These constitute the "low-hanging fruit" for clinical diagnosis and treatment. These are the gene landmarks in neurochemical pathways. In the graphic above, those conditions that share genes are sorted and colored.

A Weakness of the GIM Method

In some pairs of disorders there appears to be no common genetic ground, no shared genes. That does not mean there aren't, it only means that we have not yet discovered them if they happen to exist. This is a weakness of the GIM method. Emphasizing this point: Zero common genes do not mean none exists, only that we haven't found them yet. Thus, GIM lives downstream of gene discovery. A potential source of false-positives occurs when a gene has a false-association history. In the worst case, one would include genes on a diagnostic gene chip that are not relevant and do not need to be monitored.

Explanation of GIM Method: Step One

In step one, the terms, and keywords most frequently associated with a set of related conditions in a complex disease are identified. There is a certain amount of lexical art when it comes to term selection. Some terms such as "depression" seem ambiguous. To avoid vague results terms should pertain specifically to the condition of interest. Sometimes casting a wider net with a more general term can be beneficial. In our case the search terms are:

Search Term	Comments
"Anxiety Disorder"	
"Bipolar Disorder"	
"Tourette's"	Unique term, "syndrome" not necessary
"OCD"	Abbreviation is unique to the obsessive-compulsive disorder
"ADHD"	Abbreviation is unique to Attention Deficit Hyperactivity Disorder
"Depression"	Chosen over "Depressive Disorder" this term is vague and wide, even so, it returns only 14 matches
"Tics"	Unique term, further exposition is unnecessary
"Phobias"	Used instead of more specific "Phobic Disorders"
"Learning Disorders"	
"Language Disorders"	
"Sleep Disorders"	

Step Two: Execute the Search Using the Search Terms in OMIM

Entering a search terms in OMIM¹³ produces a list of candidate "genes". (For the naïve method, any OMIM entry is considered a "gene"). Take the term "ADHD" for example. OMIM returns 20 entries, each with three fields, the OMIM entry number, the name of the entry, and gene locus as follows:

```
#143465 ATTENTION DEFICIT-HYPERACTIVITY DISORDER; ADHD Gene map locus 17p11, 16p13, 6q12, 5p13, 5p15.3, 4p16.1-p15.3
*126452 DOPAMINE RECEPTOR D4; DRD4 Gene map locus 11p15.5
*126455 SOLUTE CARRIER FAMILY 6 (NEUROTRANSMITTER TRANSPORTER, DOPAMINE), MEMBER 3; SLC6A3 Gene map locus 5p15.3
*104210 ALPHA-2A-ADRENERGIC RECEPTOR; ADRA2A Gene map locus 10q24-q26
#192430 VELOCARDIOFACIAL SYNDROME Gene map locus 22q11.2
+126453 DOPAMINE RECEPTOR D5; DRD5 DYSTONIA, PRIMARY CERVICAL, INCLUDED Gene map locus 4p16.1-p15.3
#137580 GILLES DE LA TOURETTE SYNDROME; GTS CHRONIC MOTOR TICS, INCLUDED Gene map locus 11q23
#600202 DYSLEXIA, SUSCEPTIBILITY TO, 2; DYX2 Gene map locus 6p22.2
+305400 FACIOGENITAL DYSPLASIA FGD1 GENE, INCLUDED; FGD1, INCLUDED Gene map locus Xp11.21
*609678 SLIT- AND NTRK-LIKE FAMILY, MEMBER 1; SLITRK1
%608904 ATTENTION DEFICIT-HYPERACTIVITY DISORDER, SUSCEPTIBILITY TO, 2 Gene map locus 17p11
%608903 ATTENTION DEFICIT-HYPERACTIVITY DISORDER, SUSCEPTIBILITY TO, 1 Gene map locus 16p13
*608396 SOLUTE CARRIER FAMILY 9 (SODIUM/HYDROGEN EXCHANGER), ISOFORM A9; SLC9A9 Gene map locus 3q21
%608906 ATTENTION DEFICIT-HYPERACTIVITY DISORDER, SUSCEPTIBILITY TO, 4 Gene map locus 5p13
%608905 ATTENTION DEFICIT-HYPERACTIVITY DISORDER, SUSCEPTIBILITY TO, 3 Gene map locus 6q12
%608631 ASPERGER SYNDROME, SUSCEPTIBILITY TO, 2 Gene map locus 17p13
#605899 GLYCINE ENCEPHALOPATHY; GCE HYPERGLYCINEMIA, TRANSIENT NEONATAL, INCLUDED; TNH, INCLUDED Gene map locus 16q24, 9p22, 3p21.2-p21.1
+162200 NEUROFIBROMATOSIS, TYPE I; NF1 NEUROFIBROMIN, INCLUDED Gene map locus 2p22-p21, 17q11.2
%102300 RESTLESS LEGS SYNDROME 1 Gene map locus 12q12-q21
*191290 TYROSINE HYDROXYLASE; TH Gene map locus 11p15.5
```

