TCF4, Schizophrenia, and Pitt-Hopkins Syndrome

Derek J. Blake*, Marc Forrest, Ria M. Chapman, Caroline L. Tinsley, Michael C. O’Donovan, and Michael J. Owen

Department of Psychological Medicine and Neurology, Medical Research Council Center for Neuropsychiatric Genetics and Genomics, Cardiff University, Henry Wellcome Building, Heath Park, Cardiff CF14 4XN, UK
*To whom correspondence should be addressed; tel: 0044-2920-687051, fax: 0044-2920-687068, e-mail: blakedj@cardiff.ac.uk.

Genome-wide association studies allied with the identification of rare copy number variants have provided important insights into the genetic risk factors for schizophrenia. Recently, a meta-analysis of several genome-wide association studies found, in addition to several other markers, a single nucleotide polymorphism in intron 4 of the TCF4 gene that was associated with schizophrenia. TCF4 encodes a basic helix-loop-helix transcription factor that interacts with other transcription factors to activate or repress gene expression. TCF4 mutations also cause Pitt-Hopkins Syndrome, an autosomal-dominant neurodevelopmental disorder associated with severe mental retardation. Variants in the TCF4 gene may therefore be associated with a range of neuropsychiatric phenotypes, including schizophrenia. Recessice forms of Pitt-Hopkins syndrome are caused by mutations in NRXN1 and CNTNAP2. Interestingly, NRXN1 deletions have been reported in schizophrenia, whereas CNTNAP2 variants are associated with several neuropsychiatric phenotypes. These data suggest that TCF4, NRXN1, and CNTNAP2 may participate in a biological pathway that is altered in patients with schizophrenia and other neuropsychiatric disorders.

Key words: schizophrenia/mental retardation/transcription factor/Pitt-Hopkins syndrome/TCF4/NRXN1/CNTNAP2

Introduction

Although a small number of genetic loci have now been strongly implicated as risk factors for schizophrenia, most of these have yet to yield novel insights into the biology of the disease. A possible exception is the implication of neurexin-1 (NRXN1) and associated proteins in disease pathogenesis through the identification of copy number variants (CNVs) associated with an approximately 5-fold increased risk of schizophrenia. NRXN1 deletions are rare even in cases (0.19%), and this raises the question of the importance of NRXN1 for the disorder as a whole or at least a sizable proportion of it. Genome-wide association (GWA) studies have also led to the discovery of several new risk alleles for schizophrenia. Unlike NRXN1 deletions, these are common in the population, but the relative risks conferred are substantially lower (<1.5). For example, a synthesis of several recent GWA studies of schizophrenia with follow-up in additional samples identified a single-nucleotide polymorphism on chromosome 18q21.2 within transcription factor 4 (TCF4) that was associated with an increased risk of schizophrenia at a level of support ($P = 4.1 \times 10^{-9}$) that surpasses a widely held benchmark for genome-wide significance. TCF4 is a basic helix-loop-helix (bHLH) transcription factor that regulates gene expression in the immune system and in the brain during development. Haploinsufficiency of TCF4 causes a dominant form of Pitt-Hopkins Syndrome (PTHS), a developmental disorder associated with severe mental retardation. Remarkably, autosomal recessive forms of PTHS can also be caused by deletions and missense mutations in NRXN1 and another gene previously implicated by CNV analysis in schizophrenia and other neuropsychiatric diseases, contactin-associated protein like-2 (CNTNAP2). These findings potentially point to a functional link between TCF4 and both NRXN1 and CNTNAP2 and suggest that these proteins play a role in the pathogenic mechanisms of general relevance to schizophrenia and related disorders.

TCF4 and Brain Development

First, a word of caution concerning nomenclature is required. TCF4 (Gene ID: 6925) and TCF7L2 (Gene ID: 6934) loci are frequently confused because they share the TCF4 alias. TCF4 is the official symbol for TCF4, the protein discussed in this review, but it is also a widely used alternative name for T-cell–specific TCF4 (TCF7L2). In the present review, we have taken considerable care to include for consideration only those
data which apply to the former. TCF4 is a member of the bHLH transcription factor family homologous to the Drosophila protein daughterless. bHLH transcription factors can be divided into several phylogenetic groups based upon their sequence expression, composition pattern, and ability to interact with other bHLH proteins.7

