GENOMEWIDE GENE EXPRESSION ANALYSIS TO INVESTIGATE SYMPTOM ONSET AND EXACERBATION IN TOURETTE SYNDROME

by

Hanrui Wu

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science

Interdisciplinary Life Science (PULSe)

West Lafayette, Indiana

December 2020

THE PURDUE GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Peristera Paschou, Chair

Department of Biological Sciences

Dr. Donna M. Fekete

Department of Biological Sciences

Dr. Kari J. Ekenstedt

Department of Basic Medical Sciences

Dr. Petros Drineas

Department of Computer Science

Approved by:

Dr. Jason R. Cannon

Dr. Janice P. Evans

ACKNOWLEDGMENTS

I would like to express my deepest appreciation to my PI, Professor Peristera Paschou; her constant support and guidance shaped me into a better student and person.

To my committee members Professor Drineas, Professor Fekete and Professor Ekenstedt, I am thankful for their contributions to my thesis.

I would also like to truly thank Dr. Shanmukha Sampath Padmanabhuni for his patience and encouragement throughout my Master's study, from whom I have learned a lot about bioinformatics and human genetics. I would also like to thank the EMTICS project.

Thanks should also go to my colleagues Zhiyu Yang, Apostolia Topaloudi and Dr. Fotis Tsetsos, whose help cannot be underestimated.

Finally, I am forever grateful to my parents and my wife for always believing in me and encouraging me to follow my dreams.

TABLE OF CONTENTS

ABSTRACT

Tourette Syndrome (TS) is a neurodevelopmental disorder that starts in childhood. It is marked by multiple motor and vocal tics and a fluctuating course with recurrent time periods of symptom exacerbation followed by symptom remission. TS is considered a disorder of complex etiology that involves the interaction of multiple genetic and environmental factors. Previous studies have implicated regulation of the immune system in TS patients, although this remains a matter of debate.

Here we present a genome-wide gene expression study comparing data from TS patients at time points of symptom exacerbation and remission as well as upon symptom onset for newly diagnosed patients, with differentially gene expression analysis, co-expression analysis and pathway analysis. Although the fold change of each regulated gene was very mild, it is the first time that changes in the regulation of the immune system have been implicated at different time points in TS patients.

INTRODUCTION

Clinical Features of Tourette Syndrome

Tourette Syndrome (TS) is a complex neurodevelopmental disorder with onset in childhood (Paschou *et al*., 2014). It is marked by multiple vocal and motor tics (Cath *et al*., 2011). Tics are sudden, rapid, recurrent, non-rhythmic movements or vocalizations preceded by a premonitory urge. They are also suggestible and suppressible and can be triggered by stress, temperature changes, illness, and fatigue (Murphy *et al*., 2013). TS can cause physical discomfort, social difficulties, and it can lead to reduced quality of life for patients and their families (Julio *et al*., 2018, Eapen *et al*., 2016). In most cases, tics have an onset between four and six years old, and then reach the worst severity between 10 and 12 years old. Usually, the severity of tics decreases during adolescence; around 75% of the TS patients will have greatly decreased symptoms, and more than 30% of the patients will eventually be tic free (Leckman *et al*., 1998, Bloch & Leckman, 2014).

Apart from tics, most TS patients also present with associated neuropsychiatric comorbidities including other neurodevelopmental disorders of onset in childhood or adolescence. Around 40-60% of patients also have obsessive compulsive disorder (OCD), 60-90% of the TS patients also have obsessive compulsive behaviors (OCB), and 60% of the TS patients also have attention-deficit hyperactivity disorder (ADHD)(Robertson *et al*., 2017, Kumar *et al.*, 2016).

Genetics of Tourette Syndrome

Like many other neurological disorders, TS has a complex etiology. The exact cause of TS remains unknown and is attributed to intricate networks of genetic and environmental factors, including gene to gene and gene to environment interactions (Qi *et al*., 2017, Panoutsopoulou *et al*., 2009). Unraveling these complex biological processes has been challenging because of the increased causal and phenotypic heterogeneity (Jin *et al*., 2014).

Previous Studies have demonstrated that TS is one of the most heritable non-Mendelian neurological disorders. A population-based study revealed the estimated heritability to be 0.77(95% CI: 0.70-0.85), where 1.0 stands for a 100% heritability (Robertson *et al*., 2017, Pauls *et al*., 2014).

It has also been previously reported that the risk of TS is significantly higher in biological relatives of TS patients than individuals in the general population (Pauls *et al*., 1991). The occurrence of TS in male vs female is approximately 4:1 (Robertson, 2008).

Candidate gene studies

Prior to the genomewide era, multiple candidate gene studies reported the association between TS and genes in dopaminergic, serotonergic, and glutamatergic pathways (Qi *et al*., 2017). The dopamine hypothesis in TS suggests that TS is associated with supersensitive dopamine receptors and overactivation of dopamine transporters (Maia *et al*., 2018). Thus, dopaminergic genes have long been investigated in relation to TS etiology.

In 1991, Comings *et al*. investigated the frequency of the A1 allele of the Taq I polymorphism of the *dopamine D2 receptor* (*DRD2*) in 147 TS patients and 314 controls and reported that the frequency of A1 allele in TS patients (44.9%) is higher than controls (24.5%) (Comings *et al*., 1991). Also, in 1996, Comings *et al*. reported a significant correlation of polymorphisms in *DRD2*, *dopamine transporter* (*DAT1*) and *dopamine beta hydroxylase* (*DBH*) genes with TS (Comings *et al*., 1996).

Cruz *et al*. investigated the prevalence of the polymorphism of 48bp repeats in the *DRD4* in OCD patients with and without tics (Cruz *et al*., 1997). They found 91% of the OCD patients with tics carried at least one copy of the 48bp repeat in *DRD4* while the prevalence of this mutation is 48% in OCD patients without tics (Cruz *et al*., 1997).

Tarnok *et al*., characterized different dopamine-related polymorphisms in 103 TS patients and their parents (Tarnok *et al*., 2007). They found a 40bp variable number tandem repeat in the *DAT1*(*dopamine transporter*) gene associated with tic severity as measured by the Yale Global Tic Severity Scale, which is a clinical tool that is often used in TS patient characterization (Tarnok *et al*., 2007). In addition, Gade *et al*. investigated 229 TS patients and 57 controls and reported that the VNTR polymorphism at the X-linked *MAO-A* gene is also associated with TS (Gade *et al*., 1998). The *monoamine oxidase A* gene (*MAO-A*) participates in dopamine and serotonin metabolism.

It has also been reported that the brain serotonin level is decreased in TS patients (Anderson *et al*., 1992). Lam *et al*., interrogated the presence of an association between TS and serotonin receptors (Lam *et al*., 1996). They found a missense nucleotide change in the *serotonin 1A receptor* (*HTR1A)* in DNA extracted from one TS patient (Lam *et al*., 1996). Also, Dehning *et al*. found a nominally significant association between two polymorphisms in *serotonin receptor 2C* (*HTR2C*) and TS patients, indicating serotonin receptor genes also play an important role in TS(Dehning *et al*., 2010).

TPH2 encodes for *tryptophan 5-hydroxylase 2*, which catalyzes the rate-limiting step in the biosynthesis of serotonin. Mössner *et al* studied 98 unrelated TS patients and 178 healthy controls and found the rs4565946 T allele in the *tryptophan hydroxylase-2* gene (*TPH2*) to be associated with TS. This suggested the synthesis of serotonin via *TPH2* is regulated in the development of TS (Mössner *et al*., 2007).

Besides dopamine and serotonin-related pathways, glutamate has also been previously implicated in TS pathophysiology; glutamate is a major excitatory neurotransmitter that plays a critical role in cortico-striatal-thalamo-cortical circuitry (Singer *et al*., 2013). The *SLC1A3* gene encodes an excitatory amino acid transporter 1 (EAAT1), which transports glutamate in the brain. A rare functional *SLC1A3* variant was found in a small number of individuals with TS with a frequency 2.4 times higher in TS patients (0.022) than healthy controls (0.009). This suggested that the regulation of glutamate in the brain related to this mutatied glutamate transporter was associated with TS (Adamczyk *et al*., 2011).

A rare functional mutation in the *HDC* gene encoding *l-histidine decarboxylase* was identified by Ercan-Sencicek *et al*. as associated with TS, after analyzing a nonconsanguineous two-generation pedigree, where the father and eight offspring were diagnosed with TS (Ercan-Sencicek *et al*., 2010). A follow-up association study of 520 European families found over-transmission of two SNPs and a significant haplotype across the histidine decarboxylase gene *(HDC*) region associated with TS (Karagiannidis *et al*., 2013).

Genomewide scans for Tourette Syndrome

More recently, several genomewide scans have implicated additional genes in TS etiology.

Yu *et al*. performed a GWAS meta-analysis in 4,839 TS patients and 9,488 controls. The study identified a genome-wide significant SNP, rs2504235, located within the *FLT3* gene on chromosome 13 (Yu *et al*., 2019). *FMS-like tyrosine kinase 3* (*FLT3*) is a transmembrane protein expressed on normal hematopoietic stem and progenitor cells and retained on malignant blasts in acute myeloid leukemia, indicating the important role of *FLT3* in immune response (*Jetani et al., 2018*).

Willsey *et al*. performed a whole exome sequencing analysis with 325 TS trios (975 samples including TS patients and their parents) from the Tourette International Collaborative Genetics cohort and 186 trios (558 samples including TS patients and their parents) from the Tourette Syndrome Association International Consortium on Genetics. They estimated that *de novo* damaging variants identified in approximately 400 genes may contribute to 12% of clinical TS cases. Specifically they reported four TS risk genes: *CELSR3* (*cadherin EGF LAG seven-pass G-type receptor 3*), *NIPBL* (*nipped-B-like*), *FN1* (*fibronectin 1*), and *WWC1* (*WW and C2 domain containing 1*). Among these genes, CELSR3 and NIPBL play important roles in neuron function: *CELSR3* codes for a receptor that plays an important role in inter-cell signaling during the development of the nervous system and participates in dopaminergic neuron axon guidance. (UniProtKB. GO) The NIPBL protein promotes cortical neuron migration during the development of CNS along with ZNF609 (Willsey *et al*., 2017).

Wang *et al* analyzed the whole-exome sequencing data in a total of 802 groups of trios (2406 samples). They replicated the association of a high confidence TS risk gene *CELSR3* identified in the above-mentioned study. They also identified a group of mutated genes enriched in cell polarity pathway in TS patients. In addition, a significant overlap of *de novo* sequence variants between OCD and TS patients was identified, and a significant overlap of *de novo* Copy Number Variation (CNV) between Autism Spectrum Disorder (ASD) and TS was also found; together this is consistent with the findings of shared genetic risk between TS and comorbidities (Wang *et al*, 2018).