The summary result is as follows:

Search Term	Number of Associated "Genes"
"Anxiety Disorder"	90
"Bipolar Disorder"	63
"Learning Disorders"	29
"Tourette's"	20
"ADHD"	20
"Depressive Disorder"	14
"OCD"	12
"Tics"	10
"Speech or Language Disorders"	10
"Phobic Disorders"	8
"Sleep Disorders"	5
TOTAL	281

Step Three: Processing the Gene List For Each Search Term

As a positive control, the genes of each search term are compared with itself. If the software is functioning properly this trivially produces the original gene count for the search term. This is mathematically similar to dotting the gene vector returned by a search term with itself. It just gives the length, that is, the number of "genes" for a given disorder. In what follows the term "vector" and "list" will be used interchangeably.

To accomplish this means performing some necessary but mundane text processing. We must construct the result vector from the raw result by stripping everything except the gene ID field. The symbols in front of each gene ID have specific meanings given in the following table. OMIM returns a single character prefix that indicates the nature of the entry. This is expanded into a slightly more mnemonic two-character prefix. The first character pertains to phenotype, the expressed trait, the second character of the prefix pertains to genotype, or possible trait. P in upper case indicates phenotype is known and described. G in upper case indicates genotype locus is known. Lower case p indicates phenotype is described with unknown locus.

Prefix	Recoded Prefix	Meaning
*608903	_608903	a gene of known sequence
#608903	p_608903	a descriptive entry, usually of a phenotype, and does not represent a unique locus.
+608903	PG608903	a gene of known sequence and a phenotype.
%608903	P_608903	a confirmed mendelian phenotype or phenotypic locus for which the underlying molecular basis is not known.
608903	_608903	a phenotype for which the mendelian basis, although suspected, has not been clearly established or that the separateness of this phenotype from that in another entry is unclear.
^608903	_608903	entry no longer exists because it was removed from the database or moved to another entry as indicated.

Only entries of the first four prefix types occurred in this example. So for ADHD we have the following result transformed according to the table above:

ADHD.txt OMIM Returns	ADHD.tmp Recoded Version
+126453	PG126453
+162200	PG162200
+305400	PG305400
%102300	P_102300
%608631	P_608631
%608903	P_608903
%608904	P_608904
%608905	P_608905
%608906	P_608906
*104210	_G104210
*126452	_G126452
*126455	_G126455
*191290	_G191290
*608396	_G608396
*609678	_G609678
#137580	p_137580
#143465	p_143465
#192430	p_192430
#600202	p_600202
#605899	p_605899

The recoded result is more readable, and eliminates problems with symbols that have special meaning to the text processing software. Each list returned from a term search is stored in a file. The file suffix is initially .txt to indicate a text file. The file suffix is changed in the process to prevent stepping on the original data. This allows supervision of the process until it is verified.

Step Four: Exhaustive Search For Common Terms

In this step, every gene list is compared with every other gene list for each of the 11 search terms that produced them. This requires $11^4=14,641$ file comparisons for the four-way case, but null files are not compared, reducing the search cost significantly. The results are formatted in tabular form. The table is sorted by the number of common genes and colored to highlight conditions which share genes. The intensity of the color indicates the degree of overlap between the conditions. The main diagonal of the original matrix is a useful navigational landmark. The dot product principle applies to higher dimensional matrices because it remains a one-dimensional vector. This is critical for the higher order cases.

Genes Shared in Tourette Related Disorders

	Anxiety Disorders	Bipolar	Tourette	OCD	ADHD	Depressive Disorders	Tics	Phobic Disorders	Learning Disorders	Language Disorders	Sleep Disorders
Anxiety	90	9	5	4	1	3	3	6	3	1	0
Bipolar	9	63	3	4	3	8	1	0	4	1	0
Tourette											

Step Five: Overlapping Genes are Inspected For Significance

A count indicating the number of common genes common is produced by the software, as well as a list of the actual genes. This partitions the results into two groups; Entries whose locus and phenotype are well characterized and those which are not. Take for example the genes shared by anxiety and bipolar disorder in the small table below. The last entry will be discarded as no gene locus is associated with it.