Briefly, group A, which includes TCF4, and group B bHLH proteins bind core DNA sequences referred to as E (Ephrussi)-boxes defined loosely by the sequence CANNTG. Group C bHLH proteins are also known as bHLH-Per-Arnt-Sim (PAS) because in addition to the bHLH domain, they also contain a PAS domain. Group D HLH proteins lack a basic domain and are hence unable to bind DNA. Group E-proteins are related to Drosophila hairy and enhancer of split proteins and bind preferentially to N-box sequences (CACGCG or CACGAG). Finally, Group F is characterized by the presence of the Collier/Olfactory-1/Early B-Cell Factor domain that is involved both in dimerization and in DNA binding.7

TCF4 is 1 of 4 mammalian E-proteins, the others being E12, E47, and HeLa E-box-binding factor. Basic amino acids in the bHLH domain of this family of transcription factors bind directly to DNA, recognizing the E-box consensus sequence in the regulatory regions of many genes. Although TCF4 can form homodimers, in common with other bHLH proteins, it appears to function as a transcriptional activator or repressor only by forming heterodimers with other group A or B bHLH proteins including atonal homolog 1 (ATOH1) and achaete-scute complex homolog 1 (ASCL1) (figure 1).9–11 These interacting proteins are often expressed in a tissue- or cell-type-specific manner. By contrast, heterodimerization of E-proteins with group D HLH proteins abrogates their transcriptional activity by sequestering them into inactive complexes that cannot bind DNA.

The majority of functional studies on TCF4 concern its role in the immune system. Here TCF4 is required for the development of lymphoid progenitors to the B- and T-cell lineages and regulates plasmacytoid dendritic cell (PDC) differentiation.12–14 PDC cells secrete interferon in response to viral nucleic acids and form part of the innate immune response. Beyond the scope of this review, the role of TCF4 and associated transcription factors in the development of the immune system has been recently reviewed.15

During brain development, bHLH proteins modulate critical events in neuronal and glial progenitor cells, controlling the transition from proliferation to differentiation.16 Although TCF4 and the other E-proteins are highly expressed in neural progenitor cells, their role in brain development has not been studied in detail. Knockouts of the genes encoding each of the 4 E-proteins have been produced. In each case, Tcf4 included, mice that are homozygous null for any of the E-protein encoding genes die at birth, whereas heterozygous mice are viable.18 Clearly then, Tcf4 is required for postnatal survival. However, it appears to be dispensable, at least in mice for brain development because at the gross histiological level brain morphology appears normal in Tcf4 null animals.9

Although little is known about the role of TCF4 in the brain, it is useful to consider the roles of the 2 proneural
genes, ATOH1 (the mammalian orthologue of the Droso-
sophila atonal gene, also known as MATH1) and ASCL1,
that have been shown to form functional heterodimers
with TCF4 in this review. The proneural genes, of which
there are less than 10 in mammals, are key transcrip-
tional regulators of neurogenesis that specify all the differ-
ent neurons in the mammalian nervous system.16 During
brain development in the mouse, Atoh1 is essential for
the establishment of a neural progenitor population in
the rhombic lip and external granule layer that gives
rise to multiple hindbrain structures.17 Although Tcf4
is widely expressed, Atoh1 interacts with Tcf4 to form
neurons specifically in the pontine nucleus, a region in
the ventral pons that conveys information between the
motor cortex and cerebellum.9 In spite of no obvious neu-
rodevelopmental abnormality in Tcf4 null mice, Tcf4 is
required for the differentiation of subsets of neurons in
the developing brain. TCF4 also interacts with ASCL1
in SH-SY5Y neuroblastoma cells to form a heterodimer
that drives transcription of E-box–containing reporter
genes (figure 1).10 The formation of heterodimers be-
tween TCF4 and different proneural genes provides a
mechanism to regulate gene expression in specific sub-
sets of neuronal precursors. Temporal and spatial regu-
lation of neurogenesis by TCF4 and other E-proteins can
also be achieved by repressing the transcriptional activity
in a dominant-negative manner.18 Again in a neuronal
cell culture model, the transcriptional activity of TCF4
is attenuated through the formation of heterodimers
with the class D bHLH protein ID2 (figure 1).19
Surprisingly, little is known about the genes regulated
by TCF4 in the central nervous system. The rat tyrosine
hydroxylase enhancer has been shown to have an E-box–
binding site for Tcf4.20 That study also showed that
TCF4 could act as a transactivator but only when co-
expressed with the homeobox transcription factor
CUX/CDP2 (CCAAT displacement protein-2). Interes-
tingly, microarray detection of Atoh1-regulated genes in
the developing cerebellum found that Atoh1 targets are
E-box–regulated genes that cluster into a few functional
categories that include transcriptional regulation, cell
proliferation, and signal transduction.21 By comparison
with Atoh1, the identification of genes regulated by
TCF4 not only in the developing brain but also in adult
brain is likely to be pivotal to understanding the role of
TCF4 in schizophrenia.