Huang *et al*. analyzed a European ancestry sample of 2,434 TS patients and 4,093 controls for rare CNVs using SNP microarray data. They identified deletions in the *NRXN1* gene and duplications in the *CNTN6* gene that were each associated with an increased risk of TS. The NRXN1 protein belongs to the neurexin family, which plays an important role in neurotransmission and the formation of synaptic contacts. The *CNT6* gene encodes a glycosylphosphatidylinositol (GPI)-anchored neuronal membrane protein, expressed primarily in the central nervous system, that participates in the cell adhesion and formation of axon connections in the developing nervous system. (Wheeler *et al*., 2007, Huang *et al*., 2017).

Although several genetic mutations have been identified associated with TS in population, the exact cause of TS is still unknown due to the complex etiology of TS. It is suggested that exposures of environmental factors may be associated with increased risk of tic onset and exacerbation of TS.

Environmental factors and Tourette Syndrome

Browne *et al*. calculated incidence rates per 1,000 people for TS and chronic tic disorder (TS/CT) and pediatric-onset obsessive-compulsive disorder (OCD) on 73,073 singleton pregnancies from the Danish National Birth Cohort. They reported a 66% increased risk for TS/CT associated with heavy smoking, which is more than 10 cigarettes per day, during pregnancy. Robertson *et al* did a study with children aged <16 who were newly diagnosed with Tourette Syndrome. The DSM-III-R questionnaire was used to assess the severity of neurological disorders. Among 57 subjects, tics were aggravated by stress in 30 patients (Robertson *et al*., 2002, Hoekstra *et al*., 2013, Browne *et al*., 2016, Mattews *et al*., 2014).

Autoimmunity and Tourette Syndrome

In the year 1929, Selling investigated three patients with tic disorders and sinus infection. The severity of tics was reduced on those patients after they underwent sinus surgery, which suggested that tic onset partially originated from infections of certain tissues. Beta hemolytic streptococcus was cultured from operation-obtained material obtained at operation from two patients. Pediatric Autoimmune Neuropsychiatric Disorders Associated with Streptococcus infections (PANDAS) is hypothesized to describe a subset of childhood tic disorders and OCD triggered by group-A beta-hemolytic Streptococcus (GAS) pyogenes infection. Although the exact mechanism remains

unclear, the most convincing theory to explain the etiology of PANDAS is molecular mimicry, which results in the antibodies intended to target GAS mistakenly targeting brain protein(s) instead. These antibodies cause direct stimulation or blockade of basal ganglia receptors and promote inflammation in basal ganglia (Murphy *et al*., 2010).

In addition, a protein array study probed 102 GAS proteins with sera from tic patients and children without tics, and revealed that a portion of GAS antigens were recognized by sera from tic patients. But this portion of GAS antigen were not recognized by sera from children without tic, indicated that patients with tic disorders showed an enhanced immune response against GAS antigens (Bombaci *et al*., 2009). Although it is still not clear how GAS infections impact the central nervous system (CNS), it is assumed that GAS infections are not the primary cause of the occurrence of CNS diseases. Instead, GAS infections may act as co-factors to increase the risk for disease occurrence or trigger the disorder outcome (Bombaci *et al*., 2009, Martino *et al*., 2011, Hoekstra *et al*., 2013).

On the other hand, a histaminergic etiology hypothesis for TS has also been recently debated. As mentioned earlier, the *HDC* gene, which is a rate-limiting enzyme that catalyzes the decarboxylation of histidine to form histamine, has been previously implicated in TS genetic risk (Ercan-Sencicek *et al*., 2010, Karagiannidis *et al*., 2013). Histamine is produced by immune cells and a group of neurons in the tuberomammillary nucleus (Frick *et al*., 2016). Histamine is involved in the inflammatory response, the immune response, and neurotransmission in the brain, spinal cord, and uterus (Marieb *et al*., 2001, Andersen *et al*., 2015). The allergen-triggered release of histamine mediates allergic symptoms (Ausländer *et al*., 2014).

The knockout of the *Hdc* gene produces a mouse model that exhibits the behavioral and neurochemical symptoms seen in TS patients (Maximiliano *et al*., 2014, Rapanelli *et al*., 2016). This suggested that altered histamine levels regulate brain activity (Baldan *et al*., 2014). In *Hdc* knockout mice, the level of histamine is reduced and thus the striatal dopamine level is upregulated. Tic-like stereotypies of *Hdc* knockout mice can be eliminated by the dopamine D2 antagonist haloperidol, which is used to treat the symptoms of TS, and tics can also be eliminated by brain infusion of histamine. In *Hdc* knockout mice, increased dopamine levels were detected, the expression of the *Fos* gene was up-regulated, and dopamine receptor D2/D3 binding to dopamine

was altered. These findings indicate that *Hdc* knockout mice can be used to mimic TS outcomes in humans.

Further pointing to immunity playing a role in TS etiology, allergies often appear in TS patients. To study the increased association of TS with symptoms of allergy, Ho *et al*. evaluated 72 TS patients in Taiwan, including 65 boys and seven girls, using the Multiple Allergens Simultaneous Tests (MAST). A total of 41 patients (56.9%) were diagnosed positive for allergies. Compared with the prevalence of allergy in Taiwan reported by The International Study of Asthma and Allergy in Childhood (1994), (44.3%,) there was a significantly increased prevalence of allergy in TS patience (Ho *et al*., 1999).

Another study in Taiwan investigated the linkage between TS, allergy, and immune response activation with larger sample size. They studied 845 patients from 2-18 years old with newly diagnosed TS and 3,378 controls matched by age and sex with the TS patients. The correlation between TS and allergic diseases (allergic rhinitis, atopic dermatitis, asthma, and allergic conjunctivitis) were evaluated. They concluded that TS is significantly correlated with the tested allergic diseases, and TS patients are more likely to have allergic rhinitis, atopic dermatitis, asthma, and allergic conjunctivitis (Chang *et al*., 2011).

These findings indicate that hidden layers in the etiology of TS may include a complex mechanism involving the dysregulation of the immune system.

Gene Expression Studies of Tourette Syndrome

RNA expression may be studied to investigate regulation of genes and biological pathways. Gene expression profiling of RNA is a powerful tool to identify biomarkers. According to the Central Dogma of molecular biology, RNA is the bridge between information in DNA sequence and protein, and the expression of RNA can be affected by genetic factors and environmental factors (Qi *et al*., 2017).

RNA microarrays, first developed in 1995, were designed to measure a set of pre-defined transcripts, which were. Microarrays allow the detection of thousands of transcriptions with high sensitivity, measurement accuracy, and low cost per gene (Lowe *et al*., 2017).

Lennington *et al*. analyzed the basal ganglia transcriptome from TS patients by RNA sequencing in the caudate nucleus and putamen in seven post-mortem TS subjects and 7 controls. This study found 309 down-regulated and 822 up-regulated genes. Next they employed co-expression analysis to help identify sets of genes that follow similar patterns in regulation. This analysis indicated that the top-scoring down-regulated co-expression module contained genes associated with neurotransmitter production and the function of GABAergic interneurons. This finding was consistent with the decreased numbers of cholinergic and GABAergic interneurons, demonstrated by immunohistochemistry in the caudate nucleus and putamen. The top upregulated co-expression module was enriched in genes that participate in immune response, consistent with a hypothesis of microglia activation in brain striatum from TS patients. This study revealed a cross-talk between immune and neural activity (Lennington *et al*., 2016).

TS is considered a neurodevelopmental disorder; thus, brain tissue is ideal to investigate the gene expression profile of TS. But brain tissue is not accessible *in vivo* for most neurologic studies in humans. Blood RNA is the most accessible sampling method for a longitudinal study and is often used as a proxy. Previous studies of the blood transcriptome of patients with TS and other neurological disorders revealed that specific profiles in the blood co-expressed in blood and brain related to neurological disorders (Frank *et al*., 2006, Du *et al*., 2006). Tang *et al*. compared blood gene expression between TS patients and controls. A total of six up-regulated genes and one down-regulated gene were identified by comparing the blood transcriptome between 16 TS patients and 8 controls. IMPA2 was down-regulated; this gene is involved in signal transduction (Ashburner *et al*.. 2000, The Gene Ontology Consortium, 2019, Bostein *et al*., 2000). Four immune-related genes *granzyme B, NKG2E, CD94* and *NK-p46* were up-regulated (Tang *et al*., 2005).

Mina *et al*. analyzed blood transcriptome data from 33 controls and 27 Huntington Disease (HD) and post-mortem brain transcriptome data from 44 HD patients and 36 controls, comparing co-expression genes from the blood dataset and the brain dataset. They found genes related to the immune response are shared between blood and brain (Mina *et al*., 2016).

Study Objective

Gene expression studies can help shed light on aspects of TS etiology and help to unravel dysregulations of the transcriptome. The aim of this study is to analyze genome-wide gene

expression profiles that are associated with tic exacerbation or tic onset, with the goal of identifying genes and gene pathways that possibly underlie the cause of these events. Based on prior studies that implicate infection as a triggering factor for TS, we hypothesize that our study may uncover genes or co-expression modules that correlate with dysregulation of the immune system.

METHODS AND MATERIALS

European Multicentre Tics in Children Study (EMTICS)

The European Multicentre Tics in Children Study is an international study including 17 clinical sites from across Europe. EMTICS aims to study the genetic and environmental factors that may lead to new tic onset and tic exacerbation (Georgitsi *et al*., 2016, Roessner & Hoekstra, 2013, Schrag *et al*., 2019). A central hypothesis tested by the EMTICS study is whether or not the onset and exacerbation of tics is associated with a specific genetic background or gene expression profile as well as infections of GAS.

In this study we leverage two unique cohorts sampled as part of the EMTICS: 1. the COURSE cohort, including longitudinal follow-up of patients and biological sample collection at points of tic exacerbation and subsequent remission, and 2. the ONSET cohort, including longitudinal follow-up of almost 350 fist degree relatives (mostly siblings) of patients with TS who did not have the disorder at baseline and biological sample collection upon symptom onset (Roessner & Hoekstra, 2013). The severity of tics was measured based on the Yale Global Tic Severity Scale, which gave tic scores based on the number, frequency, intensity, complexity, and interference of tics.

Participants

We used expression data sets generated by EMTICS and described in Padmanabhuni *et al*., (2017). For the original data collection, samples were separated into two cohorts: ONSET cohort and COURSE cohort as described above.

For the COURSE cohort, EMTICS collected blood samples from 715 TS patients, who were followed-up on a monthly basis for a 18-month period. Blood samples were drawn when the patients experienced tic exacerbation (TicExac1) and then subsequent tic remission (TicRemis1). Out of the 715 patients, 153 patients presented with tic exacerbation and then tic remission within the 18-month period. For the purposes of this protocol, tic remission was defined as a subsequent decrease of at least six points on the total tic severity score.

Out of the 153 TS patients, 26 patients also went through a second tic exacerbation event (TicExac2) and tic remission (TicRemis2) event within the follow-up period. For these patients blood samples were also taken at these additional tic exacerbation and tic remission time points.

