Journal of
Personalized
Medicine

MDPI

Reframing Psychiatry for Precision Medicine

Elizabeth B Torres [1,2,3]*

1   Rutgers University Department of Psychology; ebtorres@psychology.rutgers.edu
2   Rutgers University Center for Cognitive Science (RUCCS)
3   Rutgers University Computer Science, Center for Biomedicine Imaging and Modelling (CBIM)

*   Correspondence: ebtorres@psychology.rutgers.edu; Tel.: (011) +858-445-8909 (E.B.T)

Abstract: The art of observing and describing behaviors has driven diagnosis and informed basic science in Psychiatry. In recent times, studies of mental illness are focused on understanding the brain's neurobiology but there is a paucity of information on the potential contributions from peripheral activity to mental health. In Precision Medicine, this common practice leaves a gap between bodily behaviors and genomics that we here propose to address with a new layer of inquiry that includes genes' expression on tissues inclusive of brain, heart, muscle-skeletal and organs for vital bodily functions. We interrogate genes' expression on human tissue as a function of disease-associated genes. By removing genes linked to disease from the typical human set, and recomputing the genes' expressions on the tissues, we can compare the outcomes across mental illnesses, well-known neurological conditions, and non-neurological ones. We find that major neuropsychiatric conditions that are behaviorally defined today (e.g. Autism, Schizophrenia, Depression) through DSM-observation criteria, have strong convergence with well-known neurological ones (e.g. Ataxias, Parkinson), but less overlap with non-neurological ones. Surprisingly, tissues majorly involved in the central control, coordination, adaptation and learning of movements, emotion and memory are maximally affected in psychiatric diagnoses along with peripheral heart and muscle-skeletal tissues. Our results underscore the importance of considering both the brain-body connection and the contributions of the peripheral nervous systems to mental health.

Keywords: Autism; Schizophrenia; Mental Depression; Ataxia; Fragile X; Parkinson's disease; Mitochondria; Genes' expression; Tissues; neurological disorders; nervous systems disorders

1. Introduction

Modern medicine is at an inflexion point [1], whereby advances in computational methods, wearable sensing technology and open access to Big Data are reshaping the ways in which we inform basic science and rapidly translate our knowledge to actionable treatments. Psychiatry is one of those medical fields that is rapidly evolving, while adapting traditional models to help advance their main goal of helping patients improve quality of life. Along those lines, Computational Psychiatry [2] a nascent sub-field within Psychiatry, is merging methods from Computational Neuroscience with clinical approaches, through successful collaborations. These new developments are bound to open new frontiers in therapeutic treatments. And as part of a more general effort in the medical field, Precision Medicine (PM) [1] has emerged as a new platform to combine expertise from multiple layers of the knowledge network, to ultimately design personalized targeted treatments (**Figure 1A**). Integrating the personalized concept of PM with the new advances in Computational Psychiatry could give us a new way to approach mental illness and help patients cope with lifelong changing needs.

The task ahead is challenging because there is no proper roadmap to connect the layers of the knowledge network in PM and produce personalized diagnoses and measures of treatment outcomes that truly separate disease progression from treatment's effectiveness, according to age and

47 development. Part of the problem is that most brain science has focused on experimental assays and
48 methods that curtail natural movements. As such, our knowledge about the dynamics of natural
49 behaviors is very limited, particularly in reference to those aspects of behavior that remain hidden to
50 the naked eye of the clinician -trained to observe specific expected behavioral landmarks of a
51 Psychiatric disorder, conceived exclusively as a mental illness. In so doing, the clinically trained eye
52 may miss important information that is perhaps common across different disorders of the nervous
53 systems and rather relevant to help improve the patient's quality of life. For example, motor
54 coordination and volitional control are critical ingredients of autonomy in any natural behavior
55 underlying activities of daily living. Yet, these are not considered part of these diagnostics criteria for
56 mental illnesses such as Autism, Schizophrenia and Depression -as per the Diagnostics Statistical
57 Manual (DSM-5) [3] (and see Supplementary Material.)
58
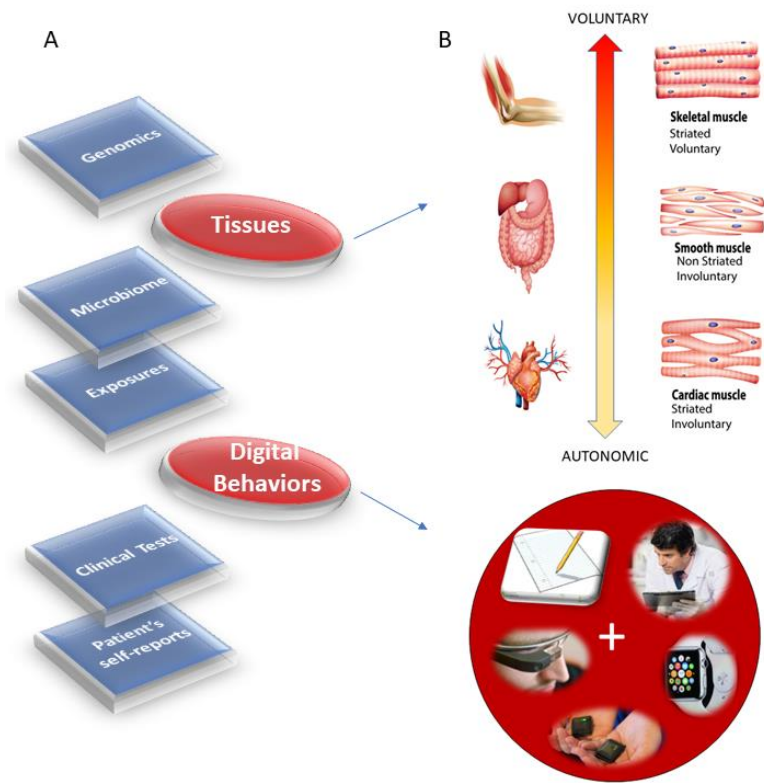
59
60 **Figure 1. Roadmap to implement the Precision Medicine Model for diagnoses and treatments of mental**
61 **illnesses**. (A) The PM's interconnected knowledge network can contribute information about the individual's
62 medical history, behaviors, environment, microbiome, and genetic makeup. Importantly, the new proposed
63 layer of *digitized behaviors* leveraging the wearable biosensors revolution can transform medicine by creating
64 truly personalized assessments. Additionally, the layer of behaviors can be connected to the nervous systems
65 functioning via fundamental levels of neuromotor control that span along a phylogenetically orderly taxonomy.
66 (B) This taxonomy is based on levels of maturation in autonomous neuromotor control, linked to three
67 fundamental muscle types: autonomic (by cardiac muscles); involuntary (by smooth muscles) and voluntary (by
68 skeletal muscles.) By linking the fundamental muscle types to the levels of control in the nervous systems, digital
69 behaviors can then be mapped to bodily autonomy, bodily autonomy mapped to muscle types and muscle types
70 to genes/proteins. Any measure of treatment's effectiveness for mental illnesses can then map back to
71 improvements in observable behaviors embedded in activities of daily social life.

72   Research on the underlying neurobiology of mental illnesses has revealed their associated
73 genetics and/or helped characterize patterns of brain activity in response to external stimuli (while
74 curtailing naturalistic bodily motions to avoid instrumentation artifacts in imaging data or in EEG,
75 MEG, etc.) This central approach to brain science has left us with a paucity of information about the
76 possible contributions to mental illness from the peripheral nervous systems, and from vital organs
77 important for autonomous living. The peripheral activity, however, continuously feeds back to the

78 brain via afferent (body-to-brain) channels and is in turn, dynamically updated through efferent
79 (brain-to-body) activity, self-generated by the system itself. This recursive loop, whereby re-entrant
80 information that is partly self-produced by the organism and partly influenced by external
81 environmental conditions, would provide important clues about truly evolving dynamics and
82 stochastic (variability) across all-natural behaviors. Approaching the problem through this lens could
83 bring a new quantifiable layer of granularity to basic research. This would include the design of age-
84 appropriate metrics reflecting the development of the organism as it ages and as it copes with a
85 disorder. The micro- and macro-motion data from the nervous systems biorhythms is the low
86 hanging fruit that we can easily attain by leveraging the wearable sensors revolution. And because
87 these quantifiable digitized activities and signals therein are partly self-generated, self-monitored
88 and self-corrected by and within the nervous systems, this quantitative approach has the potential to
89 switching us from a purely correlational science to a science that is based on causal relations between
90 nervous systems activities and external / contextual stimuli. In this sense, the new proposed approach
91 to mental illness is amenable to intervene and modify the system with well-informed, near-optimal
92 means capable of improving its performance.

93 Micro- and macro-motions that underly all aspects of human behavior depend on the intactness
94 of fundamental tissues, many of which have already been characterized in genomics. Here we
95 combine these underlying aspects of behavior with the current genomics knowledge, to inform
96 Psychiatry of possible ways to improve quantification of nervous systems activities. Ultimately, we
97 seek to build accommodations and support for the patient population, while reconceptualizing
98 mental illness as a physically quantifiable disorder of *the nervous systems*.

99 The nervous systems already offer a taxonomy of function and control that is phylogenetically
100 ordered and well organized along several axes. Some of these axes are accessible today with non-
101 invasive means and as such, we can obtain signals and build computational models to understand
102 mechanisms and translate them to actionable societal solutions. One possible orderly structure is
103 suggested in **Figure 1B**, where we propose to map levels of neuromotor control (voluntary,
104 involuntary, and autonomic) to fundamental types of muscles (skeletal, smooth, and cardiac) linked
105 to commonly sampled tissues in genomic datasets. Combining information about genes expression
106 on tissues that involve key components of the central nervous systems (the brain and the spinal cord),
107 key organs for vital bodily functions (including smooth muscle lining internal organs), muscle
108 skeletal tissues and nerves, and cardiac tissues (for autonomic heart functioning), we explore the
109 effects of removing disease-associated genes, on the overall remaining genome expression on these
110 tissues. As a first step in this exercise, we reasoned that the genes associated to a given disorder ought
111 to be important in the functioning of certain systems, which in turn depend on certain tissues. We
112 also reasoned that such stochastic variations and combinations could be measured relative to the
113 presence of all genes and to the absence of genes across neurological or non-neurological conditions.

114 *What is the tissue distribution of genes' expression in neuropsychiatric disorders like Autism,*
115 *Schizophrenia and Depression, in relation to well-characterized neurological conditions? Is there convergence*
116 *in the remaining genes' expression on the tissues, upon removal of the genes associated to that disease?*
117 *Furthermore, how would the genes' expression change across the tissues in non-neurological conditions like*
118 *various forms of cancer, immunodeficiencies, endocrine systems deficiencies and so forth? How would it change*
119 *in acquired disorders like PTSD, currently diagnosed through observation?*

120 Take Autism for example. Autism is an umbrella term for a very heterogeneous set of
121 neurodevelopmental disorders, but no gold standard criteria includes core neurological symptoms
122 that could help us create early accommodations and support for the nascent nervous systems of the
123 infant (during pre-cognitive stages of neurodevelopment.) The rule of thumb is to assume that the
124 child has odd, socially inappropriate behaviors and that they should be modified through operant
125 and cognitive conditioning techniques -often translated from lab animals to human babies, without
126 any type of collaboration with other fields studying infant development. Current methods of
127 diagnoses and treatments in autism are not based on normative neurodevelopmental data charts to
128 understand age-dependent departures from typical neurodevelopment. Without any systematic way
129 to build age-appropriate metrics, to capture highly non-linear, stochastic patterns and rates of change

130   in the (rather accelerated) infant neurodevelopment, entire generations of infants, children and
131   adolescents have been exposed to such means of behavioral treatments and no information can tie
132   these back to the underlying genomic pool of this population.
133   In Schizophrenia, delusions, avolition and catatonia are at the core of the disorder, but as in
134   Autism above, no criteria in the DSM highlights the profound somatic sensory-motor issues that have
135   been found in patients [4] -even without the use of psychotropic medications, known to alter motility.
136   Interestingly, historical accounts of Psychiatry (in pre-Freudian times) show the reliance on motor
137   aspects of the behaviors that defined several mental illness from a neurological perspective [5].
138   Depression is also currently treated purely as a mental illness, but it may be important to
139   understand potential contributions to various forms of depression, from the peripheral nervous
140   systems and from the body in general. Genetic information may give us a way to link tissues affected
141   in these neuropsychiatric conditions with those affected in neurological ones, for which treatments
142   and interventions of various forms may be effective. These may be in the form of drugs, or in the
143   form of physical, mindfulness and occupational therapies aimed at helping support the person's
144   bodily autonomy and overall increase the chances for independent living.
145   We here offer a new lens to help balance psychiatric with neurological criteria derived from
146   genomic information specific to each disorder. In a first (crude) step of many to come, we start by
147   comparing across well-known neuropsychiatric and neurological conditions, the results from
148   eliminating the genes associated with each disorder and quantifying the degree of convergence in the
149   maximally affected tissues, in relation to those resulting from eliminating the genes associated with
150   non-neurological ones. We focus our discussion on possible ways to continue this path of inquiry and
151   highlight current caveats for future improved iterations of the proposed methods.