Pitt-Hopkins Syndrome

Genetic studies of other common disorders such as type
II diabetes have shown that genes carrying common risk
alleles frequently contain rare, more highly penetrant
variants that can be associated with more severe pheno-
types.22 For example, loss of function mutations in WFS1
cause autosomal dominant Wolfram Syndrome whose
symptoms include early-onset non-autoimmune diabetes,
optic atrophy, and deafness.22 By contrast, common var-
iants in WFS1 confer risk of type II diabetes.22,23
Heterozygous deletions of the TCF4 gene cause PTHS
(OMIM: 610954); a neurodevelopmental disorder char-
acterized by severe mental retardation, microcephaly,
epilepsy, poor motor development, and breathing abnor-
malities.24 In these families, the disease is inherited in
a dominant manner and is the result of haploinsufficiency
of TCF4.25–27 In addition to chromosomal deletions,
TCF4 nonsense and missense mutations also cause dom-
inant forms of PTHS. In comparison to forms of PTHS
caused by deletions or nonsense mutations, TCF4 mis-
sense mutations are associated with an increased inci-
dence of seizures, suggesting subtle differences in the
disease mechanisms by class of mutation.28 Importantly,
most TCF4 missense mutations are located within the ba-
sic region of the bHLH domain and have been shown to
abrogate transcriptional activity in cells co-expressing
ASCL1.26,29 Heterodimerization of TCF4 with other
bHLH transcription factors such as ATOH1 and ASCL1
may explain one of the cardinal features of PTHS; a respi-
atory abnormality associated with hyperventilation and
apnea. Mice lacking either Atoh1 or Ascl1 die shortly after
birth due to an apparent inability to initiate respiration.17,30
In these mice and patients with PTHS, the functional
dependence of heterodimer formation between the products
of these proneural genes and TCF4 may result in a shared de-
ficit in formation or activity of subpopulations of neurons
that control breathing.

In addition to TCF4, autosomal recessive forms of
PTHS have recently been found to be caused by muta-
tions in NRXN1 and CNTNAP2 (Casp2).5 Neurexins
are synaptic cell adhesion molecules found on axon ter-
minals that together with their cognate neuroligins con-
nect the pre- and postsynaptic membranes of synapses.31
Three neurexin genes exist in humans, each one encodes 2
major isoforms (in the case of NRXN1, NRXN1α, and
NRXN1β) that are transcribed from different promoters.32
There is now strong evidence that deletions of NRXN1,
or parts of that gene, increase risk of schizophrenia, im-
pliying one or both of the major NRXN1 isoforms in the
etiolo2gy of that disorder.1

The second autosomal recessive PTHS gene CNTNAP2
has been implicated in the genetic etiology of several dis-
eases, again including schizophrenia (table 1). Truncating
mutations in CNTNAP2 cause autosomal recessive cortical
dysplasia-focal epilepsy (CDFE). CDFE is a rare congeni-
tal epilepsy syndrome with neuropsychiatric comorbidities
that include mental retardation, autism, and attention-
deficit hyperactivity disorder.33 By contrast, heterozygous
genomic deletions of CNTNAP2 have been described in
patients with mental retardation, epilepsy, and schizo-
phrenia.34 The expression of CNTNAP2 is regulated in
part by FOXP2; a forkhead transcription factor that has an
important role in the neurobiology of speech and lan-
guage acquisition.35 CNTNAP2 is also a member of the