The goal of the ONSET cohort was to study patients with new onset of tics. Subjects were selected as being high-risk for tic onset because their siblings were previously diagnosed with TS. A total of 260 tic-free children who were first-degree relatives of patients with a tic disorder, aged 3-10 years were recruited for follow-up and monitored for tic onset. Out of the 260 children, 48 developed tics within the three year follow-up period. Blood samples were collected at baseline visit (Baseline) upon recruitment and at the point of tic onset (ONSET). Table 1 shows the samples collected for downstream analysis.

Table 1: Samples collected for gene expression analysis through the EMTICS study.

COHORT	GROUP	CODE	TOTAL	TOTAL AFTER QC
ONSET	Baseline	Baseline	42	38
	Onset of symptoms	ONSET	42	42
COURSE	First Exacerbation	TicExac1	153	153
	First Remission	TicRemis1	153	138
	Second Exacerbation	TicExac2	26	26
	Second Remission	TicRemis2	26	24

Processing Method

Blood samples were drawn at different time points for all 42 patients in the ONSET cohort and 153 patients in COURSE cohort. RNA was extracted from the blood for transcriptomic analysis following standard protocols. Four samples in the baseline group (Baseline), 15 samples in the subsequent tic remission (TicRemis1) and two samples in the second tic remission (TicRemis2) failed at RNA quality control. The service provider Biolytix performed transcript analysis. RNA quality control was also performed by Biolytix on an Agilent Bioanalyzer, using the RNA 6000

Nano Kit before assaying for transcriptome expression. RNA samples were treated using the standard protocol of GeneChip® Human Transcriptome Array (HTA) 2.0 arrays. The Signal Space Transformation-Robust Multi-Chip Analysis algorithm from Affymetrix Power Tools software was used to reduce the background noise and increase the fold change by applying a GC content correction; this program also transformed the microdata signal to a similar signal space (Lockstone, 2011). For probe level quality control, means of perfect matched probes and background probes were used to detect unusually bright or dim probes and those probes were removed from the downstream process. For probe set level quality control, the true/false positives were checked for separation using the area under the curve (AUC) computed from perfectly matched and background probes. Any sample with AUC < 0.8 was removed from downstream analysis. Samples with higher than 2% missing probes were also removed from downstream analysis. The batch effect was corrected using ComBat from R package Surrogate Variable Analysis (SVA) (Johnson, Li, & Rabinovic, 2007). Filtered data was next processed for differential gene expression analysis.

Differential gene expression analysis. Differential gene expression analysis involves analyzing the normalized data to discover genes that show quantitative changes in expression levels between study groups satisfied by a statistical threshold (Kanz *et al*., 2005).

Three comparison groups were selected for investigation: 1. comparison within ONSET cohort between Baseline visit (Baseline) and Onset visit (ONSET), 2. comparison within COURSE cohort between First Exacerbation visit (TicExac1) and First Remission visit (TicRemis1), and 3. comparison between one or two Exacerbation visits (TicExac1 + TicExac2) and one or two Remission visits (TicRemis1 + TicRemis2). Probes were annotated using NetAffx transcript annotation of Affymetrix HTA 2.0 array. Fold change of each gene was computed based on the average signal strength:

$$\text{Fold Change} = 2^{[\text{Condition1 Gene Average Signal (log2)} - \text{Condition2 Gene Average Signal (log2)}]}$$

False discovery rate (FDR) adjusted p-value was calculated by Transcriptome Analysis Console from Thermo Scientific based on the Benjamini-Hochberg procedure. The differentially expressed genes, including positively and negatively regulated genes, were ranked and the top 10 up/down-regulated genes were checked for their biological function.

Weighted gene co-expression network analysis

Weighted gene co-expression network analysis was performed to investigate the co-expression patterns of sets of genes; this is complementary to a single-gene analysis approach. A weighted gene co-expression network is a systems biology method for describing the correlation of gene expression patterns across microarray samples(Langfelder *et al*., 2008). Nodes in each network stand for gene expression profiles, and connections between nodes were computed by the pairwise correlations between gene expression profiles. Network construction emphasizes high correlations between nodes and sacrifices those with low correlations by increasing the level of the β value, which is regarded as the soft threshold. Connection strength, also known as adjacency, between each pair of nodes is defined as:

$$\alpha_{ij} = |(1 + \mathrm{cor}(x_i, x_j))/2|^{\beta}$$

The connectivity for each node is defined as the sum of the connection strengths with all other nodes in the network (Carlson *et al.*, 2006):

$$k_i = \Sigma_{u \neq i} a_{ui}$$

The R package Weighted Correlation Network Analysis (WGCNA) was used, as it finds modules of highly correlated genes instead of looking for individual genes, constructing a cluster of highly correlated genes (Langfelder *et al*., 2008). WGCNA also can calculate the correlation between each module and external sample traits using eigengene, which is the first principle component of the corresponding co-expression network. In this study, the co-expression of annotated genes for each comparison group were computed using WGCNA (Zhang *et al*., 2005). After each module is constructed, modules contain > 100 genes with relatively high module-trait relationship. Significant p-values were selected for enrichment analysis.

Pathway analysis

Pathway analysis is widely used to give biological meaning to high-throughput genomic data and was used in the present study to provide a better biological understanding of the list of the genes from the differential expression analysis and co-expression analysis. The R package ClusterProfiler (Yu *et al*., 2012) was used to conduct the pathway analysis. ClusterProfiler calculates enrichment for GO terms based on the hypergeometric distribution. In this study, we

applied pathway analysis to the gene list from differential gene expression analysis and from significant co-expression modules from the WGCNA analysis. The differentially expressed genes in our three comparison groups were analyzed for functional profiles. Four different databases were used to extract gene-sets and pathways for analysis:

1. Cell Marker a comprehensive and accurate resource of cell marker genes for various cell types, which includes 13,605 cell markers of 467 cell types in 158 human tissues/subtissues (Zhang *et al*. 2019).

2. KEGG (Kyoto Encyclopedia of Genes and Genomes) database, which aims for understanding high-level function of the biological system from molecular-level function (Kanehisa et al., 2017).

3. Immunologic signatures collection (also called *ImmuneSigDB*) including gene sets that represent immune cell states and perturbations within the immune system, which were generated by manual curation of published studies (Godec *et al*., 2016).

4. Gene Ontology (GO) database which uses ontologies to support biologically meaningful annotation of genes and their products. The GO database including the following three GO ontologies: molecular function (MF), cellular component (CC), andr biological process (BP). In this study, BP was used to evaluate the enriched pathway of a given list of genes (Ashburner et al., 2000, The Gene Ontology Consortium, 2019).

RESULTS

Differential gene expression analysis

We investigated genome-wide gene expression profiles with the aim of identifing genes differentially expressed between comparison groups. In the comparison group of 42 tic-free siblings of patients (Baseline) vs 42 children with onset of symptoms (ONSET) (ONSET cohort), we found 378 up-regulated genes and 439 down-regulated genes with ANOVA P-value <0.05 upon onset of symptoms (Table 2 and Table3). However, none of the tests for up- or down-regulation were significant after multiple testing correction via false discovery rate (FDR) estimation. The top five up-regulated genes from baseline to exacerbation are *MAP2K4*, *EXPH5*, *SLC5A12*, *RC3H1*, and *CMTM2*. The top five down-regulated genes are *HIST1H3B*, *CD38*, *HIST1H3C*, *MIR1277*, and *MIR513B*.

Table 2: Top ten down-regulated genes between tic-free siblings of patients (Baseline) (42 samples) vs onset of symptoms (ONSET) (42 samples) (ONSET cohort) (gene ontology 0.05 ANNOVA p-value). None of the tests are statistically significant.

GENE	GENE NAME	BASELINE	ONSET	FOLD CHANGE	ANOVA P-value	FDR P-value	Expressed in both brain and blood?
HIST1H3B	*histone cluster 1*	6.5	5.83	1.59	0.0083	0.9999	No
CD38	*CD38 molecule*	10.05	9.6	1.36	0.0479	0.9999	Yes
HIST1H3C	*histone cluster 1*	5.22	4.83	1.31	0.0439	0.9999	No
MIR1277	*microRNA 1277*	5.75	5.4	1.28	0.0451	0.9999	No
DKFZP586I1420	*uncharacterized protein DKFZp586I1420*	8.57	8.25	1.25	0.0391	0.9999	Yes
MIR513B	*microRNA 513b*	4.23	3.92	1.24	0.0229	0.9999	No
TRNT1	*tRNA nucleotidyl transferase*	6.64	6.37	1.21	0.0436	0.9999	Yes
HIST1H2BN	*histone cluster 1*	6.86	6.59	1.2	0.0079	0.9999	No
HSPA1B	*heat shock 70kDa protein 1B*	7.1	6.84	1.2	0.0172	0.9999	Yes
HIST1H3J	*histone cluster 1*	7.16	6.91	1.2	0.0339	0.9999	No

Table 3: Top ten up-regulated genes between tic-free siblings of patients (Baseline)(42 samples) vs onset of symptoms (ONSET)(42 samples) (ONSET cohort). None of the tests are statistically significant.

GENE	GENE NAME	BASELINE	ONSET	FOLD CHANGE	ANOVA P-value	FDR P-value	Expressed in both brain and blood?
MAP2K4	*mitogen-activated protein kinase kinase 4*	10.4	10.7	-1.23	0.0032	0.9999	Yes
EXPH5	*exophilin 5*	4.58	4.85	-1.2	0.0083	0.9999	Yes
SLC5A12	*solute carrier family 5 (sodium/monocarboxylate cotransporter), member 12*	4.44	4.71	-1.2	0.0244	0.9999	Yes
RC3H1	*ring finger and CCCH-type domains 1*	11.28	11.54	-1.2	0.0496	0.9999	Yes
CMTM2	*CKLF-like MARVEL transmembrane domain containing 2*	6.66	6.92	-1.2	0.0262	0.9999	Yes
PRR9	*proline rich 9*	3.57	3.82	-1.19	0.0009	0.9999	Yes
LINC01402	*long intergenic non-protein coding RNA 1402*	4.39	4.64	-1.19	0.0177	0.9999	Yes
OR13D1	*olfactory receptor, family 13, subfamily D, member 1*	4.13	4.37	-1.18	0.0066	0.9999	Yes

Table 3 continued

OR51A4	*olfactory receptor, family 51, subfamily A, member 4*	3.97	4.19	-1.17	0.0183	0.9999	Yes
MIR3935	*microRNA 3935*	4.55	5.77	-1.17	0.0228	0.9999	Yes

In the comparison group of 153 first exacerbation (TicExac1) vs 138 first remission (TicRemis1) (COURSE cohort), we identified 245 genes up-regulated and 268 genes down-regulated with ANOVA P-value <0.05 from first exacerbation to first remission (Table 4 and Table5). None of the regulated genes were significant after multiple testing correction via false discovery rate (FDR) estimation. The top five up-regulated genes are *N4BP2L2-IT2*, *PLAGL1*, *DUSP1*, *DPEP2*, and *KRTAP9-2*. The top five down-regulated genes are *NCRUPAR*, *LOC101928530*, *SNORD115-35*, *RPS15AP10*, and *MIR548F5*.