2. Materials and Methods

153   We combine the data sets from genes associated with mental illnesses, with well-known
154   neurological disorders and with illnesses that are not directly associated with the nervous systems.
155   We also include genes associated with manifestations of acquired Post Traumatic Stress Syndrome
156   Disorders (PTSD). Among mental illnesses defined by the DSM-5, we include Autism, Schizophrenia
157   and Mental Depression of different types, (*e.g.* general, bipolar and unipolar).   Among neurological
158   conditions we include Ataxias (e.g. cerebellar, spino-cerebellar, progressive, and gait) and
159   Parkinson's disease. Among non-neurological disorders we include colon cancer, breast cancer,
160   diabetes, congenital heart disease, hematologic neoplasm, and various auto-immune ones (lupus
161   systemic erythematosus, psoriasis, and irritable bowel syndrome.)
162   We use as reference the genes, genes' expression, and tissues from the GTEx Portal human RNA-
163   Seq (Transcripts Per Million TPM)[1] specifically using the files denoted in the Appendix B. In Autism,
164   we use the genes scoring module of the Simons Foundation Autism Research Initiative (SFARI)
165   scored according to evidence from the literature. We also use Ataxia genes, the X-genes and the FX-
166   genes taken from various literature reviews [6, 7]. Furthermore, we use genes associated to
167   mitochondrial disorders [8] and genes identified in Parkinson's disease, taken from [9-14]. Besides

---

[1] Taken from the site: "Transcripts Per Million (TPM) is a normalization method for RNA-seq, should be read as *for every 1,000,000 RNA molecules in the RNA-seq sample, x came from this gene/transcript*. For each transcript in the gene model, the number (raw count) of reads mapped is divided by the transcript's length, giving a normalized transcript-level expression. The distribution of ambiguous reads (between transcripts of the same gene, or between different genes) is handled by OmicSoft's RSEM implementation. The sum of ALL normalized transcript expression values is divided by 1,000,000, to create a scaling factor. Each transcript's normalized expression is divided by the scaling factor, which results in the TPM value. Gene-level TPM's are calculated by summing up the transcript-level TPM for each gene. In this scaling, the sum of all TPMs (transcript-level or gene-level) should always equal 1,000,000. For cells that have approximately the same number of transcripts-per-cell, the TPM expression values can be compared between these cells to estimate relative abundance. For a given sample, TPM values will linearly scale with FPKM values for genes or transcripts, but FPKM will not add up to 1,000,000, so TPM can also be thought as FPKM, scaled to sum to 1,000,000."

168   the Autism SFARI genes and the genes reported in literature reviews, we take the genes associated
169   to Autism, Schizophrenia and Depression reported in https://www.disgenet.org/home/ along with
170   other genes from the above mentioned non-neurological disorders. The latter will inform us of
171   fundamental differences in genes' expression between these diseases and those which affect the
172   neuromotor control and basic functioning, as mediated by interactions between the brain and the
173   peripheral nervous systems (including the autonomic nervous system.)

174   The SFARI Autism categories that we used were those reported as of 03-04-2020. Quoting from
175   their site:

- CATEGORY 1
  Genes in this category are all found on the SPARK gene list. Each of these genes has been clearly implicated in ASD—typically by the presence of at least three de novo likely-gene-disrupting mutations being reported in the literature—and such mutations identified in the sequencing of the SPARK cohort are typically returned to the participants. Some of these genes meet the most rigorous threshold of genome-wide significance; all at least meet a threshold false discovery rate of < 0.1.
- CATEGORY 2
  Genes with two reported de novo likely-gene-disrupting mutations.
  A gene uniquely implicated by a genome-wide association study, either reaching genome-wide significance or, if not, consistently replicated and accompanied by evidence that the risk variant has a functional effect.
- CATEGORY 3
  Genes with a single reported de novo likely-gene-disrupting mutation.
  Evidence from a significant but unreplicated association study, or a series of rare inherited mutations for which there is not a rigorous statistical comparison with controls.
- SYNDROMIC
  The syndromic category includes mutations that are associated with a substantial degree of increased risk and consistently linked to additional characteristics not required for an ASD diagnosis. If there is independent evidence implicating a gene in idiopathic ASD, it will be listed as "#S" (e.g., 2S, 3S). If there is no such independent evidence, the gene will be listed simply as "S".

199   The GTEx dataset is as of 06-05-2017 v8 release. For every gene in the Autism, Ataxia, X, FX,
200   Mitochondrial diseases, Parkinson's disease, and the non-neurological diseases, we first confirmed
201   the presence of the gene in the GTEx data set and then incorporated it to the analyses.

202   The genes from the DisGeNet portal were found by interrogation of their data set under disease
203   type and saving the outcome to excel files containing all pertinent information. All sample files used
204   in our analyses are provided under Supplementary Materials.

Count Normalization

207   The GTEx matrix of RNA-seq genes along the rows (*56,146*) x the tissues (*54*) along the columns
208   was transposed (*54* x *56,246*), such that we expressed each tissue as a function of the gene expression
209   denoted by the count (TPM). Each individual count value was then normalized using equation (1)

$$NormalizedCount = \frac{count_i}{count_i + \dfrac{Avrg_{Global}Count}{Max_{Global}Count}} \tag{1}$$

211   Here $count_i$ is the count value of the *gene$_i$*, $Avrg_{Global}Count$ is the overall average of the matrix of
212   values taken along the columns and the rows. $Max_{Global}Count$ is the maximum count value, also taken
213   globally across the matrix values. Figure 2 shows the original count numbers (Figure 2A) and the
214   normalized version (coined Micro-Movement Spikes MMS) in Figure 2B. Figure 2C shows the MMS
215   derived from the fluctuations in counts normalized by equation (1) while Figure 2D shows the
216   histograms of the peaks (marked in red dots) for different tissues and genes scored by SFARI.

217

218
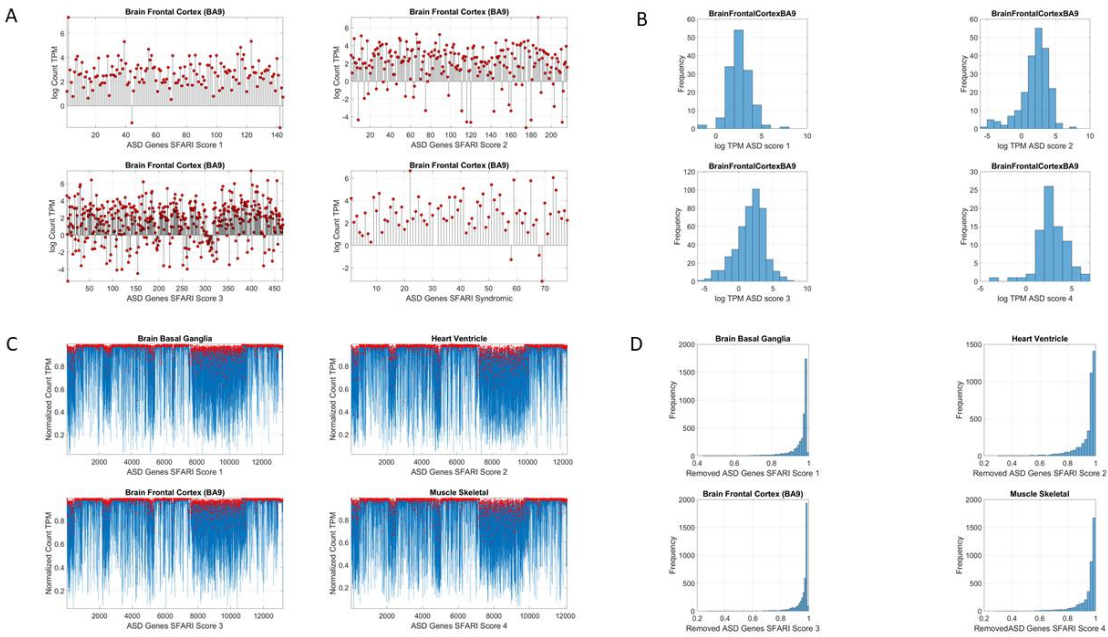
219 **Figure 2. Analytical methods**. (A) Sample raw data consisting of log count (TPM) for different scored genes
220 expressed in the brain frontal cortex (Brodmann Area 9.) (B) Histograms of the log Count TPM for each case in
221 (A). (C) Upon removal of the SFARI Autism genes, micro-fluctuations' spikes in the normalized count, with
222 deviations taken relative to empirically estimated mean, global averaged count, and global maximal count in
223 Equation (1). (D) Histograms of the normalized micro-fluctuations' spikes.

224 *Genes Removal*

225 For each of the disorders of interest in Tables 2 (Mental Illness and Neurological) and Table 3 (Non-
226 Neurological), we remove from the human GTEx data the set the genes associated with each
227 condition, as reported in the various databases (SFARI, DisGeNet https://www.disgenet.org/home/
228 and the literature meta reviews.) We then treat the resulting count series as a random process. We
229 use the exponential distribution to characterize it and to assess the differential expression across the
230 tissues relative to non-removal in the original human genome.
231
232 The question that we ask is, given the known neurological phenotypes, is there convergence
233 between the most affected tissues upon associated genes' removal and the changes in the tissues that
234 will be obtained by the removal of genes associated with mental illnesses? Furthermore, is there
235 convergence with the outcome from removing the genes associated with non-neurological illnesses?
236 Tables 2-3 show the number of genes removed in each respective case. It also shows the source of the
237 reported genes associated with each condition/disease.
238
239 *Stochastic Analyses*

240 Since the count values for each tissue can be conceived as a random series of numbers, we use
241 maximum likelihood estimation (MLE) to model the numbers representing the counts, as generated
242 by the exponential distribution using equation (2)

243 $$y = \lambda e^{-\lambda x} \qquad (2)$$

244 Here x represents the normalized count value (as per Equation (1)) and y the value from the
245 exponential distribution. We seek the value of λ to model this counting random process representing
246 the genes' expression in the tissue. To that end, we estimate the likelihood $L(\lambda \mid x_1, x_2, ..., x_n)$ where
247 the series of counts x_i, with i ranging from 1 to n represent the normalized counts (according to
248 Equation 1) across all genes for one tissue. Appendix B shows the steps to find λ. And this is

249 computed for each of the 54 tissues. We then rank the departure of λ (resulting from genes removal)
250 from the λ obtained for the full human genome (see below). This is explained in Figure 3.
251
252   **Table 2.** Genes distributions used in the removal process and literature sources.

Neurological	Number of Genes	Source
Parkinson	17	Lit Review [9, 11-14]
Ataxia Autosomal Recessive	70	Lit Review [6, 7]
Ataxia Autosomal Dominant	46	Lit Review [6, 7]
X-Chromosome	6	Lit Review [6, 7]
Fragile X (SFARI scores)	73 [1] (17,11,30,15)	SFARI Genes Module
Fragile X Syndrome	194	DisGeNet
FXTAS	62	DisGeNet
Ataxia	813	DisGeNet
Cerebellar Ataxia	421	DisGeNet
Gait Ataxia	159	DisGeNet
Progressive Cerebellar Ataxia	134	DisGeNet
Spinocerebellar Ataxia	145	DisGeNet
Neuropsychiatric (DSM)	**Number of Genes**	**Source**
Autism (SFARI scores)	906 [1] (144,216,468,78)	SFARI Genes Module
Autistic Disorder	1,043	DisGeNet
Schizophrenia	2,697	Lit Review [15-20] and DisGeNet
Mental Depression	1,468	DisGeNet
Depression Unipolar	641	DisGeNet
Depression Bipolar	116	DisGeNet

253   [1] SFARI scores for Autism and FX genes were also used to assess each scored module separately. FXTAS stands
254   for Fragile X Tremor Ataxia Syndrome

255   ***Stochastic Analyses – Visualization of change relative to the normative data of the full human***
256   ***genome***

257   Using MLE we also obtain for each of the 54 tissues, the frequency histograms of the normalized
258   counts across all genes and fit the continuous Gamma family of probability distributions with shape
259   (*a*) and scale (*b*) values, to obtain the Gamma moments and plot them on a parameter space. We do
260   this to visualize the spread of the tissues and their shift upon genes removal. To that end, we plot the
261   mean, the variance, and the skewness across the x-, y- and z-axis respectively. We plot the size of the
262   marker representing the tissue proportional to the kurtosis value, and we color the marker based on
263   the change relative to the original genome count (*i.e.* containing all the genes, without removal.)
264   To measure the stochastic shift between the tissues from the full genome and those upon
265   removal of the genes identified with each known neurological condition, we take the absolute
266   difference between the MLE λ for the full GTex genome and that for the genome upon removal of the
267   genes associated with each condition, disorder or disease, as shown in Figure 3.
268   We median rank the $\Delta\lambda$ for each tissue, sorting $\Delta\lambda$ in ascending order across the 54 tissues. Then
269   we create 4 median ranked blocks and plot the maximally affected block of tissues (Figure 3A).
270   The highest ranked group is then compared across all conditions, mental illness *vs.* those
271   neurologically defined *vs.* those non-neurological ones. We annotate the neurological functions that
272   such tissues are known to maximally disrupt. And we ask if there is convergence between the tissue
273   outcome in mental illnesses, upon removing the associated genes from the human genome-tissue
274   model, and the outcome upon removing those genes tied to the other known disorders of the nervous
275   systems. We repeat this interrogation process using the genes associated with disorders in the
276   DisGeNet portal, including Autism, Schizophrenia, Depression, the neurological in Table 2, and the
277   non-neurological ones in Table 3.
278   To assess tissue outcome upon removal of genes associated to various non-neurological diseases,
279   we follow these procedures and compare these to the above results. These diseases include colon