Genes Shared By Anxiety and Bipolar Disorders
PG116790
PG122560
PG138040
PG309550
_G113505
_G126452
_G182138
_G607478
p_608516

Here we can use the prefix codes. We are most interested in those conditions for which we can establish a gene chip normal baseline and mutated cases. This will be true for entries prefixed with PG, meaning phenotype and genotype are well characterized, and for entries prefixed with _G, meaning some gene association with this disorder exists in the literature. Entries prefixed with P_ and p_ are discarded, since no gene specific information is available yet for these disorders. The remaining gene ID's are retained.

OMIM ID	Name	Condition	Gene locus
+116790	CATECHOL-O-METHYLTRANSFERASE	LOW, IN RED CELL	22q11.2
+122560	CORTICOTROPIN-RELEASING HORMONE	DEFICIENCY	8q13
+138040	GLUCOCORTICOID RECEPTOR	DEFICIENCY	8q13
+309550	FRAGILE SITE MENTAL RETARDATION 1 GENE	FRAGILE X SYNDROME	Xq27.3
*113505	BRAIN-DERIVED NEUROTROPHIC FACTOR		11p13
*126452	DOPAMINE RECEPTOR D4		11p15.5
*182138	SEROTONIN TRANSPORTER		17q11.1-q12
*607478	TRYPTOPHAN HYDROXYLASE 2		12q21.1

Author's Note: It was with excitement that the names of genes shared by anxiety and bipolar disorder were retrieved. One often reaches for but rarely experiences that sense of progress. It is enriching to see precursors, degradation enzymes and NT transport systems implicated as well as long-known NT. The similarity between hormones and neurotransmitters is striking. Long distance and short-range cell-to-cell signaling share mechanism, extending the boundary of the "nervous system". Hormones do their work binding to free cytosolic receptors which dimerize, enter the nucleus, and bind to DNA and upregulate transcription. Neurotransmitters bind to transmembrane receptors facilitating long and short-term state changes in cell state. In some sense, the synapse itself looks cytosolic. The difference being that hormones are ligands for transcription factors while neurotransmitters enable ion cascades. Hormones can trigger growth cascades which also have ionic involvement, for example the role of calcium ion concentration in nuclear envelope breakdown.

The take home lesson is that we must understand all players in the pathways these genes share. Compensation of the entire pathway will reduce disorders of the mind. Single factor solutions have problems due to the long delay between dose changes and assessment.¹⁴

We must pull these sweater threads to look for other gene loci with similar roles. This provides handles in the control of neurochemical equilibrium.

Two Clinical Approaches

Two clinical approaches are suggested:

Procedure 1: Macro to Micro - Clinically Driven Approach

In this approach, we move from clinical diagnosis to gene identification to gene tailored treatment. This is especially useful when a constellation of symptoms is present, each of symptoms can be used to indicate a screen to reveal which genes are involved. When specific functional genes are identified, therapies that are targeted to the specific gene products (or lack of!) can be performed.

The patient is screened with one or more gene chips to determine if the candidate genes are in fact mutated. This confirms and verifies the clinical assessment and diagnosis, providing a biochemical "second opinion" and confirming or refuting diagnostic theories being applied to a specific patient. A powerful advantage is that this works whether the patient is being seen for the first time, or has been in long-term treatment.

Procedure 2: Micro to Macro - Gene Driven Approach:

When patients are referred into a justice or psychiatric system without a solid history, they can be screened and eventually treated for a broad range conditions by looking directly at the results of gene chip assays. This enables understanding of the forces affecting them and faster routing to effective treatment. Many of the privacy and sampling issues relating to DNA have already been defined, so this is becoming more practical.

This kind of custom tailored genetic profile will also allow for ideal patient follow up to occur and life-damaging or threatening events to be predicted, assessed, and avoided.

Once genetic confirmation is achieved, it is not necessary to retest frequently. Annual or bi-annual retesting should be done, since genetic methods are evolving rapidly and this process is becoming more comprehensive. The entire diagnostic procedure should be repeated as improvements in the gene databases occur. At this writing, six months is probably adequate for running the deductive diagnostic software, unless new gene functions are discovered and announced. The range of conditions addressed can be broadened as well. One could allow clinicians to enter a variety of patient symptoms and come up with a "custom profile", for example asthma and anxiety.

Conclusions

In summary, it is possible to work in either direction; from clinical condition to phenotype to genotype and visa versa. The goal is personally tailored treatment. A bioinformatics method of finding genes common to related disorders has been described and tested. This method makes possible the definition of custom gene chips using for screening disorders of the mind in the eleven conditions reviewed above. A potential schematic is given below.