Table 4: Top ten down-regulated genes in first exacerbation (TicExac1)(153 samples) vs first remission (TicRemis1)(138 samples) (COURSE cohort). None of the tests are statistically significant.

GENE	GENE NAME	BASELINE	ONSET	FOLD CHANGE	ANOVA P-value	FDR P-value	Expressed in both brain and blood?
NCRUPAR	*non-protein coding RNA*	5.19	5.04	1.11	0.0219	1	No
LOC101928530	*uncharacterized LOC101928530*	4.16	4.03	1.1	0.0016	1	No
SNORD115-35	*small nucleolar RNA*	4.02	3.89	1.09	0.0242	1	No
RPS15AP10	*ribosomal protein S15a pseudogene 10*	6.94	6.82	1.09	0.0359	1	Yes
MIR548F5	*microRNA 548f-5*	5.44	5.31	1.09	0.0331	1	No
LOC100131497	*uncharacterized LOC100131497*	3.87	3.75	1.09	0.0005	1	No
SNORD115-27	*small nucleolar RNA*	4.78	4.67	1.07	0.0397	1	No
CUBN	*cubilin (intrinsic factor-cobalamin receptor)*	6.71	6.61	1.07	0.0154	1	Yes
SNORD19	*small nucleolar RNA*	4.22	4.12	1.07	0.0438	1	No
RAD54L2	*RAD54-like 2 (S. cerevisiae)*	5.52	5.42	1.07	0.0095	1	Yes

Table 5: Top ten up-regulated genes in first exacerbation (TicExac1) (153 samples) vs first remission (TicRemis1) (138 samples) (COURSE cohort). None of the tests are statistically significant.

GENE	GENE NAME	BASELINE	ONSET	FOLD CHANGE	ANOVA P-value	FDR P-value	Expressed in both brain and blood?
N4BP2L2-IT2	*N4BPL2 intronic transcript 2*	11.17	11.33	-1.11	0.0419	1	Yes
PLAGL1	*pleiomorphic adenoma gene-like 1*	8.87	9.01	-1.11	0.0466	1	Yes
DUSP1	*dual specificity phosphatase 1*	10.23	10.37	-1.1	0.0155	1	Yes
DPEP2	*dipeptidase 2*	9.54	9.67	-1.09	0.031	1	Yes
KRTAP9-2	*keratin associated protein 9-2*	5.52	5.64	-1.09	0.0058	1	Yes
MIR938	*microRNA 938*	4.18	4.29	-1.09	0.0174	1	No
CEP295NL; TIMP2	*CEP295 N-terminal like; TIMP metallopeptidase inhibitor 2*	7.78	7.89	-1.08	0.015	1	Yes
SGK1	*serum/glucocorticoid regulated kinase 1*	7.71	7.82	-1.08	0.0117	1	Yes
LOC606724	*coronin*	8.87	8.98	-1.08	0.0113	1	Yes
IMPA2	*inositol(myo)-1(or 4)-monophosphatase 2*	8.85	8.95	-1.08	0.0406	1	Yes

In the comparison group of one or two exacerbations (TicExac1 + TicExac2) (181 samples) vs one or two remissions (TicRemis1 + TicRemis2) (162 samples), there are 189 genes up-regulated from two exacerbations to two remissions (Table 6 and Table7). None of the regulated genes were significant after multiple testing correction via the false discovery rate (FDR) estimation. There are 734 genes down-regulated from two exacerbations to two remissions with ANOVA P-value <0.05. The top up-regulated genes are *RNF217*, *KRTAP9-2*, *ABI1*, *SNRPEP2*, and *LY6G5B*. The top down-regulated genes are *RNU6ATAC*, *RNY4P22*, *MIR190A*, *ZFP57*, and *GJC2*.

Table 6: The top ten down-regulated genes in two exacerbations (TicExac1 + TicExac2)(181 samples) vs two remissions (TicRemis1 + TicRemis2)(162 samples) (COURSE cohort). None of the tests are statistically significant.

GENE	GENE NAME	BASELINE	ONSET	FOLD CHANGE	ANOVA P-value	FDR P-value	Expressed in both brain and blood?
RNU6ATAC	*RNA, U6atac small nuclear (U12-dependent splicing)*	10.25	9.93	1.25	0.0475	0.9997	Yes
RNY4P22	*RNA, Ro-associated Y4 pseudogene 22*	7.00	6.87	1.09	0.0175	0.9997	No
MIR190A	*microRNA 190a*	4.70	4.58	1.09	0.0344	0.9997	No
ZFP57	*ZFP57 zinc finger protein*	6.39	6.28	1.08	0.0432	0.9997	Yes
GJC2	*gap junction protein gamma 2*	6.63	6.52	1.08	0.0204	0.9997	Yes
KLRF2	*killer cell lectin-like receptor subfamily F, member 2*	3.78	3.68	1.07	0.0492	0.9997	Yes
MIR651	*microRNA 651*	4.25	4.15	1.07	0.0175	0.9997	No
MAP7	*microtubule-associated protein 7*	4.07	3.98	1.07	0.0097	0.9997	Yes
OR5H6	*olfactory receptor, family 5, subfamily H, member 6 (gene/pseudogene)*	4.10	4.01	1.07	0.0229	0.9997	Yes
MIR606	*microRNA 606*	5.67	5.58	1.07	0.0442	0.9997	No

Table 7: The top ten up-regulated genes in two exacerbations (TicExac1 + TicExac2) (181 samples) vs two remissions (TicRemis1 + TicRemis2) (162 samples) (COURSE cohort). None of the tests are statistically significant.

GENE	GENE NAME	BASELINE	ONSET	FOLD CHANGE	ANOVA P-value	FDR P-value	Expressed in both brain and blood?
RNF217	*ring finger protein 217*	5.17	5.31	-1.11	0.029	0.9997	Yes
LOC101928324	*uncharacterized LOC101928324*	6.78	6.9	-1.09	0.0126	0.9997	No
LOC101927069	*uncharacterized LOC101927069*	6.10	6.21	-1.08	0.0296	0.9997	No
KRTAP9-2	*keratin associated protein 9-2*	5.51	5.63	-1.08	0.0294	0.9997	Yes
ABI1	*abl-interactor 1*	6.37	6.48	-1.08	0.0205	0.9997	Yes
SNRPEP2	*small nuclear ribonucleoprotein polypeptide E pseudogene 2*	6.00	6.11	-1.08	0.0398	0.9997	Yes
LY6G5B	*lymphocyte antigen 6 complex, locus G5B*	8.85	8.95	-1.07	0.0169	0.9997	No
ZBED6	*zinc finger, BED-type containing 6*	4.76	4.87	-1.07	0.0346	0.9997	Yes
DSTNP4	*destrin (actin depolymerizing factor) pseudogene 4*	3.98	4.07	-1.06	0.0050	0.9997	Yes
ATP5L2	*ATP synthase, H+ transporting, mitochondrial Fo complex subunit G2*	6.36	6.43	-1.06	0.0327	0.9997	No

The pathway analysis of each comparison group mapped the up- and down-regulated genes to a list of biological pathways (Table 8).

For the comparison group Baseline vs. ONSET up-regulated genes are enriched in an ImmuneSigDBe pathway, including a gene set of genes down-regulated in comparison of "unstimulated peripheral blood mononuclear cells (PBMC) three days after stimulation with YF17D vaccine versus PBMC 21 days after the stimulation". In addition, according to the GO database, up-regulated genes are enriched in the "detection of chemical stimulus involved in sensory perception pathway", which represents a series of events in which a chemical stimulus is received and converted into a molecular signal as part of sensory perception. According to the Cell Markers database, down-regulated genes are enriched in the marker genes of neural progenitor cell in embryonic prefrontal cortex pathway. Also, according to ImmuneSigDB, down-regulated genes are enriched in genes up-regulated in comparison to "wild type CD8 effector T cells at day six versus those from mice deficient for TRAF6 at day ten". In addition, down-regulated genes are also enriched in the spliceosome genes in the KEGG database and enriched in the RNA splicing pathway according to the GO database.

For the comparison group TicExac1 vs. TicRemis1, no significant pathways were identified in up-regulated genes; down-regulated genes are enriched in the pathway neural progenitor cell in embryonic prefrontal cortex. Also, enriched in the ImmuneSigDB gene set that was reported down-regulated in "Macrophages differentiated for 5 days in the presence of IL4 versus IL4 and dexamethasone". In addition, down-regulated genes were also enriched in the "sister chromatid segregation" pathway, according to the GO database.

For the comparison group TicExac1 + TicExac2 vs TicRemis1 + TicRemis2, no significant pathways were enriched in up-regulated genes; down-regulated genes are reported enriched in "neural progenitor cell in embryonic prefrontal cortex" according to the Cell Marker database.

Table 8: Biological classification for up-regulated or down-regulated genes from each comparison group. The Biological Classification column shows the most significant term from each database. The "Number of genes enriched" column indicates the number of differentially expressed genes enriched in the corresponding pathway, the "Total number of genes in the pathway" column indicates the total number of genes included in the pathway. The following sources were used for the enrichment analysis: Cell Markers, ImmuneSigDB, KEGG, Gene Ontology (Biological Process). Corrected p-value: Bonferroni corrected p-values.

Group Name/ down-regulated or up-regulated	Biological Classification	Database	Corrected p-values	Number of genes enriched	Total number of genes in the pathway
Baseline vs. ONSET (up-regulated genes)	NA	Cell Markers	NA	NA	NA
	GSE13485 Genes down-regulated in comparison to unstimulated peripheral blood mononuclear cells (PBMC) three days after stimulation with YF17D vaccine versus PBMC 21 days after the stimulation.	ImmuneSigDB	1.482e-4	13	184
	NA	KEGG	NA	NA	NA
	GO:0050907 Detection of chemical stimulus involved in sensory perception	GO (Biological Process)	1.417e-2	19	477
Baseline vs. ONSET (down-regulated genes)	Embryonic prefrontal cortex, Normal, Neural progenitor cell	Cell Markers	1.542e-10	23	166
	GSE15750: Genes up-regulated in comparison to wild type CD8 effector T cells at day six versus those from mice defficient for TRAF6 at day 10.	ImmuneSigDB	2.049e-17	31	200
	hsa03040: Spliceosome	KEGG	7.197e-3	12	151
	GO:0008380: RNA splicing	GO (Biological Process)	1.164e-05	29	469
TicExac1 vs. TicRemis1 (up-regulated genes	NA	Cell Markers	NA	NA	NA
	NA	ImmuneSigDB	NA	NA	NA
	NA	KEGG	NA	NA	NA

Table 8 continued

	NA	**GO (Biological Process)**	**NA**	**NA**	**NA**
TicExac1 vs. TicRemis1 (down-regulated genes)	Embryonic prefrontal cortex, Normal, Neural progenitor cell	Cell Markers	2.293e-2	7	166
	GSE7568: Genes up-regulated in macrophages differentiated for 5 days in the presence of: IL4 versus IL4 and dexamethasone.	ImmuneSigDB	9.322e-3	10	180
	NA	KEGG	NA	NA	NA
	GO:0000819: sister chromatid segregation	GO (Biological Process)	5.891e-3	10/151	189
TicExac1 + TicExac2 vs TicRemis1 + TicRemis2 (up-regulated genes)	NA	Cell Markers	NA	NA	NA
	NA	ImmuneSigDB	NA	NA	NA
	NA	KEGG	NA	NA	NA
	NA	GO (Biological Process)	NA	NA	NA
TicExac1 + TicExac2 vs TicRemis1 + TicRemis2 (down-regulated genes)	Embryonic prefrontal cortex, Normal, Neural progenitor cell	Cell Markers	2.293e-2	17	344
	NA	ImmuneSigDB	NA	NA	NA
	NA	KEGG	NA	NA	NA
	NA	GO (Biological Process)	NA	NA	NA

NA: No significant association.