280   cancer, breast cancer, psoriasis, diabetes, congenital heart disease, hematologic neoplast and systemic
281   lupus. Table 3 describes the number of genes associated with each of these diseases and the sources.
282   We also examine other mental illnesses described by the DSM. These include schizophrenia,
283   depression, unipolar depression, and bipolar depression. We ask if there are tissues that overlap with
284   those affected in the neurological disorders. Lastly, we examine mitochondria related disorders and
285   PTSD using these methods. We reasoned that these may be disorders that potentially have affected
286   tissues across a broader range of functions, including those from the brain and other bodily organs.
287
288   **Table 3.** Genes associated with non-neurological diseases.

| Non-Neurological	Number of Genes	Source
Colon Cancer	3,669	DisGeNet
Diabetes Mellitus (noninsulin dependent)	3,134	DisGeNet
Oestrogen Receptor Positive Breast Cancer	510	DisGeNet
Congenital Heart Disease	267	DisGeNet
Hematologic Neoplasm	827	DisGeNet
Systemic Lupus Erythematosus	1,883	DisGeNet
Psoriasis	1,308	DisGeNet
Irritable Bowel Syndrome	1,483	DisGeNet
Mixed	**Number of Genes**	**Source**
Mitochondria	41	Lit Review [8]
Mitochondrial Myopathies	121	DisGeNet
Mitochondrial Diseases	284	DisGeNet
MELAS Syndrome	81	DisGeNet
PTSD	418	DisGeNet

289   MELAS (a form of dementia) stands for Mitochondrial Encephalopathy, Lactic Acidosis, and Stroke-like episodes.

290
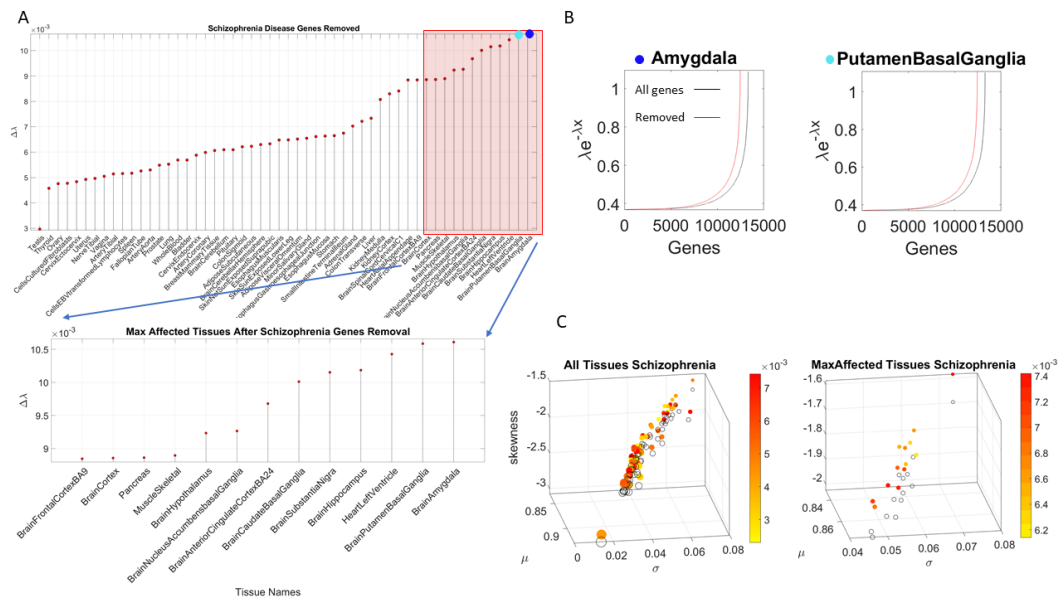
291   **Figure 3.** Sample metrics used for the stochastic analyses of a data sample (using 2,697schizophrenia associated
292   genes reported in DisGenNet portal and in the literature.) (A) Effect of removing the schizophrenia genes from
293   the GTex human genome set expressed across 54 tissues. Tissues are sorted in ascending order, by the absolute
294   difference $\Delta\lambda$ between genes' expression on the 54 tissues before and after removal. Red square highlights the
295   top 13 median ranked tissues shown in the panel below and dark and light blue circles mark the top two
296   tissues affected (the brain amygdala and brain putamen in the basal ganglia. (B) The exponential distribution
297   curve is fit to the sorted normalized count representing the gene's expression in TMP on the top median-
298   ranked affected tissues (as in Figure 2C, taking the peaks highlighted in red and fitting the exponential
299   distribution to the frequency histogram, as in Figure 2D) before (black line) and after (red line) the removal of
300   the genes associated with the disease. The absolute value difference between the curves is the $\Delta\lambda$ used to rank

301 the tissues by the effect size. (C) The fitting of the Gamma distribution yields the shape and scale parameters
302 used to compute the Gamma moments. The axes represent the mean, the variance, and the skewness of the
303 distribution of the normalized values and the color map represents the Earth's Movers Distance values
304 measuring the difference between the resulting exponential frequency histograms in (B). The size of the circle
305 is proportional to the kurtosis and the color-filled circles represent the tissue (54 in the left panel) with the
306 original genes' expression from GTex (our reference template) *vs.* the open circles representing the stochastic
307 shift, i.e. upon the removal of the genes associated with the disease in DisGenNet. Right panel contains the top
308 ranked tissues (13) according to the median values of the Δλ.

309
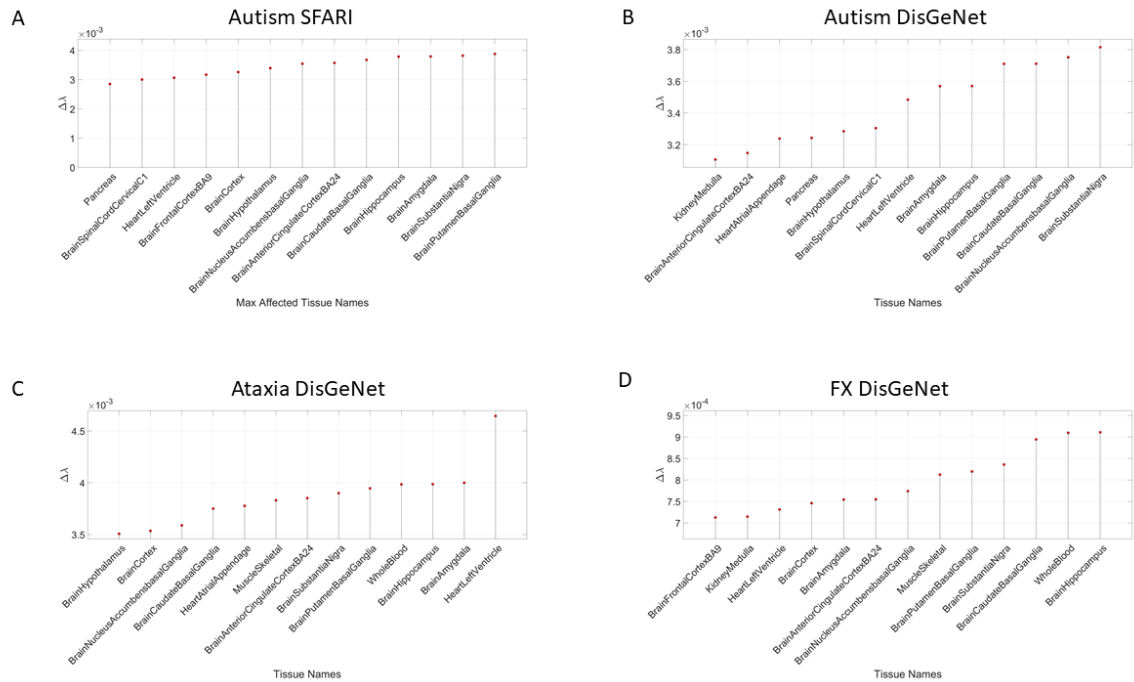
310 **Figure 4.** Convergence between Autism and known Neurological disorders shown by comparison of maximally
311 affected tissues after removing genes associated with the disorder in Autism (from SFARI), Autistic Disorders
312 (from the DisGeNet portal) *vs.* Ataxia and Fragile X (See Tables 4-5)

313 **3. Results**

314 **Autism, Ataxia and FX have convergence in maximally affected tissues by the removal of**
315 **associated genes**

316 　　The maximally affected tissues upon genes removal, according to the genes' stochastic
317 expression (count in Transcripts Per Million, TPM) are depicted in Table 4. This tissues' gene
318 expression was modelled by the exponential distribution $y = \lambda e^{-\lambda x}$, with x as the genes' combination
319 expressed in the tissues, and λ as the exponential rate parameter. The Δλ between the neurotypical
320 template case from the GTex portal (containing all genes) and the modeled disorder case (upon
321 removal of the SFARI genes) provides a sense for the departure from the normative case. This
322 difference, taken for the removal of the SFARI genes, is depicted in Figure 4A, with samples of
323 maximally affected tissues in Figure 4B that are known to be critical for motor control, regulation,
324 adaptation/learning, and coordination.
325 　　The Δλ-median ranking quantified the difference between neurotypical tissues' gene expression
326 *vs.* tissues' gene expression upon removal of the genes corresponding to the disorders in question,
327 with 4 groups ordered by the size of Δλ. This λ-quantity was first obtained relative to the neurotypical
328 population tissues, *i.e.* including all the counts (gene expression) from all genes, *i.e.* to model an
329 exponential process. The computation of λ using maximum likelihood estimation (MLE) is explained
330 in the Appendix A. It does not assume any order of the counts, but rather seeks to identify the

331 resulting λ for each tissue, treating the genes' count (expression) as a random, memoryless stochastic
332 process. Typically, the exponential distribution is used to model times between events, but here we
333 used it to model the fluctuations in the values of the counts across the genes as they randomly
334 fluctuate their expression across each of the 54 tissues reported in the GTex portal. We note this to
335 underscore that the results spontaneously self-emerge from the random combination of the genes
336 involved (with and without removal), rather than from the clinical criteria used to denote the genes'
337 relevance to Autism, or the evidence from the literature used to determine their association. There is
338 in fact no scoring of such relevance for the genes associated to the other neurological disorders under
339 consideration (*e.g.* the Ataxias, Parkinson's, etc.) or for the Autistic disorders reported in the
340 DisGenNet portal. Using those genes instead of the SFARI reveals tissues in Table 5, where we report
341 the convergence.

342

SFARI Autism	Ataxias, X, FX, PD, Mitochondria	SFARI Autism without overlapping Ataxia, PD genes
Putamen Basal Ganglia	Hippocampus	Putamen Basal Ganglia
Substantia Nigra	Amygdala	Substantia Nigra
Amygdala	Substantia Nigra	Amygdala
Hippocampus	Heart Left Ventricle	Hippocampus
Caudate Basal Ganglia	Whole Blood	Caudate Basal Ganglia
Anterior Cingulate Cortex	Muscle Skeletal	Nucleus Accumbent Basal Ganglia
Nucleus Accumbent Basal Ganglia	Nucleus Accumbent Basal Ganglia	Anterior Cingulate Cortex
Hypothalamus	Putamen Basal Ganglia	Hypothalamus
Brain Cortex	Anterior Cingulate Cortex	Brain Cortex
Frontal Cortex	Brain Cortex	Frontal Cortex
Heart Left Ventricle	Hypothalamus	Spinal Cord
Spinal Cord	Heart Atrial Appendage	Heart Left Ventricle
Pancreas	Caudate Basal Ganglia	Kidney Cortex

343
344 **Table 4.** The 13 top median ranked tissues in descending order of Δλ value, the highest affected tissues upon
345 removal of the SFARI genes (906) linked to Autism from the human GTEx database in column 1. Ataxias, X-
346 Chromosome, Fragile X, Parkinson's disease and Mitochondrial disease extracted from the literature and
347 column 3 is the same as in column 1 while removing from the SFARI Autism set 14 genes that overlap with the
348 Ataxias and PD (see those genes listed in Supplementary Table 2). Tissues are grouped by CNS (brain and spinal
349 cord in blue); muscle skeletal (green), heart (pink) and peripheral organs (gray). SFARI Autism (11/13 (84.6%)
350 CNS, 1/13 (7.6%) heart and 1/13 (7.6%) peripheral organ); neurological disorders 10/13 (76.9%) CNS, 2/13 heart
351 (15.3%) and 1/13 (7.6%) muscle skeletal); SFARI Autism without the overlapping genes from the neurological
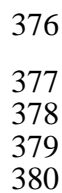352 disorders 11/13 (76.9%) CNS, 1/13 (7.6%) heart and 1/13 (7.6%) peripheral organs.