Schizophrenia and PTHS
neurexin superfamily and is encoded by one of the largest genes in the human genome spanning in excess of 2 Mb. CNTNAP2 and its ligand contactin form a receptor signaling complex that mediates neuron-glial interactions and neuronal migration in the developing cortex and regulates the clustering of Kv1 channels at nodes of Ranvier.37 Histological examination of mice lacking CNTNAP2 revealed no gross abnormalities in the brain of homozygous animals contrasting with the defect in cortical lamination found in patients with CDFE.35,38

In a recent report, a small proportion of patients diagnosed with Angelman Syndrome, a neurodevelopmental disorder affecting the epigenetic regulation of the ubiquitin ligase UBE3A, were in fact found to have TCF4 mutations.39 The phenotypic similarities shared by these disorders include mental retardation and motor dysfunction. This study suggests that TCF4 mutations may be found in other mental retardation syndromes including Rett Syndrome that have a similar spectrum of phenotypes.

Conclusions

There is increasing evidence that schizophrenia results from a combination of rare mutations of relatively large effect and common variants that confer low risk. For example, common variants in TCF4 and rare CNVs in NRXN1 and, though the evidence is weaker, also CNTNAP2 are associated with schizophrenia. It can be argued that rare, highly penetrant variants are useful for defining biological pathways whose disruption can lead to schizophrenia, but their presence in only a limited number of cases means that further evidence is required to determine their relevance to schizophrenia more generally. The observation that rare mutations in TCF4, NRXN1, and CNTNAP2 can result in similar neurodevelopmental phenotypes suggests a functional link between the proteins they encode. The implication of a common variant in TCF4 in schizophrenia is therefore evident that the functions of NRXN1 and CNTNAP2 might also be of general importance in this disorder. In doing so, it demonstrates the utility of seeking both common and rare variants in complex genetic disorders.

The identification of biological pathways that may be altered in schizophrenia is a fundamental aim in deciphering the complex genetic factors that contribute to disease susceptibility. Several of the CNVs associated with schizophrenia and other neuropsychiatric disorders are found in genes that encode synaptic and neurodevelopmental genes.40 Moreover, some of the CNVs associated with schizophrenia are also found in patients with mental retardation, autism spectrum disorder, or bipolar disorder, suggesting that these may represent a continuum of overlapping phenotypes that range in severity of neurodevelopmental insult and age of onset.41,42 In common with many other GWA candidates in complex diseases, schizophrenia-associated mutations or alterations in gene expression have yet to be described for TCF4. Alterations in transcript levels, alternative splicing, or alternative promoters may generate subtle functional variants of TCF4 that could be altered in schizophrenia. While it is tempting to speculate that the expression of NRXN1 and CNTNAP2 may be regulated by TCF4, more work is required to further delineate the genetic role of TCF4 in schizophrenia.

Funding

Wellcome Trust (WT088866); Medical Research Council.

References


Table 1. Pitt-Hopkins Syndrome (PTHS) Genes and Their Association with Schizophrenia and Other Neuropsychiatric and Neurological Diseases

<table>
<thead>
<tr>
<th>Genes</th>
<th>Inheritance</th>
<th>Locus</th>
<th>Mutations</th>
<th>Genome-Wide Association Studies</th>
<th>CNV</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCF4</td>
<td>AD</td>
<td>18q21.1</td>
<td>PTHS</td>
<td>SZ*</td>
<td>—</td>
</tr>
<tr>
<td>NRXN1</td>
<td>AR</td>
<td>2p16.3</td>
<td>(i) PTHS</td>
<td>—</td>
<td>SZ1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(ii) ASD</td>
<td>ASD2</td>
<td></td>
</tr>
<tr>
<td>CNTNAP2</td>
<td>AR</td>
<td>7q35</td>
<td>(i) PTHS</td>
<td>—</td>
<td>SZ3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(ii) CDFE</td>
<td>ASD3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(iii) ASD</td>
<td>Epilepsy33</td>
<td></td>
</tr>
</tbody>
</table>

Note: This table summarizes the association of the 3 PTHS genes with different neuropsychiatric disorders. References are indicated in superscripts. AD, autosomal dominant; AR, autosomal recessive; ASD, autism spectrum disorder; CDFE, cortical dysplasia-focal epilepsy syndrome; SZ, schizophrenia.


34. Elia J, Gai X, Xie HM, et al. Rare structural variants found in attention-deficit hyperactivity disorder are preferentially associated with neurodevelopmental genes. *Mol Psychiatry*. In press.