Weighted Gene Co-expression Network Analysis

Weighted gene co-expression network analysis was applied to investigate groups of genes that were regulated jointly among samples between different TS stages. Microarray data can be more comprehensively explored by assessing the pair-wise correlations between gene expression profiles in groups under comparison. Here, 17,300 annotated genes were analyzed using WGCNA as described in the methods. Within the ONSET cohort, the comparison between Baseline (Baseline) and Onset visit (ONSET), seven co-expression modules were clustered.

Table 9: Co-expression modules in the ONSET cohort, from the comparisons between Baseline (Baseline) and onset visit (ONSET) and the number of genes in each module.

Module Annotation	Number of genes	Module-trait Correlation	Corrected P-value
Blue	344	0.053	0.6
Brown	271	-0.12	0.3
Green	15819	-0.15	0.2
Pink	269	0.28	0.01*
Red	32	0.17	0.1
Turquoise	30	-0.092	0.4
Yellow	42	0.26	0.02*

Table 9 shows the module-trait relationship for the comparison between Baseline vs. ONSET. The yellow module and pink module are the two significant modules with p-value < 0.05. The pink module has a corresponding correlation of 0.28 (p-value = 0.01) and the yellow module has the corresponding correlation of 0.26 (p-value = 0.02) ; the pink module contains 269 genes, while there are 42 genes within the yellow module. Consequently, we decided to focus on the pink and yellow module for pursuing enrichment analysis.

In the COURSE cohort comparison between First Exacerbation visit (TicExac1) and First Remission visit (TicRemis1), according to Table 10, the module-trait correlation shows two significant modules: light green module (correlation = 0.12, p-value = 0.05) and purple module (correlation = 0.13, p-value = 0.03).

Table 10: Co-expression modules from the COURSE cohort comparison between First Exacerbation visit (TicExac1) and First Remission visit (TicRemis1) and the number of genes in each module.

Module Colors	Number of genes	Module-trait Correlation	P-value
Black	14021	0.017	0.8
Blue	490	-0.025	0.7
Brown	306	-0.02	0.7
Cyan	133	-0.00088	0.1
Green	244	-0.082	0.2
Greenyellow	136	-0.014	0.8
Grey60	87	-0.079	0.2
LightCyan	85	-0.07	0.2
LightGreen	543	0.12	0.05*
Lightyellow	286	-0.014	0.8
Matenga	59	-0.032	0.6
Midnightblue	244	-0.068	0.2
Pink	48	0.063	0.3
Purple	31	0.13	0.03*
Red	98	-0.015	0.8
Salmon	148	0.03	0.6
Tan	62	-0.048	0.4
Turquoise	222	0.0097	0.9
Yellow	77	0.033	0.6

In the comparison between two Exacerbation visits (TicExac1 + TicExac2) and two Remission visits (TicRemis1 + TicRemis2), A total of nineteen modules were constructed by WGCNA (Table 11). The blue module was the only significant co-expression module (p-value = 0.01) with the module-trait correlation of 0.13. As it is shown in Table 11, there were 456 genes in the blue co-expression module.

Table 11: Co-expression modules in the comparison between two Exacerbation visits (TicExac1 + TicExac2) and two Remission visits (TicRemis1 + TicRemis2) and the number of genes in each module.

Module Colors	Number of genes	Module-trait Correlation	FDR Corrected P-value
Black	13782	-0.0051	0.9
Blue	456	0.13	0.01*
Brown	349	-0.057	0.3
Cyan	84	-0.016	0.8
Green	295	0.032	0.6
Greenyellow	108	-0.031	0.6
Grey60	77	-0.011	0.8
LightCyan	66	0.004	0.9
LightGreen	952	-0.031	0.6
Lightyellow	345	0.00074	1
Matenga	47	-0.0095	0.9
Midnightblue	236	-0.032	0.6
Pink	45	-0.032	0.6
Purple	31	0.049	0.4
Red	81	-0.03	0.6
Salmon	111	0.061	0.3
Tan	49	-0.0095	0.9
Turquoise	133	0.054	0.3
Yellow	53	0.052	0.3

Pathway analysis was also performed to test genes in co-expression modules to reveal the potential biological meanings of each co-expression network.

Table 12: Biological classification for the modules significantly enriched in up-regulated or down-regulated genes for each comparison group based on the over representative test. For each of the co-expression modules, the Biological Classification column shows the most significant term from each database. The "Number of genes enriched" column indicates the number of differentially expressed genes enriched in the corresponding pathway, the "Total number of genes in the pathway" column indicates the total number of genes included in the pathway.The following sources were used for the enrichment analysis: Cell Markers, ImmuneSigDB, KEGG, and Gene Ontology (Biological Process). Corrected p-value: Bonferroni corrected p-values.

Group Name	Module Name	Biological Classification	Database	Corrected p-values	Number of genes enriched	Total number of genes in the pathway
Baseline vs. ONSET	Pink	Embryo, Normal, Morula cell (Blastomere)	Cell Markers	9.083e-6	20	219
		GSE2405: Genes up-regulated in polymorphonuclear leukocytes (24h): control versus infection by A. phagocytophilum.	ImmuneSigDB	1.953e-47	48	200
		hsa05016: Huntington disease	KEGG	8.895e-03	13	306
		GO:0006613 : cotranslational protein targeting to membrane	GO (Biological Process)	2.155e-64	51	109
	Yellow	NA	Cell Markers	NA	NA	NA
		GSE40225: Genes up-regulated in pancreatic CD8 T cells from mice with: type 1 diabetes mellitus versus healthy controls.	ImmuneSigDB	9.922e-3	5	199
		NA	KEGG	NA	NA	NA
		GO:0015074 DNA integration	GO (Biological Process)	9.922e-3	2	16
TicExac1 vs. TicRemis1	Light green	Embryo, Normal, Morula cell (Blastomere)	Cell Markers	1.481e-4	15	100
		GSE2405: Genes up-regulated in polymorphonuclear leukocytes (24h): control versus infection by *A. phagocytophilum.*	ImmuneSigDB	4.551e-52	66	200
		hsa05016: Huntington disease	KEGG	8.340e-07	32	306
		GO:0006613: cotranslational protein targeting to membrane	GO (Biological Process)	6.908e-67	62	109
	Purple	Blood, Normal, CD1C+_B dendritic cell	Cell Markers	1.312e-2	3	55

Table 12 continued

		GSE22886: Genes down-regulated in comparison of neutrophils versusl monocytes.	**ImmuneSigDB**	**1.718e-8**	**9**	**200**
		NA	KEGG	NA	NA	NA
		GO:0002283: neutrophil activation involved in immune response	GO (Biological Process)	3.924e-2	6	485
TicExac1 + TicExac2 vs TicRemis1 + TicRemis2	Blue	Embryo, Normal, Morula cell (Blastomere)	Cell Markers	1.271e-4	23	219
		GSE2405: Genes down-regulated in polymorphonuclear leukocytes (9h): *S. aureus* infection versus control.	ImmuneSigDB	8.814e-53	63	200
		hsa05016: Huntington disease	KEGG	1.681e-6	29	306
		GO:0006613: cotranslational protein targeting to membrane	GO (Biological Process)	5.349e-68	60	109

For the pink module in the comparison group of Baseline vs. ONSET (Table 12), genes are most significantly enriched in "embryo morula cell" (Cell Marker database). It is also enriched in genes that are downregulated due to the "infection by *A. phagocytophilum* in polymorphonuclear leukocytes" (ImmuneSigDB database). Genes in the pink module are also enriched in the "Huntington disease" pathway (KEGG database) and "cotranslational protein targeting to membrane" pathway (GO database).

For the yellow module in the comparison group Baseline vs. ONSET, genes are enriched in "up-regulated in pancreatic CD8 T cells from mice with type 1 diabetes mellitus" (ImmuneSigDB database) and the "DNA integration" pathway (GO database).

Genes in the light green module are significantly enriched in the marker genes of "embryo morula cell" (Cell Marker database) and enriched in the "down-regulated genes in polymorphonuclear leukocytes after *S. aureus* infection" pathway (ImmuneSigDB database). Genes in the light green module were also enriched in the "Huntington disease" pathway (KEGG database) and "cotranslational protein targeting to membrane" (GO database). Genes in the purple module (Table 10) are enriched in the marker genes of "blood CD1C+ B dendritic cell" (Cell Marker database), and are also enriched in genes down-regulated in comparison to "naive B cells versus unstimulated

neutrophils" (WHICH DATABASE?)and the pathway "neutrophil activation involved in immune response" (GO database). For the blue module, genes are significantly enriched in "embryo morula cell" in Cell Marker database, "genes down-regulated in polymorphonuclear leukocytes infection by *A. phagocytophilum*" (ImmuneSigDB), "Huntington disease" pathway (KEGG database) and "cotranslational protein targeting to membrane" (GO database).

DISCUSSION

TS is a neurological disorder with a complex cause that has not been elucidated to date. Although brain tissue would be the ideal choice for investigating the gene expression profile of a neurological disorder, blood RNA which is the most accessible sampling method for a longitudinal study is often used as a proxy. Previous studies have shown levels of DNA methylation on CpG islands are highly correlated (r=0.90) between brain and blood, and transcriptomic studies revealed between 35% to 80% of transcripts being expressed in both blood and brain. Also, the correlation of expression level across different tissues ranged from 0.25 to 0.64, indicating whole blood can be used as a proxy to study brain transcriptome (Tylee *et al*., 2013).