353

DisGeNet Autistic Disorder	DisGeNet Ataxia	FX DisGeNet
Substantia Nigra	Heart Left Ventricle	Hippocampus
Nucleus Accumbent Basal Ganglia	Amygdala	Whole Blood
Caudate Basal Ganglia	Hippocampus	Caudate Basal Ganglia
Putamen Basal Ganglia	Whole Blood	Substantia Nigra
Hippocampus	Putamen Basal Ganglia	Putamen Basal Ganglia
Amygdala	Substantia Nigra	Muscle Skeletal
Heart Left Ventricle	Anterior Cingulate Cortex	Nucleus Accumbent Basal Ganglia
Spinal Cord	Muscle Skeletal	Anterior Cingulate Cortex
Hypothalamus	Heart Atrial Appendage	Amygdala
Pancreas	Caudate Basal Ganglia	Brain Cortex
Heart Atrial Appendage	Nucleus Accumbent Basal Ganglia	Heart Left Ventricle
Anterior Cingulate Cortex	Brain Cortex	Kidney Medulla
Kidney Medulla	Hypothalamus	Frontal Cortex

354
355 **Table 5.** Highest affected tissues upon removal of the DisGeNet genes associated with Autistic Disorder (1,043)
356 from the human GTEx database (column 1); Ataxia (813 genes) in DisGeNet (column 2) and FX (194 genes) in

357   DisGeNet (column 3). Convergence between autism and neurological disorders is noted in the shaded tissues
358   color coded as in Table 4, based on CNS, Heart, and peripheral organs.

359   Removal of SFARI genes, ranked by change in gene expression, reveals brain tissues linked to
360   the CNS, in brain subcortical tissues linked to motor control (basal ganglia, striatum), memory
361   (hypocampus), emotions (amygdala) and regulation (hypothalamus); and the spinal cord. This is also
362   generally the case for the removal of the DisGeNet genes associated with Autistic Disorders and with
363   FX and Ataxia. Congruent with the outcome from the removal of the SFARI genes from the GTex
364   genome, the DisGeNet genes removal also affected the tissues associated CNS function. Important
365   tissues for systemic organs' functioning such as those containing smooth muscles, cardiac and
366   skeletal muscles in the taxonomy proposed in Figure 1B were also affected (commonly) across these
367   disorders. Figure 4 shows a summary of the results, visualizing the stochastic shifts. Tables 4-5 show
368   the tissues ranked in descending order and color-coded according to CNS, brain, and spinal cord
369   (blue); heart-related (pink), muscle skeletal (green) and peripheral vital organs (gray). Most tissues
370   in Autism and the neurological disorders are from the CNS, followed by the PNS related tissues in
371   the heart and muscle skeletal and with the vital organs towards the end of the $\Delta\lambda$-ranking.

372   Supplementary Figures 1-3 show these results separately for each neurological condition. We
373   note in Supplementary Material that removal of the SFARI Autism syndromic genes from the GTex
374   genome reveals maximal differences in tissues of organs with smooth and cardiac muscles, linked to
375   involuntary and autonomic function in the proposed taxonomy of Figure 1B.

376

Figure 5. Genes' expression on maximally affected tissues (colorbar coded in log TPM) upon removal of
overlapping genes between the SFARI Autism set and the Ataxias (dominant and recessive and X-Chromosome
sets) from the literature (A) and (B) from Parkinson's disease. Horizontal axis is the tissues' names and vertical
axis the genes' names.

381   Removal of overlapping SFARI genes and Neurological disorders also reveal brain tissues linked
382   to motor control, memory, emotions, and regulation. This is depicted in Figure 4. Given the
383   congruence between the tissues maximally affected by removing the SFARI Autism genes from the
384   GTEx database and those from the Neurological conditions, we next ascertain the extent to which
385   these genes overlap with those used from the Ataxias in the literature. To that end, we divide them
386   into the autosomal dominant, the autosomal recessive and the X-Chromosome genes. Figure 5A
387   shows the result of this interrogation broken down by the Ataxia's autosomal dominant, autosomal
388   recessive and X-Chromosome genes. The genes listed in this figure overlap with those from the SFARI
389   Autism genes. Table 4-Column 3 lists the tissues while Table 6 also has the scoring from the SFARI
390   Autism genes. The Supplementary Material Table 2 lists the phenotypic information of the disorders
391   associated with these genes, as described by the clinical literature. Figure 5C lists the PD gene that

392 overlaps with the SFARI Autism genes, also depicted on Table 6 along with the score ranking from
393 the SFARI portal.
394

SFARI Autism & Ataxia Dominant	SFARI Autism & Ataxia Recessive	SFARI Autism & (Early Onset) PD
CACNA1C (1)	**KCNJ10 (2)**	**RAB39B (3)**
SCN2A (1)	**WWOX (2)**	
ATP1A3 (3S)	**GRID2 (3)**	
CCDC88C (3)	**LAMA1 (3)**	
ITPR1 (3)	**PEX7 (3)**	
	SYNE1(3S)	
	CYP27A1 (4)	
	SNX14 (4)	

395
396 **Table 6.** Overlapping genes between the Ataxias (dominant and recessive genes) and the Parkinson's disease
397 with the genes from the SFARI portal. Scores in parenthesis refer to the scoring of the gene according to the
398 SFARI site (see Methods for explanation on each category). Syndromic is (4). Supplementary Figures 5-18
399 provide the GTEx violin plots of these genes' expression in the top-ranked tissues unveiled by our analyses in
400 Table 4. Supplementary Table 2 compiles the genes' additional information from various sources in the clinical
401 literature.
402
403   We note that removing this subset of 14 overlapping genes from the SFARI Autism set (Table 6),
404 does not change the primary result whereby the most affected tissues upon removal of the SFARI
405 Autism set from the GTEx dataset are those associated with sub-cortical brain structures critical for
406 motor control, adaptation/learning, regulation, coordination and autonomic function, as well as
407 memory and emotion. This is shown in Figure 5A-B and in the third column of Table 4. We also plot
408 in Figure 4D the top ranked tissues affected by the removal from the GTEx dataset of the FX genes
409 reported in the DisGeNet Portal. Supplementary Figures 19-22 further provide details on X-
410 Chromosome genes in Figure 5A implicated in Autism according to the SFARI genes portal.
411

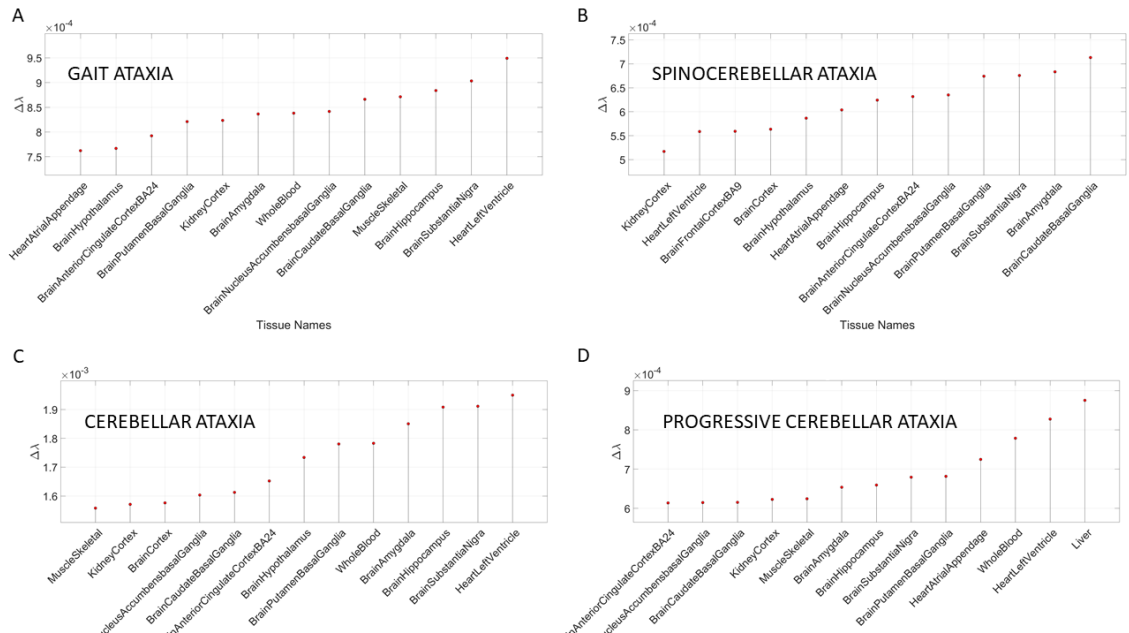| Gait Ataxia	Spinocerebellar Ataxia	Cerebellar Ataxia	Progressive-C Ataxia
Heart Left Ventricle	**Caudate Basal Ganglia**	**Heart Left Ventricle**	**Liver**
Substantia Nigra	**Brain Amygdala**	**Substantia Nigra**	**Heart Left Ventricle**
Hippocampus	**Substantia Nigra**	**Hippocampus**	**Whole Blood**
Muscle Skeletal	**Putamen Basal Ganglia**	**Amygdala**	**Heart Atrial Appendage**
Caudate Basal Ganglia	**N Accumbens BG**	**Whole Blood**	**Putamen Basal Ganglia**
N Accumbens BG	**Ant Cingulate Cortex**	**Putamen Basal Ganglia**	**Substantia Nigra**
Whole Blood	**Hippocampus**	**Hypothalamus**	**Hippocampus**
Brain Amygdala	**Heart Atrial Appendage**	**Ant Cingulate Cortex**	**Brain Amygdala**
Kidney Cortex	**Hypothalamus**	**Caudate Basal Ganglia**	**Muscle Skeletal**
Putamen Basal Ganglia	**Brain Cortex**	**N Accumbens BG**	**Kidney Cortex**
Ant Cingulate Cortex	**Frontal Cortex**	**Brain Cortex**	**Caudate Basal Ganglia**
Hypothalamus	**Heart Left Ventricle**	**Kidney Cortex**	**N Accumbens BG**
Heart Atrial Appendage	**Kidney Cortex**	**Muscle Skeletal**	**Ant Cingulate Cortex**

412
413 **Table 7.** Highest affected tissues upon removal of the DisGeNet genes associated with different types of Ataxias,
414 color-coded by CNS (brain, spinal cord), heart-related, muscle-skeletal, and peripheral vital organ for systemic
415 functioning. Predominance of CNS is evident, followed by heart-related and muscle skeletal and peripheral
416 organs.
417
418   The results that convergence in the ranked descending order of CNS (brain and spinal cord
419 tissues), followed by heart-related tissues, muscle skeletal tissue and lastly peripheral vital organs for
420 systemic functioning in SFARI-Autism and well-known neurological disorders from the literature is
421 also congruent with the results using the genes associated with these conditions in the DisGeNet
422 portal. There we interrogated Autistic Disorders, Ataxias and Fragile X, confirming the overlapping

423 in genes, their expression on the 54 tissues of the GTex database and the orderly levels of tissues
424 maximally affected by the removal of the associated genes. We grouped the tissues by CNS, heart,
425 muscle skeletal and peripheral vital organs to follow the proposed taxonomy of Figure 1B.

426   In the remaining of the paper, we consistently use this grouping to simplify visualization of the
427 tables and data presentation. Figure 6 shows the results for different types of the Ataxias and FX,
428 while Table 7 summarizes the top ranked affected tissues in different types of Ataxias, color-coded
429 according to the grouping approximating the taxonomy proposed in Figure 1B.

430

431 **Figure 6.** Different types of Ataxia showing the top $\Delta\lambda$-ranked tissues also shown in Table 7. (A) Gait Ataxia;
432 (B) Spinocerebellar Ataxia; (C) Cerebellar Ataxia; (D) Progressive Cerebellar Ataxia.