Gene Chip Schematic									
	BIPOLAR	ANXIETY	TOURETTES	ADHD	LEARNING-DIS	OCD	DEPRESSION	TICS	STRENGTH
126452	+	+	+	+					++++
191290	+		+	+			+		++++
609678			+	+		+		+	++++
182138	+	+				+	+		++++
116790	+	+			+	+			++++
309550	+	+			+				+++
113505	+	+				+			+++
104210			+	+	+				+++
126455			+	+					++
162200				+	+				++
126450		+	+						++
143100		+	+						++
138040	+	+							++
122560	+	+							++
607478	+	+							++
261600		+			+				++
606941	+		+						++
300207	+						+		++
314850	+							+	++
604569			+			+	van-wdv 106		++
POSSIBLE	12	11	9	6	5	5	3	2	

Gene Key	Name	Locus
126452	DOPAMINE RECEPTOR D4	11p15.5
191290	TYROSINE HYDROXYLASE	11p15.5
609678	SLITRK1, INTEGRAL MEMBRANE PROTEIN	13q31
182138	SEROTONIN TRANSPORTER	17q11.1-q12
116790	CATECHOL-O-METHYLTRANSFERASE	22q11.2
309550	FRAGILE X SYNDROME	Xq27.3
113505	BRAIN-DERIVED NEUROTROPHIC FACTOR	11p13
104210	ALPHA-2A-ADRENERGIC RECEPTOR	10q24-q26
126455	DOPAMINE TRANSPORTER	5p15.3
162200	NEUROFIBROMIN GENE	2p22-p21, 17q11.2
126450	DOPAMINE RECEPTOR D2	11q23
143100	HUNTINGTON CHOREA	4p16.3
138040	GLUCOCORTICOID RECEPTOR	5q31
122560	CORTICOTROPIN-RELEASING HORMONE	8q13
607478	TRYPTOPHAN HYDROXYLASE	12q21.1
261600	PHENYLKETONURIA	12q24.1
606941	DISRUPTED IN BIPOLAR DISORDER 1	11q23
300207	G PROTEIN-COUPLED RECEPTOR 50	Xq28
314850	NEUROACANTHOCYTOSIS, MCLEOD TYPE	Xp21.2-p21.1
604569	CONTACTIN-ASSOCIATED PROTEIN 2	7q35-q36

Acknowledgments

Support of this work by Lynn Mittelstaedt Warren, family and friends is gratefully acknowledged. The author would like to express deep appreciation regarding the continuing excellence of NCBI, OMIM, and PubMed data repositories. These are the great national resources of the genomic age. Thanks to Dr. Steven J. Fliesler of St. Louis University School of Medicine and Dr. Christopher Lamps of UAMS College of Medicine for their encouragement.

Abbreviations

OMIM: Online Mendelian Inheritance In Man
NCBI: National Center for Biotechnology Information
PubMed: National Library of Medicine Division
SSRI: Selective Serotonin Reuptake Inhibitor
NT: Neurotransmitter or Neurotransmitters

References

1. Lynch, Clifford: "Searching the Internet," *Scientific American*, March 1997
2. Service, Robert F.; Big Biology Is Bad Biology; *Science*: Vol. 291. no. 5507, p. 1182, 16 February 2001
3. Warren, L. Van: Knowledge Mapping the Corpus; www.wdv.com Feb 23, 2002
4. Cabeza, Roberto and Nyberg, Lars; Imaging Cognition II: An Empirical Review of 275 PET and fMRI Studies; *Journal of Cognitive Neuroscience*. Vol. 12:1-47, 2000
5. Perou, Charles M. et al; Molecular Portraits of Breast Cancer; *Nature* **406**, 747-752, (17 August 2000
6. Best, Ben; The Basal Ganglia, Chapter 2 – [Gross Neuroanatomy Web Page](#) 1990-2006
7. Osborn, Ian; Tormenting Thoughts and Secret Rituals – Dell, Random House ISBN 0-440-50847-9 (1998)
8. Kalat, James; [Biological Psychology](#) 8th edition– Wadsworth Thomson Publishing 2001
9. Lundbeck Institute; Brain Atlas - Neurotransmitters; www.brainexplorer.org (2005)
10. University of Utah; Genetic Science Learning Center; [Beyond the Reward Pathway](#) (2006)
11. Leckman, James F. and Cohen, Donald J.; [Tourette's Syndrome](#) – Tics, Obsessions, Compulsions; John Wiley & Sons (1999) ISBN – 0-471-16037-7
12. Wikipedia; [Equivalence Relation](#); Feb 14, 2006
13. OMIM; [Online Mendelian Inheritance in Man](#); National Center for Biotechnology Information (2006)
14. Mayberg, Helen S. et al; Regional Metabolic [Effects of Fluoxetine in Major Depression](#): Serial Changes and Relationship to Clinical Response; *Biological Psychiatry* (2000) Vol 48. p 831

last updated March 1, 2013