Our study use whole blood RNA expression analysis and focused on comparing data from TS patients at time points of symptom exacerbation and remission as well as before and after patients are diagnosed with TS. None of the genes that we found to be up-regulated or down-regulated upon symptom onset or tic exacerbation versus remission achieved statistical significance after correction for multiple testing (FDR p-value <0.05). However, we found that within all three comparison groups, several of the top up- or down-regulated genes play a role in immune response and some genes have been previously reported as associated with neurological functions and/or are expressed in the brain (Table 2, Table3, Table4). For example, among top regulated genes from comparison group Baseline vs. ONSET, *MAP2K4* encodes a member of the mitogen-activated protein kinase (MAPK) family. *MAP2K4* also plays a role in the toll-like receptor signaling pathway, which is critical in innate immune responses. Up-regulation of *MAP2K4* has also the regulation of *MAP2K4* has been detected in ASD patients; together, this demonstrates that *MAP2K4* is involved in both the immune system and the neuron function (Nazeen *et al*., 2016, Barton *et al*., 2009).

CD38 encodes a type II transmembrane glycoprotein that is widely expressed in both blood and brain tissues. The *CD38* gene has also been associated with ASD risk. According to Lerer *et al*., a high reduction in *CD38* expression in ASD subjects compared to their "unaffected" parents was found, indicating the important role of CD38 in neurodevelopmental disorders (Glaría *et al*., 2020, Lerer *et al*., 2010).

Interestingly, *EXPH5* is actively transcribed in both blood and brain, but EXPH5 protein is only detected in the brain. A population-based study also previously identified *EXPH5* as an ASD-associated gene. They suggested that the *EXPH5* gene may influence neural correlates associated with the cerebellum (Lin *et al*., 2013). In this case, regulation of mRNA of *EXPH5* in blood can also indicate the level of EXPH5 in brain.

Among the top genes differing between First Exacerbation visit (TicExac1) and First Remission visit (TicRemis1), *dual specificity phosphatase 1* (*DUSP1*) is expressed in both blood and brain. This protein is enriched in neutrophils and plays a role in immune response. Up-regulation of *DUSP1* is involved in the negative feedback regulation of the MAPK signaling pathway (Uhen *et al*., 2017, Abraham *et al*., 2006, Lang *et al*., 2019, Arthur *et al*., 2013).

For the comparison group between two Exacerbation visits (TicExac1 + TicExac2) and two Remission visits (TicRemis1 + TicRemis2), *RNF217* (*Ring Finger Protein 217*) is the top up-regulated gene. *RNF217* is constitutively expressed in all studied tissues with different relative expression levels; *RNF217* protein participates in regulating immune signaling as E3 ubiquitin ligases (Stelzer *et al*., 2016, Luo *et al*., 2019). *LY6G5B* (*Lymphocyte antigen 6 family member G5B*) is expressed in both brain and white blood cells. Secreted proteins *Ly6G5b* and *Ly6G5c* form a subcluster of MHC, which indicates an immune-related role of *LY6G5B* (Mallya *et al*., 2006).

The top genes we identified as up-regulated or down-regulated upon symptom exacerbation or symptom onset matched our hypothesis that the regulation of the neural system and the immune system are both associated with TS. Although the fold change of each regulated gene was very mild, it is the first time that changes in the regulation of the immune system have been implicated at different time points in TS patients. Pathway analysis of the top genes reveals an interesting trend of regulation of neuronal development and immune pathways participating in TS etiology. Down-regulated genes in Baseline vs. ONSET and up- or down-regulated genes in TicExac1 vs. TicRemis are enriched in the “embryonic prefrontal cortex” and “neural progenitor cells” (Cell Marker database), indicating the involvement of the neurodevelopmental processes. Gene sets from ImmuneSigDB indicate a role for gene regulation before and after the stimulation of immune cells, thus the enriched genes show the likely regulation of blood immune cell activities in TS patients, specifically at time points of symptom exacerbation and remission as well as before and after patients are diagnosed with TS.

Pathway analysis of co-expression modules also revealed the crosstalk of the immune system and the nervous system. Interestingly, the pink/light green/blue co-expression modules are enriched in the "Huntington disease pathway" (KEGG database), indicating the possibility of shared gene regulations in both TS and Huntington disease, between different stages of TS. Finally, the enrichment of ImmuneSigDB pathways for all significant co-expression modules indicates the role of the regulation of blood immune cells between different stages of TS.

Limitations of This Study and Suggestions for Future Studies

This transcriptome study that was performed on blood instead of brain tissue, so the findings may potentially be affected by non-related environmental factors. A recent study of ADHD patients suggested that blood transcriptome analysis may be useful for identifying environmental factors, causative genes, and pathological changes. (Lorenzo *et al*., 2018) Blood and brain tissues may also have different expression profiles for some genes.

An additional limitation is the fact that this study did not involve a healthy control group. TS severity is associated with patient age, thus, growth and development can play a role in TS progress. Therefore, involving a healthy control group would provide additional insights and should be included in future work.

The HTA 2.0 microarray is not designed to detect microRNA (miRNA) directly: signals of miRNA are estimated based on the signals of pre-miRNA. As an important component of epigenetics, direct measurement of miRNA expression such as transcriptome sequencing would be helpful for further understanding TS etiology in relation to this component. Transcriptome sequencing would likely give us a more complete picture of the expression network involved in onset or exacerbation of TS. Lastly, RNA sequencing could yield higher resolution and provide a better coverage in the analysis since the microarray gene coverage is limited.

CONCLUSION

Here we present a genome-wide gene expression study comparing data from TS patients at time points of symptom exacerbation and remission as well as before and after patients are diagnosed with TS.

The results of the gene differential analysis suggest an impressive cross-talk between immune pathways and the nervous system. Although no single-gene comparisons retained significance after correction for multiple testing, the top differentially expressed genes upon tic onset and symptom exacerbation of TS are found to play important roles in the function of the immune system and in the nervous system. In addition, the differentially expressed genes are enriched in immune-related pathways and pathways related to neurodevelopment. Co-expression analysis revealed sets of genes that are regulated jointly at symptom fluctuation and onset. Pathway analysis of co-expression modules also revealed the crosstalk of the immune system and the nervous system.

Although the fold change of each regulated gene was very mild, it is the first time that changes in the regulation of the immune system have been implicated at different time points in TS patients. Follow-up studies involving a healthy control group would provide additional insights and should be included in future work, as well as RNA sequencing studies with an increased number of samples.

REFERENCES

Abraham, S. M., & Clark, A. R. (2006). Dual-specificity phosphatase 1: A critical regulator of innate immune responses. Biochemical Society Transactions, 34(Pt 6), 1018-1023. doi:10.1042/BST0341018

Adamczyk, A., Gause, C. D., Sattler, R., Vidensky, S., Rothstein, J. D., Singer, H., & Wang, T. (2011). Genetic and functional studies of a missense variant in a glutamate transporter, SLC1A3, in tourette syndrome. Psychiatric Genetics, 21(2), 90-97.

Ahmed, A. U., Beech, P. L., Lay, S. T., Gilson, P. R., & Fisher, P. R. (2006). Import-associated translational inhibition: Novel in vivo evidence for cotranslational protein import into dictyostelium discoideum mitochondria. Eukaryotic Cell, 5(8), 1314-1327. doi:10.1128/EC.00386-05

Anderson, G. M., Pollak, E. S., Chatterjee, D., Leckman, J. F., Riddle, M. A., & Cohen, D. J. (1992). Brain monoamines and amino acids in gilles de la tourette's syndrome: A preliminary study of subcortical regions. Archives of General Psychiatry, 49(7), 584-586.

Arthur, J. S. C., & Ley, S. C. (2013). Mitogen-activated protein kinases in innate immunity. Nature Reviews Immunology, 13(9), 679-692.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., . . . Eppig, J. T. (2000). Gene ontology: Tool for the unification of biology. Nature Genetics, 25(1), 25-29.

Barton, G. M., & Kagan, J. C. (2009). A cell biological view of toll-like receptor function: Regulation through compartmentalization. Nature Reviews Immunology, 9(8), 535-542.

Bombaci, M., Grifantini, R., Mora, M., Reguzzi, V., Petracca, R., Meoni, E., . . . Bensi, G. (2009). Protein array profiling of tic patient sera reveals a broad range and enhanced immune response against group A streptococcus antigens. PloS One, 4(7), e6332. doi:10.1371/journal.pone.0006332 [doi]

Botstein, D., Cherry, J. M., Ashburner, M., Ball, C. A., Blake, J. A., Butler, H., . . . Eppig, J. T. (2000). Gene ontology: Tool for the unification of biology. Nat Genet, 25(1), 25-29.

Browne, H. A., Modabbernia, A., Buxbaum, J. D., Hansen, S. N., Schendel, D. E., Parner, E. T., . . . Grice, D. E. (2016). Prenatal maternal smoking and increased risk for tourette syndrome and chronic tic disorders. Journal of the American Academy of Child and Adolescent Psychiatry, 55(9), 784-791. doi:10.1016/j.jaac.2016.06.010 [doi]

Carlson, M. R., Zhang, B., Fang, Z., Mischel, P. S., Horvath, S., & Nelson, S. F. (2006). Gene connectivity, function, and sequence conservation: Predictions from modular yeast co-expression networks. BMC Genomics, 7, 40-40. doi:1471-2164-7-40 [pii]

Chang, Y. T., Li, Y. F., Muo, C. H., Chen, S. C., Chin, Z. N., Kuo, H. T., . . . Chou, I. C. (2011). Correlation of tourette syndrome and allergic disease: Nationwide population-based case-control study. Journal of Developmental and Behavioral Pediatrics : JDBP, 32(2), 98-102. doi:10.1097/DBP.0b013e318208f561 [doi]

Chang, Y., Li, Y., Muo, C., Chen, S., Chin, Z., Kuo, H., . . . Chou, I. (2011). Correlation of tourette syndrome and allergic disease: Nationwide population-based case-control study. Journal of Developmental & Behavioral Pediatrics, 32(2), 98-102.

Comings, D. E., Wu, S., Chiu, C., Ring, R. H., Gade, R., Ahn, C., . . . Muhleman, D. (1996). Polygenic inheritance of tourette syndrome, stuttering, attention deficit hyperactivity, conduct, and oppositional defiant disorder: The additive and subtractive effect of the three dopaminergic genes--DRD2, D beta H, and DAT1. American Journal of Medical Genetics, 67(3), 264-288. doi:10.1002/(SICI)1096-8628(19960531)67:3<264::AID-AJMG4>3.0.CO;2-N [pii]

Comings, D. E., Comings, B. G., Muhleman, D., Dietz, G., Shahbahrami, B., Tast, D., . . . Flanagan, S. D. (1991). The dopamine D2 receptor locus as a modifying gene in neuropsychiatric disorders. JAMA : The Journal of the American Medical Association, 266(13), 1793-1800. doi:10.1001/jama.1991.03470130073032

Croft, D., O'Kelly, G., Wu, G., Haw, R., Gillespie, M., Matthews, L., . . . Stein, L. (2011). Reactome: A database of reactions, pathways and biological processes. Nucleic Acids Research, 39, D691-D697. doi:10.1093/nar/gkq1018

Cruz, C., Camarena, B., King, N., Páez, F., Sidenberg, D., de la Fuente, Juan Ramón, & Nicolini, H. (1997). Increased prevalence of the seven-repeat variant of the dopamine D4 receptor gene in patients with obsessive-compulsive disorder with tics doi:https://doi.org/10.1016/S0304-3940(97)00523-5

Dehning, S., Muller, N., Matz, J., Bender, A., Kerle, I., Benninghoff, J., . . . Zill, P. (2010). A genetic variant of HTR2C may play a role in the manifestation of tourette syndrome. Psychiatric Genetics, 20(1), 35-38. doi:10.1097/YPG.0b013e32833511ce [doi]

DeMaria, S., & Ngai, J. (2010). The cell biology of smell. The Journal of Cell Biology, 191(3), 443-452. doi:10.1083/jcb.201008163

Diaz-Anzaldua, A., Joober, R., Riviere, J. B., Dion, Y., Lesperance, P., Richer, F., . . . Montreal Tourette Syndrome Study Group. (2004). Tourette syndrome and dopaminergic genes: A family-based association study in the french canadian founder population. Molecular Psychiatry, 9(3), 272-277. doi:10.1038/sj.mp.4001411 [doi]

Diaz-Anzaldua, A., Joober, R., Riviere, J. B., Dion, Y., Lesperance, P., Richer, F., . . . Rouleau, G. A. (2004). Tourette syndrome and dopaminergic genes: A family-based association study in the french canadian founder population. Molecular Psychiatry, 9(3), 272-277.