433 **Removal of Genes Associated with Schizophrenia and Multiple Forms of Mental Depression**
434 **Reveal Convergence with Neurological Disorders**

Schizophrenia	Depression	Unipolar Depression	Bipolar Depression
Amygdala	Heart Left Ventricle	Putamen Basal Ganglia	Hippocampus
Putamen Basal Ganglia	Heart Atrial Appendage	Caudate Basal Ganglia	Putamen Basal Ganglia
Heart Left Ventricle	Hippocampus	Hippocampus	Amygdala
Hippocampus	Muscle Skeletal	Substantia Nigra	Caudate Basal Ganglia
Substantia Nigra	Putamen Basal Ganglia	Kidney Cortex	Muscle Skeletal
Caudate Basal Ganglia	Caudate Basal Ganglia	Amygdala	Substantia Nigra
Ant Cingulate Cortex	Amygdala	Heart Left Ventricle	Spinal Cord
N Accumbens BG	Pancreas	Kidney Medulla	Esophagus Muscularis
Hypothalamus	Kidney Medulla	Brain Cortex	Adrenal Gland
Muscle Skeletal	Substantia Nigra	N Accumbens BG	Frontal Cortex
Pancreas	Kidney Cortex	Colon Transverse	Minor Salivary Gland
Brain Cortex	Ant Cingulate Cortex	Ant Cingulate Cortex	Colon Sigmoid
Frontal Cortex	N Accumbens BG	Heart Atrial Appendage	Ant Cingulate Cortex

435
436 **Table 8.** Highest affected tissues upon removal of the DisGeNet genes associated with Schizophrenia (2,697)
437 from the human GTEx database; Depression genes (1,468); Unipolar Depression genes (641) and Bipolar
438 Depression genes (116). Convergence between Schizophrenia and Depression is high, with maximally affected
439 CNS tissues (10/13), followed by heart-related and muscle skeletal. Unipolar and bipolar depression also show
440 systemic effect of vital peripheral organs (N stands for Nucleus, Ant for Anterior, and BG for Basal Ganglia.)

441

442   The removal from the normative GTex genome of the DisGeNet genes associated with mental
443 illnesses such as Schizophrenia, Depression, Bipolar Depression and Unipolar Depression resulted in

444 convergence of maximally affected tissues involved in the CNS, especially those brain regions
445 necessary for neuromotor control, memory, and emotion. This is depicted in Table 8 and Figure 7.
446 Several of these tissues were also found to be affected upon removal of the SFARI genes and the genes
447 associated in DisGeNet with Autistic Disorders. Furthermore, these are maximally affected tissues in
448 the well-known neurological conditions depicted in Figures 4-5, Tables 4-6.
449
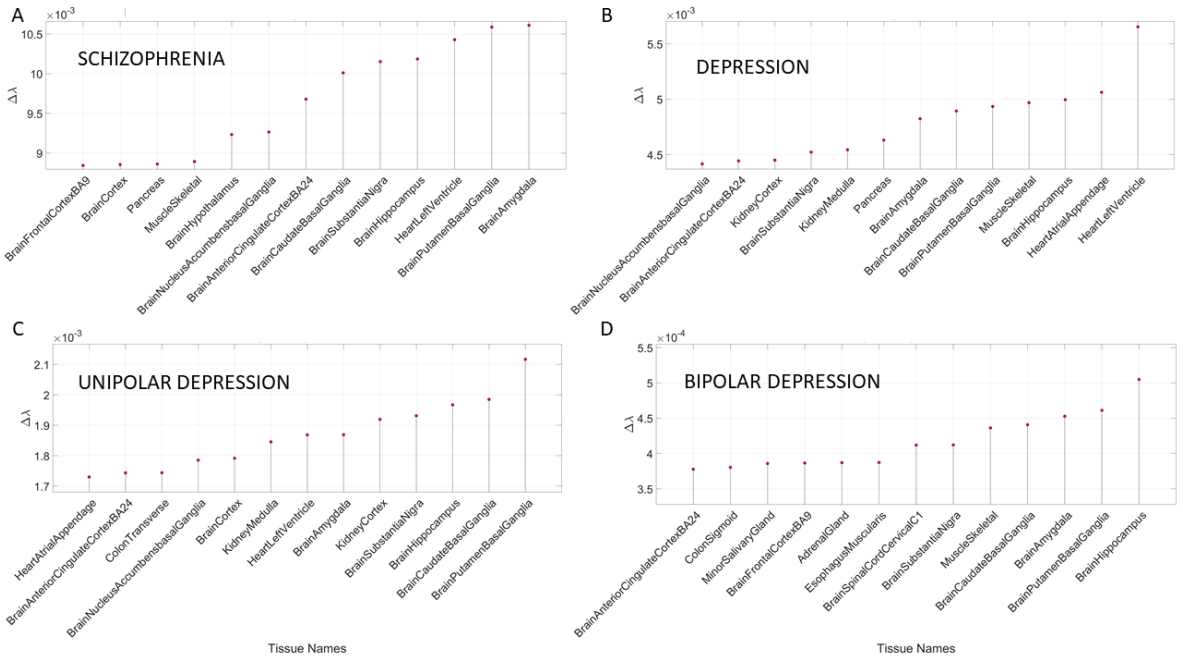
451 **Figure 7.** Maximally affected tissues in Schizophrenia and Depression, and in different types of Depression
452 (Unipolar and Bipolar) shown in Table 8.

Removal of Genes Associated with Non-Neurological Disorders Reveal Other non-CNS Tissues

454   In addition to the examination of mental illnesses and neurological disorders, we also
455 interrogated the GTex genome upon removal of genes associated with various non-neurological
456 disorders. These included various forms of cancers, inflammatory and auto-immune disorders and
457 other tissues related to the heart, the circulatory and the endocrine systems. Tables 9-10 summarize
458 the results of this interrogation and Figures 8-9 show the Δλ-ranking graphs.
459

| Colon Cancer	Irritable Bowel Syndrome	Congenital Heart Disease	Hematologic Neoplast
Pancreas	Heart Left Ventricle	Heart Left Ventricle	Heart Left Ventricle
Whole Blood	Pancreas	Heart Atrial Appendage	Whole Blood
Stomach	Kidney Cortex	Kidney Cortex	Liver
Adrenal Gland	Liver	Whole Blood	Heart Atrial Appendage
Colon Transverse	Muscle Skeletal	Kidney Medulla	Pancreas
Prostate	Heart Atrial Appendage	Substantia Nigra	Muscle Skeletal
Small Intestine T	Kidney Medulla	Hypothalamus	Kidney Cortex
Spleen	Whole Blood	Prostate	Hippocampus
Minor Salivary Gland	Adrenal Gland	Hippocampus	Kidney Medulla
Fallopian Tube	Adipose Viseral Oment	Muscle Skeletal	Esophagus Mucosa
EBT Lymphocytes	Esophagus Mucosa	Adipose Visceral O	EBT Lymphocytes
SkinNoSunExposed S	Colon Transverse	Minor Salivary Gland	Substantia Nigra
Liver	Stomach	Pancreas	Putamen Basal Ganglia

460

461 **Table 9.** Highest affected tissues upon removal of the DisGeNet genes associated with Colon Cancer (3,669) from
462 the human GTEx database; Irritable Bowel Syndrome genes (1,483); Congenital Heart Disease genes (267) and
463 Hematologic Neoplast genes (827). Color code as in previous Tables.

464

465
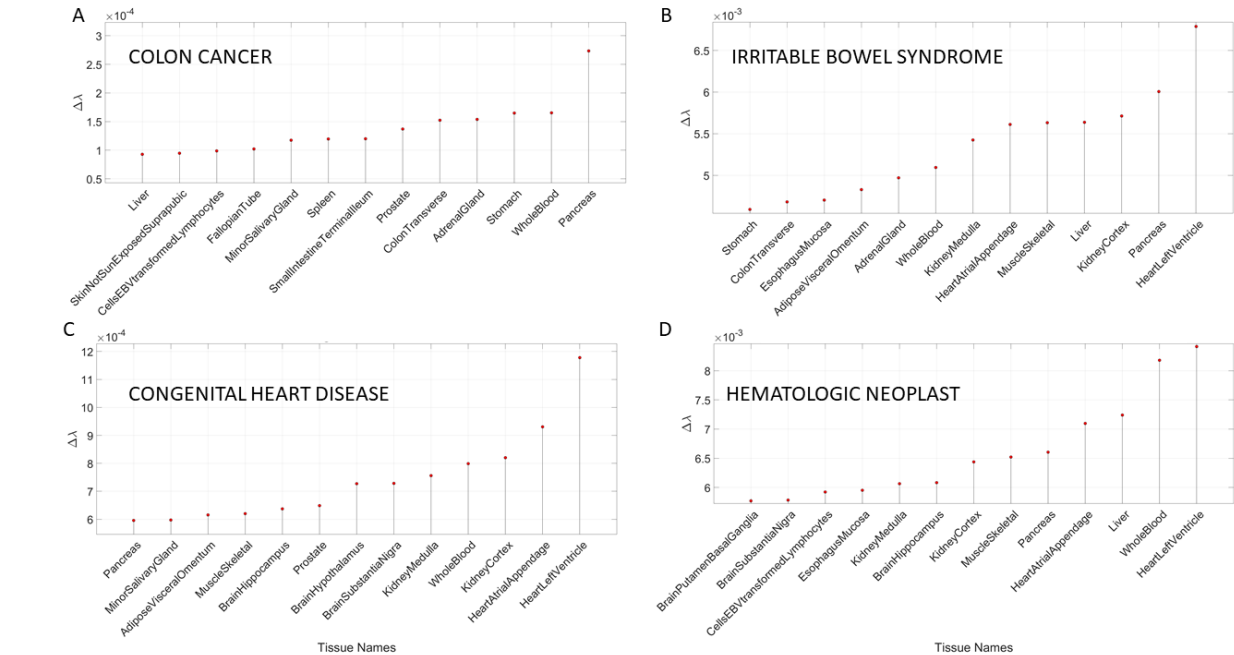
Figure 8. Maximally affected tissues in sample non-neurological diseases of Table 9 reveal primarily non-CNS tissues involving peripheral vital organs for systemic functioning, followed by heart-related and muscle skeletal tissues. As before, the interrogation of the GTex genome is based on the genes associated with diseases in the DisGeNet portal. (A) Colon cancer; (B) Irritable Bowel Syndrome; (C) Congenital heart disease and (D) Hematologic Neoplast. Color code as in previous Tables.

471

Systemic Lupus Erythematosus	Psoriasis	Breast Cancer	Diabetes
Heart Left Ventricle	Heart Left Ventricle	Pancreas	Liver
Whole Blood	Liver	Heart Left Ventricle	Pancreas
Liver	Whole Blood	Heart Atrial Appendage	Heart Atrial Appendage
Heart Atrial Appendage	Muscle Skeletal	Kidney Medulla	Heart Left Ventricle
Pancreas	Heart Atrial Appendage	Muscle Skeletal	Kidney Medulla
Muscle Skeletal	Pancreas	Whole Blood	Kidney Cortex
Kidney Cortex	Kidney Cortex	Ant Cingulate Cortex	Putamen Basal Ganglia
Hippocampus	Putamen Basal Ganglia	Esophagus Mucosa	Substantia Nigra
Kidney Medulla	Caudate Basal Ganglia	Kidney Cortex	Hippocampus
Esophagus Mucosa	Adipose Visceral O	Adrenal Gland	Caudate Basal Ganglia
EBT Lymphocytes	Substantia Nigra	Esophagus Muscularis	Amygdala
Substantia Nigra	Esophagus Mucosa	Hypothalamus	N Accumbens BG
Putamen Basal Ganglia	Hypothalamus	Putamen Basal Ganglia	Hypothalamus

472

Table 10. Highest affected tissues upon removal of the DisGeNet genes associated with Systemic Lupus Erythematosus (1,883) from the human GTEx database; Psoriasis genes (1,308); Breast Cancer (510) and Diabetes genes (3,134). The top half of the highest ranked tissues show no convergence with CNS-related tissues found in the mental illnesses and neurological disorders interrogated in this work. Instead, heart-related tissue, muscle skeletal tissue and tissues related to peripheral vital organs for systemic functioning are found. The bottom half of the top $\Delta\lambda$-ranked tissues are a mixture of tissues in peripheral bodily organs and brain-related tissues. The latter are from motor-control, coordination, and adaptation subcortical areas and from emotion, memory, and regulatory areas. Color code as in previous Tables.