Eapen, V., Cavanna, A. E., & Robertson, M. M. (2016). Comorbidities, social impact, and quality of life in tourette syndrome. Frontiers in Psychiatry, 7, 97. doi:10.3389/fpsyt.2016.00097 [doi]

Ercan-Sencicek, A. G., Stillman, A. A., Ghosh, A. K., Bilguvar, K., O'Roak, B. J., Mason, C. E., . . . Pauls, D. L. (2010). L-histidine decarboxylase and tourette's syndrome. New England Journal of Medicine, 362(20), 1901-1908.

Felling, R. J., & Singer, H. S. (2011). Neurobiology of tourette syndrome: Current status and need for further investigation. Journal of Neuroscience, 31(35), 12387-12395.

Feltrin, A. S., Tahira, A. C., Simões, S. N., Brentani, H., & Martins,David Corrêa,,Jr. (2019). Assessment of complementarity of WGCNA and NERI results for identification of modules associated to schizophrenia spectrum disorders. PloS One, 14(1), e0210431. doi:10.1371/journal.pone.0210431

Gade, R., Muhleman, D., Blake, H., MacMurray, J., Johnson, P., Verde, R., . . . Comings, D. E. (1998). Correlation of length of VNTR alleles at the X-linked MAOA gene and phenotypic effect in tourette syndrome and drug abuse. Molecular Psychiatry, 3(1), 50-60. doi:10.1038/sj.mp.4000326

García-Campos, M. A., Espinal-Enríquez, J., & Hernández-Lemus, E. (2015). Pathway analysis: State of the art. Frontiers in Physiology, 6, 383.

Gene Ontology Consortium. (2019). The gene ontology resource: 20 years and still GOing strong. Nucleic Acids Research, 47(D1), D330-D338.

Gene, O. C. (2015). Gene ontology consortium: Going forward. Nucleic Acids Research, 43, D1049-D1056. doi:10.1093/nar/gku1179

Georgitsi, M., Willsey, A. J., Mathews, C. A., State, M., Scharf, J. M., & Paschou, P. (2016). The genetic etiology of tourette syndrome: Large-scale collaborative efforts on the precipice of discovery. Frontiers in Neuroscience, 10, 351. doi:10.3389/fnins.2016.00351

Glaría, E., & Valledor, A. F. (2020). Roles of CD38 in the immune response to infection. Cells, 9(1), 228.

Godec, J., Tan, Y., Liberzon, A., Tamayo, P., Bhattacharya, S., Butte, A. J., . . . Haining, W. N. (2016). Compendium of immune signatures identifies conserved and species-specific biology in response to inflammation. Immunity, 44(1), 194-206.

Herrera, R., Morris, M., Rosbe, K., Feng, Z., Weinberg, A., & Tugizov, S. (2016). Human beta-defensins 2 and -3 cointernalize with human immunodeficiency virus via heparan sulfate proteoglycans and reduce infectivity of intracellular virions in tonsil epithelial cells. Virology, 487, 172-187. doi:10.1016/j.virol.2015.09.025

Ho, C. S., Shen, E. Y., Shyur, S. D., & Chiu, N. C. (1999). Association of allergy with tourette's syndrome. Journal of the Formosan Medical Association= Taiwan Yi Zhi, 98(7), 492.

Hoekstra, P. J., Dietrich, A., Edwards, M. J., Elamin, I., & Martino, D. (2013). Environmental factors in tourette syndrome. Neuroscience and Biobehavioral Reviews, 37(6), 1040-1049. doi:10.1016/j.neubiorev.2012.10.010 [doi]

Huang, A. Y., Yu, D., Davis, L. K., Sul, J. H., Tsetsos, F., Ramensky, V., . . . Chen, J. A. (2017). Rare copy number variants in NRXN1 and CNTN6 increase risk for tourette syndrome. Neuron, 94(6), 1101-1111. e7.

Jetani, H., Garcia-Cadenas, I., Nerreter, T., Thomas, S., Rydzek, J., Meijide, J. B., . . . Einsele, H. (2018). CAR T-cells targeting FLT3 have potent activity against FLT3− ITD AML and act synergistically with the FLT3-inhibitor crenolanib. Leukemia, 32(5), 1168-1179.

Ji, X., Wu, M., & Zhang, X. (2016). Dopamine level in the stratium exhibits circadian rhythms in the rat model of tourette syndrome. Int J Clin Exp Med, 9(6), 9015-9020.

Johnson, W. E., Li, C., & Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical bayes methods. Biostatistics, 8(1), 118-127.

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., & Morishima, K. (2017). KEGG: New perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Research, 45, D353-D361. doi:10.1093/nar/gkw1092

Kanz, C., Aldebert, P., Althorpe, N., Baker, W., Baldwin, A., Bates, K., . . . Apweiler, R. (2005). The EMBL nucleotide sequence database. Nucleic Acids Research, 33, D29-D33. doi:10.1093/nar/gki098

Karagiannidis, I., Dehning, S., Sandor, P., Tarnok, Z., Rizzo, R., Wolanczyk, T., . . . Lehmkuhl, G. (2013). Support of the histaminergic hypothesis in tourette syndrome: Association of the histamine decarboxylase gene in a large sample of families. Journal of Medical Genetics, 50(11), 760-764.

Kumar, A., Trescher, W., & Byler, D. (2016). Tourette syndrome and comorbid neuropsychiatric conditions. Current Developmental Disorders Reports, 3(4), 217-221.

Lam, S., Shen, Y., Nguyen, T., Messier, T. L., Brann, M., Comings, D., . . . O'Dowd, B. F. (1996). A serotonin receptor gene (5HT1A) variant found in a tourette's syndrome patient. Biochemical and Biophysical Research Communications, 219(3), 853-858.

Lang, R., & Raffi, F. A. (2019). Dual-specificity phosphatases in immunity and infection: An update. International Journal of Molecular Sciences, 20(11), 2710.

Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. BMC Bioinformatics, 9, 559-559. doi:10.1186/1471-2105-9-559 [doi]

Lee, C. C., Chou, I. C., Tsai, C. H., Wang, T. R., Li, T. C., & Tsai, F. J. (2005). Dopamine receptor D2 gene polymorphisms are associated in taiwanese children with tourette syndrome. Pediatric Neurology, 33(4), 272-276. doi:S0887-8994(05)00258-4 [pii]

Lennington, J. B., Coppola, G., Kataoka-Sasaki, Y., Fernandez, T. V., Palejev, D., Li, Y., . . . Vaccarino, F. M. (2016). Transcriptome analysis of the human striatum in tourette syndrome. Biological Psychiatry, 79(5), 372-382. doi:10.1016/j.biopsych.2014.07.018

Lerer, E., Levi, S., Israel, S., Yaari, M., Nemanov, L., Mankuta, D., . . . Ebstein, R. P. (2010). Low CD38 expression in lymphoblastoid cells and haplotypes are both associated with autism in a familyâ€• based study. Autism Research, 3(6), 293-302.

Lin, P., Kuo, P., Chen, C., Wu, J., Gau, S. S., Wu, Y., & Liu, S. (2013). Runs of homozygosity associated with speech delay in autism in a taiwanese han population: Evidence for the recessive model. PLoS One, 8(8), e72056.

Lockstone, H. E. (2011). Exon array data analysis using affymetrix power tools and R statistical software. Briefings in Bioinformatics, 12(6), 634-644. doi:10.1093/bib/bbq086

Lorenzo, G., Braun, J., Munoz, G., Casarejos, M. J., Bazan, E., & Jimenez-Escrig, A. (2018). RNA-seq blood transcriptome profiling in familial attention deficit and hyperactivity disorder (ADHD). Psychiatry Research, 270, 544-546. doi:S0165-1781(18)30783-2 [pii]

Lowe, R., Shirley, N., Bleackley, M., Dolan, S., & Shafee, T. (2017). Transcriptomics technologies. PLoS Computational Biology, 13(5), e1005457.

Lu, Q. (2013). The critical importance of epigenetics in autoimmunity. Journal of Autoimmunity, 41, 1-5. doi:10.1016/j.jaut.2013.01.010 [doi]

Luo, L., Zhu, D., Huang, R., Xiong, L., Mehjabin, R., He, L., . . . Wang, Y. (2019). Molecular cloning and preliminary functional analysis of six RING-between-ring (RBR) genes in grass carp (ctenopharyngodon idellus). Fish & Shellfish Immunology, 87, 62-72.

Maia, T. V., & Conceição, V. A. (2018). Dopaminergic disturbances in tourette syndrome: An integrative account. Biological Psychiatry, 84(5), 332-344.

Mallya, M., Campbell, R. D., & Aguado, B. (2006). Characterization of the five novel Lyâ€• 6 superfamily members encoded in the MHC, and detection of cells expressing their potential ligands. Protein Science, 15(10), 2244-2256.

Martino, D., Chiarotti, F., Buttiglione, M., Cardona, F., Creti, R., Nardocci, N., . . . Italian Tourette Syndrome Study Group. (2011). The relationship between group A streptococcal infections and tourette syndrome: A study on a large service-based cohort. Developmental Medicine and Child Neurology, 53(10), 951-957. doi:10.1111/j.1469-8749.2011.04018.x [doi]

Martino, D., Zis, P., & Buttiglione, M. (2015). The role of immune mechanisms in tourette syndrome. Brain Research, 1617, 126-143. doi:10.1016/j.brainres.2014.04.027 [doi]

Mathews, C. A., Scharf, J. M., Miller, L. L., Macdonald-Wallis, C., Lawlor, D. A., & Ben-Shlomo, Y. (2014). Association between pre- and perinatal exposures and tourette syndrome or chronic tic disorder in the ALSPAC cohort. The British Journal of Psychiatry : The Journal of Mental Science, 204(1), 40-45. doi:10.1192/bjp.bp.112.125468 [doi]

Mina, E., van Roon-Mom, W., Hettne, K., Van Zwet, E., Goeman, J., Neri, C., . . . Roos, M. (2016). Common disease signatures from gene expression analysis in Huntington's disease human blood and brain. Orphanet Journal of Rare Diseases, 11(1), 1-13.

Mössner, R., Müller-Vahl, K. R., Döring, N., & Stuhrmann, M. (2007). Role of the novel tryptophan hydroxylase-2 gene in tourette syndrome. Molecular Psychiatry, 12(7), 617-619.

Moya, P. R., Dodman, N. H., Timpano, K. R., Rubenstein, L. M., Rana, Z., Fried, R. L., . . . Wendland, J. R. (2013). Rare missense neuronal cadherin gene (CDH2) variants in specific obsessive-compulsive disorder and tourette disorder phenotypes. European Journal of Human Genetics : EJHG, 21(8), 850-854. doi:10.1038/ejhg.2012.245 [doi]

Murphy, T. K., Kurlan, R., & Leckman, J. (2010). The immunobiology of tourette's disorder, pediatric autoimmune neuropsychiatric disorders associated with streptococcus, and related disorders: A way forward. Journal of Child and Adolescent Psychopharmacology, 20(4), 317-331.

Murphy, T. K., Lewin, A. B., Storch, E. A., & Stock, S. (2013). Practice parameter for the assessment and treatment of children and adolescents with tic disorders. Journal of the American Academy of Child & Adolescent Psychiatry, 52(12), 1341-1359.

Nakamura, N. (2011). The role of the transmembrane RING finger proteins in cellular and organelle function. Membranes, 1(4), 354-393. doi:10.3390/membranes1040354

Nazeen, S., Palmer, N. P., Berger, B., & Kohane, I. S. (2016). Integrative analysis of genetic data sets reveals a shared innate immune component in autism spectrum disorder and its co-morbidities. Genome Biology, 17(1), 228.

O'Leary, N.,A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., . . . Pruitt, K. D. (2016). Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. Nucleic Acids Research, 44, D733-D745. doi:10.1093/nar/gkv1189

Padmanabhuni, S. S. (2017). Computational methods of analysis for gene expression and genomewide variation data (Doctoral dissertation, Δημοκρίτειο Πανεπιστήμιο Θράκης (ΔΠΘ). Σχολή Επιστημών Υγείας. Τμήμα Μοριακής Βιολογίας και Γενετικής. Εργαστήριο Γονιδιωματικής Ποικιλότητας και Γενετικής Επιδημιολογίας).

Panula, P., & Nuutinen, S. (2013). The histaminergic network in the brain: Basic organization and role in disease. Nature Reviews Neuroscience, 14, 472. Retrieved from https://doi.org/10.1038/nrn3526

Paschou, P. (2013). The genetic basis of gilles de la tourette syndrome. Neuroscience and Biobehavioral Reviews, 37(6), 1026-1039. doi:10.1016/j.neubiorev.2013.01.016 [doi]

Pauls, D. L., Raymond, C. L., Stevenson, J. M., & Leckman, J. F. (1991). A family study of gilles de la tourette syndrome. American Journal of Human Genetics, 48(1), 154-163.

Pauls, D. L., Fernandez, T. V., Mathews, C. A., State, M. W., & Scharf, J. M. (2014). The inheritance of tourette disorder: A review. Journal of Obsessive-Compulsive and Related Disorders, 3(4), 380-385.

Qi, Y., Zheng, Y., Li, Z., & Xiong, L. (2017). Progress in genetic studies of tourette's syndrome. Brain Sciences, 7(10), 10.3390/brainsci7100134. doi:E134 [pii]

Qi, Y., Zheng, Y., Li, Z., Liu, Z., & Xiong, L. (2019). Genetic studies of tic disorders and tourette syndrome. In F. H. Kobeissy (Ed.), Psychiatric disorders: Methods and protocols (pp. 547-571). New York, NY: Springer New York. doi:10.1007/978-1-4939-9554-7_32 Retrieved from https://doi.org/10.1007/978-1-4939-9554-7_32

Quezada, J., & Coffman, K. A. (2018). Current approaches and new developments in the pharmacological management of tourette syndrome. CNS Drugs, 32(1), 33-45. doi:10.1007/s40263-017-0486-0

Rapanelli, M., Frick, L. R., & Pittenger, C. (2017). The role of interneurons in autism and tourette syndrome. Trends in Neurosciences, 40(7), 397-407. doi:10.1016/j.tins.2017.05.004

Robertson, M. M. (2008). The prevalence and epidemiology of gilles de la tourette syndrome. part 1: The epidemiological and prevalence studies. Journal of Psychosomatic Research, 65(5), 461-472. doi:10.1016/j.jpsychores.2008.03.006 [doi]

Robertson, M. M., Banerjee, S., Eapen, V., & Fox-Hiley, P. (2002). Obsessive compulsive behaviour and depressive symptoms in young people with tourette syndrome. European Child & Adolescent Psychiatry, 11(6), 261-265.

Robertson, M. M., Eapen, V., Singer, H. S., Martino, D., Scharf, J. M., Paschou, P., . . . Mathews, C. A. (2017). Gilles de la tourette syndrome. Nature Reviews Disease Primers, 3(1), 1-20.

Roessner, V., & Hoekstra, P. J. (2013). European multicenter tics in children studies (EMTICS): Exploring the onset and course of tic disorders. European Child & Adolescent Psychiatry, 22(7), 451-452. doi:10.1007/s00787-013-0437-3 [doi]

Schrag, A., Martino, D., Apter, A., Ball, J., Bartolini, E., Benaroya-Milshtein, N., . . . Efstratiou, A. (2019). European multicentre tics in children studies (EMTICS): Protocol for two cohort studies to assess risk factors for tic onset and exacerbation in children and adolescents. European Child & Adolescent Psychiatry, 28(1), 91-109.

Schrag, A., Martino, D., Apter, A., Ball, J., Bartolini, E., Benaroya-Milshtein, N., . . . EMTICS, C. G. (2019). European multicentre tics in children studies (EMTICS): Protocol for two cohort studies to assess risk factors for tic onset and exacerbation in children and adolescents. European Child & Adolescent Psychiatry, 28(1), 91-109. doi:10.1007/s00787-018-1190-4

Selling, L. (1929). The role of infection in the etiology of tics.

Singer, H. S. (2013). The neurochemistry of tourette syndrome.

Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., . . . Mazor, Y. (2016). The GeneCards suite: From gene data mining to disease genome sequence analyses. Current Protocols in Bioinformatics, 54(1), 1.30. 1-1.30. 33.

Tang, Y., Gilbert, D. L., Glauser, T. A., Hershey, A. D., & Sharp, F. R. (2005). Blood gene expression profiling of neurologic diseases: A pilot microarray study. Archives of Neurology, 62(2), 210-215.

Tarnok, Z., Ronai, Z., Gervai, J., Kereszturi, E., Gadoros, J., Sasvari-Szekely, M., & Nemoda, Z. (2007). Dopaminergic candidate genes in tourette syndrome: Association between tic severity and 3′ UTR polymorphism of the dopamine transporter gene. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 144B(7), 900-905. doi:10.1002/ajmg.b.30517

Thompson, M., Comings, D. E., Feder, L., George, S. R., & O'Dowd, B. F. (1998). Mutation screening of the dopamine D1 receptor gene in tourette's syndrome and alcohol dependent patients. American Journal of Medical Genetics, 81(3), 241-244. doi:10.1002/(SICI)1096-8628(19980508)81:3<241::AID-AJMG7>3.0.CO;2-Z [pii]

Tylee, D. S., Kawaguchi, D. M., & Glatt, S. J. (2013). On the outside, looking in: A review and evaluation of the comparability of blood and brain “‐omes”. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 162(7), 595-603.

Uhlen, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhori, G., . . . Edfors, F. (2017). A pathology atlas of the human cancer transcriptome. Science, 357(6352)

van Dam, S., Vosa, U., van der Graaf, A., Franke, L., & de Magalhaes, J. P. (2018). Gene co-expression analysis for functional classification and geneâ€“disease predictions. Briefings in Bioinformatics, 19(4), 575-592.

Vartanian, K., Slottke, R., Johnstone, T., Casale, A., Planck, S. R., Choi, D., . . . Harrington, C. A. (2009). Gene expression profiling of whole blood: Comparison of target preparation methods for accurate and reproducible microarray analysis. BMC Genomics, 10, 2. doi:10.1186/1471-2164-10-2

Wang, S., Mandell, J. D., Kumar, Y., Sun, N., Morris, M. T., Arbelaez, J., . . . Zhao, X. (2018a). De novo sequence and copy number variants are strongly associated with tourette disorder and implicate cell polarity in pathogenesis. Cell Reports, 24(13), 3441-3454. e12.

Wheeler, D. L., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Chetvernin, V., . . . Federhen, S. (2007). Database resources of the national center for biotechnology information. Nucleic Acids Research, 36(suppl_1), D13-D21.

Willsey, A. J., Fernandez, T. V., Yu, D., King, R. A., Dietrich, A., Xing, J., . . . Richer, P. (2017). De novo coding variants are strongly associated with tourette disorder. Neuron, 94(3), 486-499. e9.

Yoon, D. Y., Gause, C. D., Leckman, J. F., & Singer, H. S. (2007). Frontal dopaminergic abnormality in tourette syndrome: A postmortem analysis. Journal of the Neurological Sciences, 255(1-2), 50-56. doi:S0022-510X(07)00094-9 [pii]

Yu, D., Sul, J. H., Tsetsos, F., Nawaz, M. S., Huang, A. Y., Zelaya, I., . . . Hirschtritt, M. E. (2019). Interrogating the genetic determinants of tourette's syndrome and other tic disorders through genome-wide association studies. American Journal of Psychiatry, 176(3), 217-227.

Yu, G., Wang, L. G., Han, Y., & He, Q. Y. (2012). clusterProfiler: An R package for comparing biological themes among gene clusters. Omics : A Journal of Integrative Biology, 16(5), 284-287. doi:10.1089/omi.2011.0118 [doi]

Zhang, B., & Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. Statistical Applications in Genetics and Molecular Biology, 4, Article17-6115.1128. Epub 2005 Aug 12. doi:10.2202/1544-6115.1128 [doi]

Zhang, X., Lan, Y., Xu, J., Quan, F., Zhao, E., Deng, C., . . . Yan, M. (2019). CellMarker: A manually curated resource of cell markers in human and mouse. Nucleic Acids Research, 47(D1), D721-D728.

Zilhão, N. R., Olthof, M. C., Smit, D. J., Cath, D. C., Ligthart, L., Mathews, C. A., . . . Dolan, C. V. (2017). Heritability of tic disorders: A twin-family study. Psychological Medicine, 47(6), 1085-1096. doi:10.1017/S0033291716002981 [doi]