481
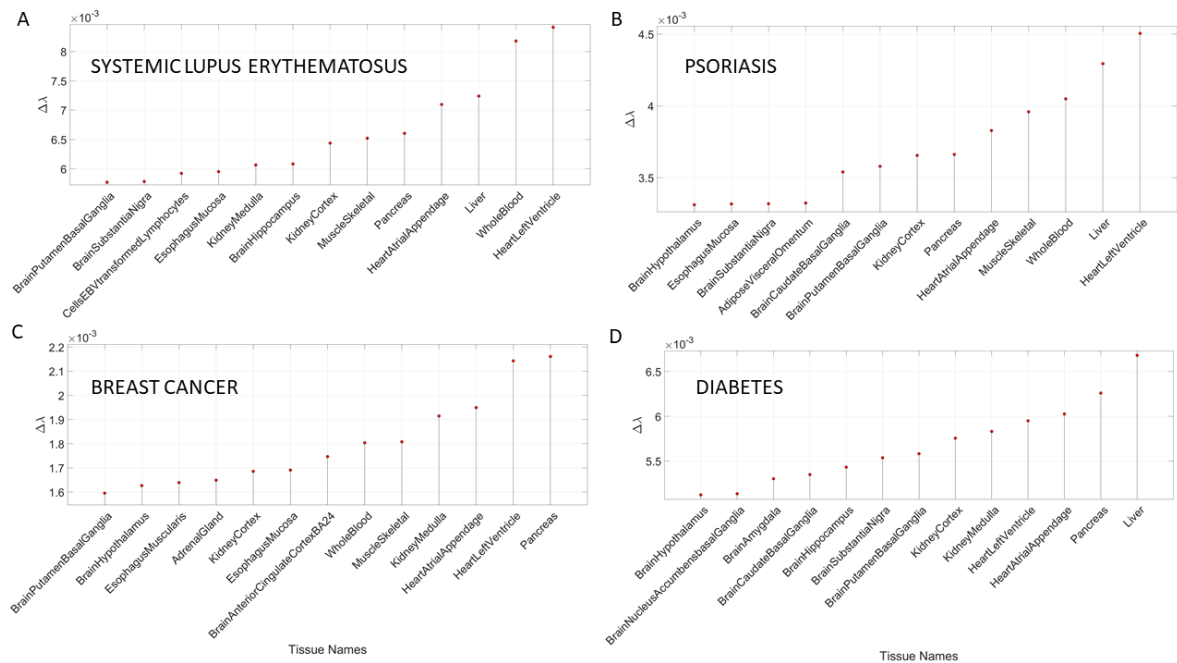
Figure 9. Maximally affected tissues in non-neurological diseases depicted in Table 10 median-ranked according to the $\Delta\lambda$-values obtained from the absolute difference between the tissues according to the full genome of the GTex database and the GTex genome without the genes associated to each disease, according to the queries to the DisGeNet portal.

The results of the maximally affected tissues upon the removal of the genes associated with these non-neurological disorders revealed a very different picture than those upon removal of the genes associated with the mental illnesses (Autism, Schizophrenia and the Depressions) and those associated with the known neurological conditions (the various forms of Ataxia, FX and Parkinson's disease.) Namely, the CNS-related tissues were less affected in these non-Neurological diseases than those related to the PNS (muscle skeletal and ANS heart) and those linked to peripheral bodily organs were the most visibly affected. The exception was Diabetes, maximally affected tissues in peripheral organs, but also CNS and PNS tissues in the tail of the top $\Delta\lambda$-ranked tissues. We next interrogate the genome in relation to mitochondrial disorders of several kinds and acquired PTSD.

Removal of Genes Associated with Mitochondrial Diseases Reveal Heart-related Tissues are Maximally Affected but PTSD is mixed

Mitochondrial Disease	Mitochondrial Myopathies	MELAS	PTSD
Heart Left Ventricle	Liver	Heart Left Ventricle	Caudate Basal Ganglia
Muscle Skeletal	Heart Left Ventricle	Liver	Kidney Medulla
Pancreas	Muscle Skeletal	Kidney Cortex	Kidney Cortex
Heart Atrial Appendage	Hippocampus	Pancreas	Pancreas
Kidney Cortex	Whole Blood	Heart Atrial Appendage	Putamen Basal Ganglia
Whole Blood	Ant Cingulate Cortex	Kidney Medulla	Heart Left Ventricle
Kidney Medulla	Brain Amygdala	Putamen Basal Ganglia	Hypothalamus
Ant Cingulate Cortex	Substantia Nigra	Esophagus Mucosa	Substantia Nigra
Putamen Basal Ganglia	Heart Atrial Appendage	Brain Amygdala	Heart Atrial Appendage
N Accumbens BG	Putamen Basal Ganglia	Adipose Visceral Omen	N Accumbens BG
Adrenal Gland	Spinal Cord	Artery Coronary	Hippocampus
Hippocampus	Pancreas	Stomach	Colon Sigmoid
Caudate Basal Ganglia	Hypothalamus	Colon Traverse	Amygdala

Table 11. Highest affected tissues upon removal of the DisGeNet genes associated with Mitochondrial disease (284) from the human GTEx database; Mitochondrial Myopathies genes (121); Mitochondrial Encephalopathy, Lactic Acidosis, and Stroke-like episodes, MELAS (81) and Post-Traumatic Stress Disorder genes (418).
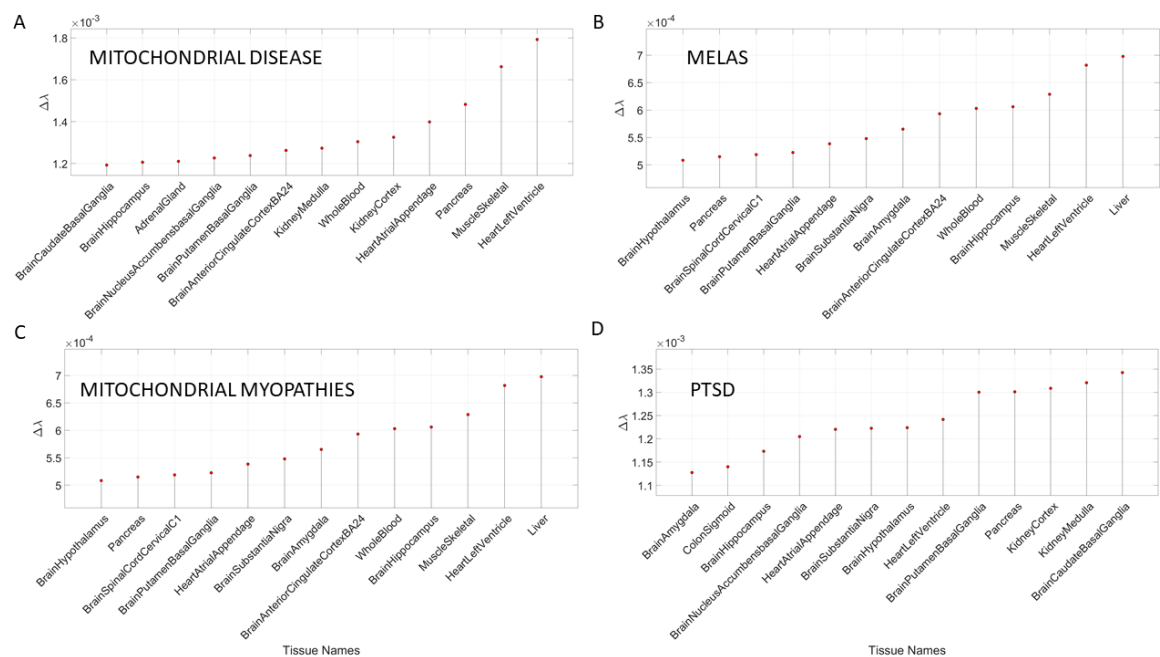
Figure 10. Maximally affected tissues in Mitochondrial Diseases and in PTSD of Table 11, shown in graphical form according to the Δλ values.

Removal of genes associated with Mitochondrial Disorders of various types from the GTex genome, according to the genes in the DisGeNet portal, reveal a mixture of tissues associated with peripheral vital organs for systemic functions, heart-related and muscle skeletal tissues and CNS-related tissues. The top half of the highest ranked tissues in Mitochondrial Disease shows affected tissues related to the heart, muscle skeletal and peripheral organs, while the bottom half shows more involvement of brain-related tissues in sub-cortical regions of motor control. In contrast, Mitochondrial Myopathies show a predominance of CNS-related tissues, including the brain and spinal cord, with top Δλ-ranked tissues related to the heart and muscle skeletal tissues. MELAS shows a predominance of tissues associated with peripheral vital organs for systemic function and heart-related tissues. Only two brain regions for motor control and emotion are present in the bottom ranked tissues of the topmost affected tissues.

The case of acquired PTSD also reveals a mixture of tissues from brain, heart, and peripheral organs. There we see maximally affected tissues linked to sub-cortical regions of the brain involved in motor control, adaptation, learning, and coordination intermixed with tissues linked to peripheral bodily organs (like the kidneys) and the autonomic systems' heart. Furthermore, we see also tissues linked to the hypothalamus, a regulatory brain structure. Figure 10 shows the graphs of the Δλ-difference median ranked as in the previous cases, for these disorders.

We summarize the results across all 54 tissues (in alphabetical order from left to right) in Figure 11. Here a color map depicts the values of Δλ normalized for each disease (along the rows) across the tissues (columns) by dividing by the maximum Δλ value of each row. The pattens reveal that neurological disorders and mental illnesses shared the tissues related to the brain as the maximally affected. They also reveal that the whole blood tissue is not as affected in the mental illnesses as in the neurological disorders (marking a point of divergence that warrants further investigation). Heart-related tissues and muscle skeletal are also shared between these mental illnesses and neurological disorders when the genes specific to each disorder are removed from the GTex genome. Interestingly, the pancreas shows as a tissue from a peripheral bodily organ commonly affected across most disorders and diseases interrogated here (with lower value in the neurological ones.)

The non-neurological diseases reveal less involvement of the CNS-related tissues but highly overlapping with the heart and muscle skeletal tissues. Tissues linked to the kidneys, liver and pancreas are also maximally affected in these diseases, with colon cancer showing an interesting pattern whereby the pancreas reveals top normalized Δλ value. Figure 12 summarizes the patterns

535   in binary by turning ON (yellow) values above .8 considered high and OFF those below (blue). This
536   cutoff is chosen to further highlight overlapping and differences across diseases based on high $\Delta\lambda$.
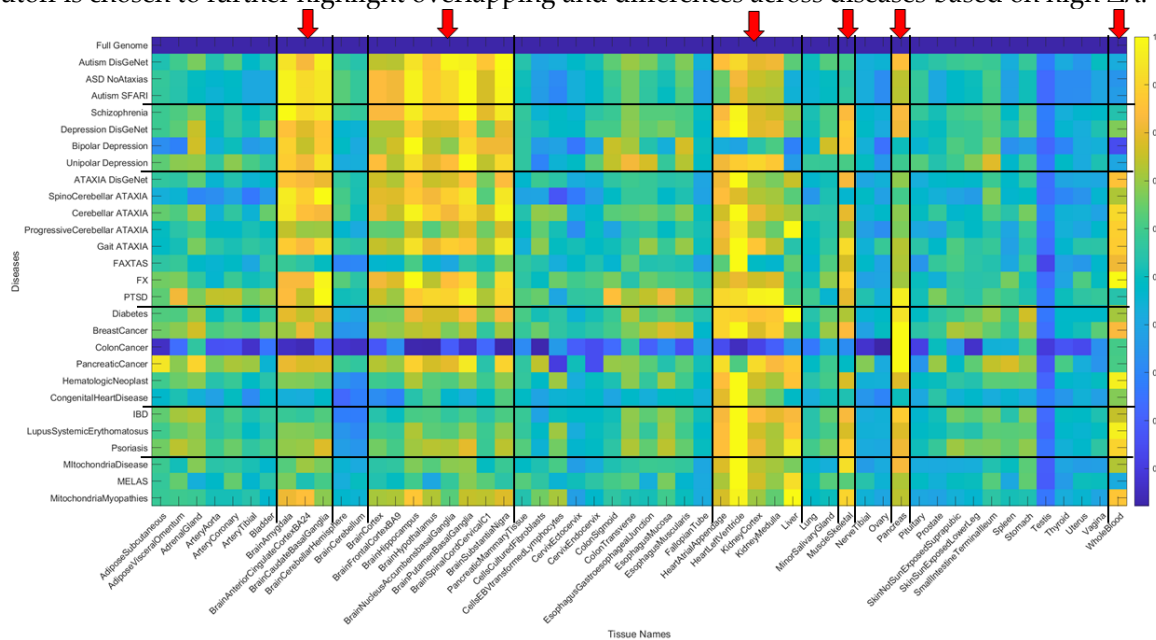

537
538   **Figure 11**. Summary of disorders/diseases (27 rows) x tissues (54 columns) in alphabetical order. Entries are $\Delta\lambda$
539   (difference with respect to the genes' expression values from the full Gtex genome), values normalized by the
540   maximum across the tissues for each row (disorder/disease.) First row is the 0-$\Delta\lambda$ difference reference from full
541   genome. Red arrows mark the maximally affected tissues across all diseases, showing mental illnesses on the
542   top, followed by neurological disorders, including several types of cancer, auto-immune disorders, diabetes, etc.
543   Black lines delineate blocks of diseases (rows) and blocks of genes' expression on tissues (columns.)
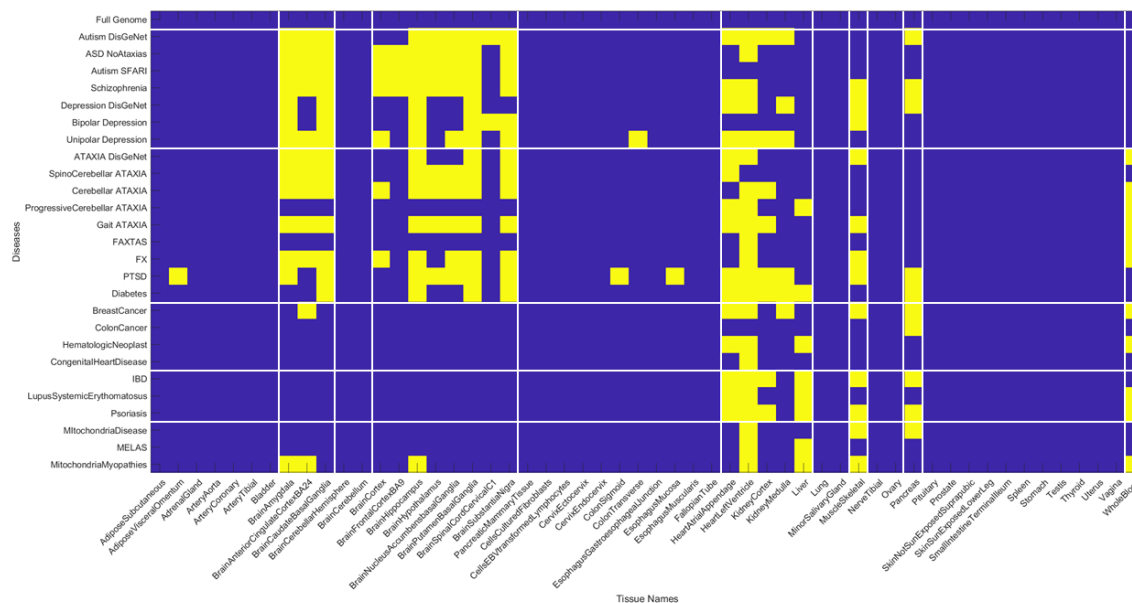
544
545   **Figure 12**. Binary version of matrix in Figure 11 upon thresholding by high normalized $\Delta\lambda$ value of .8 shows
546   the overlap across mental illnesses, neurological disorders and non-neurological diseases are primarily in the
547   heart tissues, the muscle skeletal and organs like the pancreas, liver, and kidney. Whole blood tissue is shared
548   between neurological and non-neurological disorders but not present in the mental illnesses. Brain tissues are
549   shared between mental illnesses and neurological disorders.

550   The patterns revealed by the normalized $\Delta\lambda$ high values show convergence of maximally
551   affected brain tissues in mental illnesses with the neurological disorders but not with the non-
552   neurological ones (except for diabetes which does affect some brain regions). The Mitochondrial
553   diseases do not show the same intensity of the CNS-related $\Delta\lambda$ values as the mental illnesses and the

554　neurological disorders do, but they do share the heart and muscle skeletal patterns with all the
555　examined diseases and disorders. This is interesting, given that some of the children with various
556　forms of mitochondrial disorders may receive diagnoses of autism. In summary, there is clear overlap
557　between mental illness and neurological disorders suggesting involvement of the central nervous
558　systems in both, but also contributions from the peripheral nervous systems, particularly the heart,
559　the muscle skeletal tissues and to a lesser degree, tissues of peripheral organs. The latter are most
560　affected in the non-neurological diseases. Figure 13 shows the most affected tissue in each
561　disease/disorder (also depicted on Appendix B Table.)
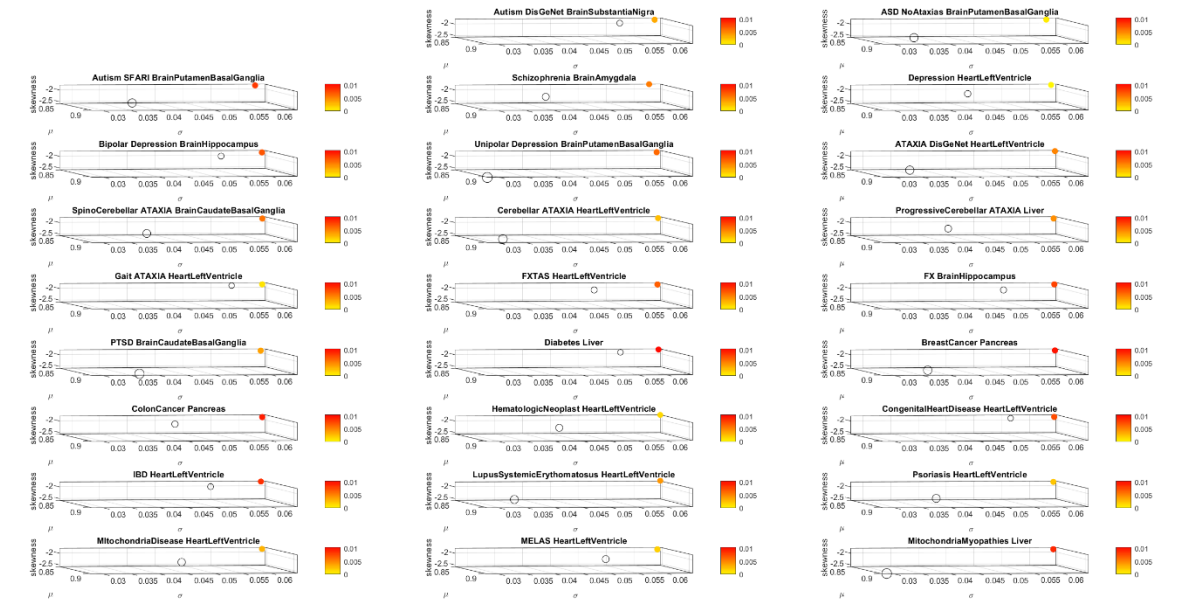
562
563　**Figure 13.** Stochastic shift of genes expression on tissue, after genes removal for each disease/disorder under
564　consideration (colored circle) relative to the genes expression on that tissue using the full genome in GTex (open
565　circles). The top affected tissue in each disorder/disease under consideration was selected according to the
566　maximum $\Delta\lambda$ value, the absolute difference between the empirically estimated λ using MLE of the exponential
567　distribution rate parameter for the full GTex genome and for the genome minus the genes associated to each
568　disease/disorder. Since the exponential distribution is a particular case of the continuous Gamma family of
569　probability distributions (when the Gamma shape parameter is 1), we also used MLE to estimate the shape and
570　scale Gamma parameters and the 4 Gamma moments, plotted here in a 5 dimensional parameter space. Along
571　the x-axis we plot the empirically estimated Gamma mean; along the y-axis we plot the Gamma variance; along
572　the z-axis we plot the Gamma skewness and the kurtosis is used to represent the size of the marker (more kurtotic
573　distributions have higher value, i.e. larger circle.) The fifth dimension is the color representing the $\Delta\lambda$ (see also
574　Methods Figure 3C visualizing one single disease and all 54 tissues or summarizing the top 13 median-ranked
575　tissues as those most affected).

4. Discussion

577　　　This work interrogated the human genome by removing genes associated with various diseases
578　and comparing the outcome from the remaining genes' expression on 54 tissues commonly examined
579　in the GTex portal. These tissues involve parts of the central nervous systems (the brain and the spinal
580　cord), parts of the peripheral nervous systems (the muscle skeletal tissue and the heart tissues as part
581　of the autonomic nervous system) which we grouped as in the Figure 1B. Other tissues are from
582　peripheral vital organs for systemic bodily functions (whole blood, pancreas, liver, kidneys, lungs,
583　etc.)
584　　　We compared mental illnesses such as Autism, Schizophrenia, and various types of Mental
585　Depression (including Unipolar and Bipolar), with well-known neurological disorders such as
586　different types of Ataxias, Fragile X and Parkinson's Disease. We found convergence between the
587　tissues maximally affected by the removal of disease-associated genes across these mental illnesses
588　and neurological disorders. CNS-related tissues in subcortical regions of the brain related to motor

589   control, motor learning, motor coordination, and motor adaptation, as well as memory and emotion,
590   were predominantly maximally affected by the corresponding genes' removal across both the mental
591   illnesses and the neurological disorders. This convergence demonstrates overlap between Psychiatric
592   and Neurological conditions with specific involvement of motor, memory, and emotional axes. In
593   Autism and FX, we obtained congruent results on maximally affected tissues both using for removal
594   the genes from the SFARI-Autism database, and the genes upon querying the DisGeNet portal. In
595   addition to the genes reported by querying DisGeNet, we also used the genes reported in the
596   literature for Schizophrenia and Depression, and for Parkinson's disease. We found congruence in
597   all cases.

598   To further test our hypothesis that these mental illnesses are *disorders of the nervous systems* and
599   that removing the gene pool associated with them gives rise to overlapping tissues related to the CNS
600   functioning, we also queried DisGeNet about other non-neurological diseases. We found that in such
601   cases, the predominance of maximally affected tissues was on tissues associated with peripheral vital
602   bodily organs related to the disease, as pancreas, kidney, liver, and colon transverse in colon cancer.
603   Furthermore, several of these diseases had maximally affected heart-related tissues and whole blood.
604   Other cases also revealed a predominance of peripheral organs. We lastly, interrogated the genome
605   in relation to Mitochondrial diseases and to acquired PTSD. In these cases, we hypothesized and
606   confirmed a mixture of tissues related to peripheral organs (for mitochondrial diseases) and to the
607   CNS (for PTSD.)

608   In the case of Mitochondrial diseases, the heart-related tissues were revealed as the most affected
609   along with muscle-skeletal tissue. Furthermore CNS-related tissues were more affected by the genes'
610   removal in Mitochondrial Myopathies (i.e. the amygdala and the anterior cingulate cortex), as
611   compared to MELAS or to Mitochondrial disease. The common thread across all three types of
612   Mitochondrial related disorders was the heart-related tissues. The case of acquired PTSD showed a
613   mixture of CNS-related tissues, tissues related to bodily peripheral organs, and heart-related tissues.
614   The kidney and pancreas were also affected in PTSD. When we examined the maximum $\Delta\lambda$ for each
615   disorder/disease under examination, we found that the basal ganglia was maximally affected in
616   Autism, Unipolar Depression, Spinocerebellar Ataxia, and PTSD), while the heart left ventricle tissue
617   was maximally affected in Depression, Ataxia, Cerebellar Ataxia, Gait Ataxia and Fragile X Tremor
618   Ataxia Syndrome (FXTAS). This result indicates overlap between the Psychiatric mental illnesses and
619   the Neurological disorders. It also shows the importance of examining mental illnesses in a more
620   systemic way that includes the Autonomic Nervous Systems of the PNS.

621   This test on non-neurological illnesses served as a control to show that removal of the genes
622   associated with each disorder did have specificity with the disorder, and yet, very different outcome
623   when comparing the mental illnesses to the non-neurological disorders. Among the top tissues
624   affected across non-neurological diseases, the pancreas was maximally affected by the removal of
625   disease-associated genes in Breast and Colon Cancer, while the liver was maximally affected in
626   Diabetes. The heart left ventricle was maximally affected across autoimmune disorders like psoriasis,
627   Lupus Systemic Erythematosus, and Irritable Bowel Disease (IBD). The heart was also maximally
628   affected in Hematologic Neoplast and Congenital Heart Disease, Mitochondrial Disease and MELAS,
629   in contrast to the Mitochondrial Myopathies which showed the liver as the maximally affected tissue.

630   This exercise demonstrated that despite the stochastic nature of genes' expression, upon their
631   removal and random recombination, there is convergence across Psychiatric and Neurological
632   disorders, thus potentially rendering both as *disorders of the nervous systems*. We found in both cases
633   a strong prevalence of the CNS, but also found important differences in tissues from the PNS,
634   including the heart and the muscle skeletal tissues involved in both mental illnesses and neurological
635   disorders. Because of these convergences, and the fact that there are treatments and accommodations
636   to help persons with neurological disorders, it may be possible to leverage some of those types of
637   bodily-based supports to help persons with mental illnesses. Behaviors that are described by
638   observation to define mental illnesses can now be connected with underlying tissues involved in
639   voluntary, involuntary, and autonomic function across the CNS and the PNS and mapped to the
640   genome, thus closing the present gap between behaviors and genomics in the Precision Medicine

641 knowledge network. In this sense, the present methods offer a new way to interrogate the genome
642 and link tissues with behavioral phenotyping.
643   A surprising finding here is the potential contributions of peripheral structures and organs to
644 mental illness. Tissues of the autonomic nervous systems were maximally impacted by the removal
645 of the genes associated with these mental illnesses, as was the muscle skeletal tissue among the top
646 ranked ones. Tissues associated to sub-cortical brain regions necessary for motor control, learning,
647 adaptation, and coordination (basal ganglia and striatum) were highly impacted by the removal of
648 the genes in both mental illnesses and neurological disorders, along with those tissues important for
649 memory (hippocampus), emotion (amygdala) and regulation (hypothalamus). Surprisingly, we did
650 not see cerebellum-related tissues among the top affected by the removal of the genes (even in the
651 ataxias) where we do know that the cerebellum plays a large role [21-24]. This was also the case in
652 autism, where we know the cerebellum has been implicated [25, 26].
653   Lastly, the autoimmune disorders that we examined had very different brain-tissue patterns
654 from the mental illnesses and neurological conditions but shared the heart-related tissues and the
655 muscle skeletal tissue. In this sense, the contributions from the peripheral systems to mental illnesses
656 and to auto-immune disorders seems important. While blood tissue marked a departure of
657 neurological disorders from mental illnesses, as it was maximally affected in neurological disorders
658 but not in the mental illnesses. All and all these genes' removal revealed surprising results that invite
659 rethinking how we may want to describe, diagnose, and treat mental illnesses in general.

Caveats and Future Directions

661   Although we found evidence that the mental illnesses and neurological disorders have
662 remarkable overlap in the types of brain tissues that are maximally affected by the removal of their
663 corresponding associated genes, we recognize that genes' removal is a crude way to interrogate the
664 human genome and its expression of the 54 tissues of the GTex database. Future work will aim at
665 developing more sophisticated methods to explore genes' over expression too, and to build
666 simulations of the use of these methods in *e.g.* combination with transcriptome evolution during
667 neuronal differentiation in the development of cell lines from induced pluripotent stem cells. This
668 will be important to assess asynchronous genes' expression behaviors over time when cell lines
669 differentiate into neuronal types. Full transcriptome dynamic interrogation longitudinally, over time
670 is now possible using these stochastic analyses in combination with the various data repositories
671 featuring disease-associated genes.
672   The present work merely scratches the surface on possible new ways to interrogate the human
673 genome in relation to diseases of all types (not just mental or neurological), to possibly build
674 comparative models of outcomes *in tissues* that can be related to behavioral phenotypic
675 manifestations of the clinical disorder. In this sense, the work presented here can help bridge the gap
676 between behavioral description of a mental illness, or a neurological disorder and its genetic
677 underpinnings via the affected tissues. Combining this approach with the new wave of digital
678 biomarkers that describe human behavior digitally at a microscopic level [4, 27-30], using objective
679 means and a finer level of granularity beyond naked eye detection, could help us redefine many
680 Psychiatric disorders and medical conditions under the Precision Medicine paradigm.

5. Conclusions

682   We here offer a new roadmap to reframe Psychiatry using the Precision Medicine paradigm. The
683 new stochastic approach can initiate the steps to connect behavioral phenotypic description from
684 clinical observation, with the underlying neurobiology of mental illnesses. Borrowing knowledge
685 from Neurology and Brain Science, it will be possible to shift Psychiatry from an art to a quantitative
686 objective science under the tenets of Precision Medicine, by integrating all layers of the knowledge
687 network. This would help design personalized targeted treatments utilizing the person's genome,
688 localizing the most affected tissues defining central nervous systems functions and distinguishing
689 those from tissues related to vital organs for systemic functions. This new approach could potentially
690 mark the beginning of a transformative era in Mental Health.

691   **6. Patents**

692   EBT holds the US Patent "Methods and Systems for the Diagnoses and Treatments of Nervous
693   Systems Disorders" combined in the paper as micro-movement spikes, MMS data type and
694   continuous Gamma probability distributions family empirical estimation.

706   **Appendix A**

707   We estimate the likelihood $L\left(\lambda \mid x_1, x_2, ..., x_n\right)$ where x_i is the series of counts representing the
708   gene expression on each given tissue and i ranges from *1* to *n*, the number of genes.
709

$$
\begin{aligned}
L\left(\lambda \mid x_1, x_2, ..., x_n\right) &= \lambda e^{-\lambda x_1} \lambda e^{-\lambda x_2} ... \lambda e^{-\lambda x_n} \\
&= \lambda^n \left(e^{-\lambda x_1} e^{-\lambda x_2} ... e^{-\lambda x_n}\right) \\
&= \lambda^n \left(e^{-\lambda\left(x_1 + x_2 + ... + x_n\right)}\right)
\end{aligned}
\tag{A1}
$$

711
712   To obtain the maximum likelihood, we take the derivative of the likelihood in equation (A1) and
713   set it to 0 (since the derivative is 0 at the maximum likelihood value).

$$
\frac{d}{d\lambda} L\left(\lambda \mid x_1, x_2, ..., x_n\right) = \frac{d}{d\lambda} \lambda^n \left(e^{-\lambda\left(x_1 + x_2 + ... + x_n\right)}\right)
\tag{A2}
$$

715   We take the *log* here because the derivative of the function and the derivative of the *log* of the
716   function equals *0* at the same point, so for the purposes of finding where the derivative is *0*, the
717   original function in equation (A2) and the *log* of it are interchangeable.
718

$$
\begin{aligned}
&\frac{d}{d\lambda} \log\left(\lambda^n \left(e^{-\lambda\left(x_1 + x_2 + ... + x_n\right)}\right)\right) = \frac{d}{d\lambda} \log \lambda^n + \log\left(e^{-\lambda\left(x_1 + x_2 + ... + x_n\right)}\right) \\
&\frac{d}{d\lambda} n \log \lambda - \lambda\left(x_1 + x_2 + ... + x_n\right) = 0 \\
&n \frac{1}{\lambda} - \left(x_1 + x_2 + ... + x_n\right) = 0 \\
&\lambda = \frac{n}{\left(x_1 + x_2 + ... + x_n\right)}
\end{aligned}
\tag{A3}
$$

720
721
722   And with this result in equation (A3), we can obtain the maximum likelihood estimate of each
723   λ_j, given all the 56,146 genes expressed with some random value for each of the $j = 1:54$ tissues.
724

725   **Appendix B**

726   The data file name from the GTEx Portal https://www.gtexportal.org/home/datasets used in this
727   paper is GTEx_Analysis_2017-06-05_v8_RNASeQCv1.1.9_gene_median_tpm.gct.csv
728   The data file name from SFARI Gene is located at https://gene.sfari.org/database/human-gene/
729   and named SFARI-Gene_genes_03-04-2020release_03-05-2020export.csv
730
731
732   Appendix B Table: Top affected tissue in each disease/disorder according to the maximum $\Delta\lambda$ value
733   (See Figure 13 in the main text.)
734

Disease/Disorder	Maximally Affected Tissue
Autism DisGeNet	BrainSubstantiaNigra
ASD NoAtaxias	BrainPutamenBasalGanglia
Autism SFARI	BrainPutamenBasalGanglia
Schizophrenia	BrainAmygdala
Depression DisGeNet	HeartLeftVentricle
Bipolar Depression	BrainHippocampus
Unipolar Depression	BrainPutamenBasalGanglia
ATAXIA DisGeNet	HeartLeftVentricle
SpinoCerebellar ATAXIA	BrainCaudateBasalGanglia
Cerebellar ATAXIA	HeartLeftVentricle
ProgressiveCerebellar ATAXIA	Liver
Gait ATAXIA	HeartLeftVentricle
FAXTAS	HeartLeftVentricle
FX	BrainHippocampus
PTSD	BrainCaudateBasalGanglia
Diabetes	Liver
BreastCancer	Pancreas
ColonCancer	Pancreas
HematologicNeoplast	HeartLeftVentricle
CongenitalHeartDisease	HeartLeftVentricle
IBD	HeartLeftVentricle
LupusSystemicErythomatosus	HeartLeftVentricle
Psoriasis	HeartLeftVentricle
MItochondriaDisease	HeartLeftVentricle
MELAS	HeartLeftVentricle
MitochondriaMyopathies	Liver

735

736   **References**

737   1.   Hawgood, S., et al., *Precision medicine: Beyond the inflection point.* Sci Transl Med, 2015. 7(300):
738   p. 300ps17.
739   2.   Adams, R.A., Q.J. Huys, and J.P. Roiser, *Computational Psychiatry: towards a mathematically*
740   *informed understanding of mental illness.* J Neurol Neurosurg Psychiatry, 2016. 87(1): p. 53-63.

741 3.   *American Psychiatric Association: Diagnostic and statistical manual of mental disorders (5th ed.).*
742   2013.

743 4.   Nguyen, J., et al., *Schizophrenia: The micro-movements perspective.* Neuropsychologia, 2016. 85:
744   p. 310-26.

745 5.   Rogers, D.M., *Motor disorder in psychiatry : towards a neurological psychiatry.* 1992, Chichester ;
746   New York: J. Wiley & Sons. viii, 159 p.

747 6.   Bird, T.D., *Hereditary Ataxia Overview*, in *GeneReviews((R))*, M.P. Adam, et al., Editors. 1993:
748   Seattle (WA).

749 7.   Jayadev, S. and T.D. Bird, *Hereditary ataxias: overview.* Genet Med, 2013. 15(9): p. 673-83.

750 8.   Chinnery, P.F., *Mitochondrial Disorders Overview*, in *GeneReviews((R))*, M.P. Adam, et al.,
751   Editors. 1993: Seattle (WA).

752 9.   Domingo, A., et al., *New insights into the genetics of X-linked dystonia-parkinsonism (XDP,*
753   *DYT3).* Eur J Hum Genet, 2015. 23(10): p. 1334-40.

754 10.   Klein, C. and A. Westenberger, *Genetics of Parkinson's disease.* Cold Spring Harb Perspect Med,
755   2012. 2(1): p. a008888.

756 11.   Scott, W.K., et al., *Complete genomic screen in Parkinson disease: evidence for multiple genes.*
757   JAMA, 2001. 286(18): p. 2239-44.

758 12.   Hicks, A.A., et al., *A susceptibility gene for late-onset idiopathic Parkinson's disease.* Ann Neurol,
759   2002. 52(5): p. 549-55.

760 13.   Valente, E.M., et al., *PARK6-linked parkinsonism occurs in several European families.* Ann Neurol,
761   2002. 51(1): p. 14-8.

762 14.   Kay, D.M., et al., *Parkinson's disease and LRRK2: frequency of a common mutation in U.S.*
763   *movement disorder clinics.* Mov Disord, 2006. 21(4): p. 519-23.

764 15.   Salleh, M.R., *The genetics of schizophrenia.* Malays J Med Sci, 2004. 11(2): p. 3-11.

765 16.   Schizophrenia Working Group of the Psychiatric Genomics, C., *Biological insights from 108*
766   *schizophrenia-associated genetic loci.* Nature, 2014. 511(7510): p. 421-7.

767 17.   Kotlar, A.V., et al., *New discoveries in schizophrenia genetics reveal neurobiological pathways: A*
768   *review of recent findings.* Eur J Med Genet, 2015. 58(12): p. 704-14.

769 18.   Fabbri, C. and A. Serretti, *Role of 108 schizophrenia-associated loci in modulating*
770   *psychopathological dimensions in schizophrenia and bipolar disorder.* Am J Med Genet B
771   Neuropsychiatr Genet, 2017. 174(7): p. 757-764.

772 19.   Forsyth, J.K. and R.F. Asarnow, *Genetics of Childhood-onset Schizophrenia 2019 Update.* Child
773   Adolesc Psychiatr Clin N Am, 2020. 29(1): p. 157-170.

774 20.   Zamanpoor, M., *Schizophrenia in a genomic era: a review from the pathogenesis, genetic and*
775   *environmental etiology to diagnosis and treatment insights.* Psychiatr Genet, 2020. 30(1): p. 1-9.

776 21.   Soong, B.W. and P.J. Morrison, *Spinocerebellar ataxias.* Handb Clin Neurol, 2018. 155: p. 143-
777   174.

778 22.   Fujioka, S., C. Sundal, and Z.K. Wszolek, *Autosomal dominant cerebellar ataxia type III: a review*
779   *of the phenotypic and genotypic characterics.* Orphanet J Rare Dis, 2013. 8: p. 14.

780 23.   Sailer, A. and H. Houlden, *Recent advances in the genetics of cerebellar ataxias.* Curr Neurol
781   Neurosci Rep, 2012. 12(3): p. 227-36.

782 24.   Mariotti, C. and S. Di Donato, *Cerebellar/spinocerebellar syndromes.* Neurol Sci, 2001. 22 Suppl
783   2: p. S88-92.

784   25.   Mosconi, M.W., et al., *The role of cerebellar circuitry alterations in the pathophysiology of autism*
785   *spectrum disorders.* Front Neurosci, 2015. 9: p. 296.

786   26.   Mostofsky, S.H., et al., *Decreased connectivity and cerebellar activity in autism during motor task*
787   *performance.* Brain, 2009. 132(Pt 9): p. 2413-25.

788   27.   Ryu, J., et al., *Dynamic Digital Biomarkers of Motor and Cognitive Function in Parkinson's Disease.*
789   J Vis Exp, 2019(149).

790   28.   Wu, D., et al., *A Biomarker Characterizing Neurodevelopment with applications in Autism.* Sci Rep,
791   2018. 8(1): p. 614.

792   29.   Torres, E.B., J. Vero, and R. Rai, *Statistical Platform for Individualized Behavioral Analyses Using*
793   *Biophysical Micro-Movement Spikes.* Sensors (Basel), 2018. 18(4).

794   30.   Torres, E.B., et al., *Autism: the micro-movement perspective.* Front Integr Neurosci, 2013. 7: p.
795   32.

796

797