DOCTORAL THESIS

Studies of the Drosophila Basic Helix-loop-helix Transcription Factor Daughterless and its Mammalian Homologue Transcription Factor 4

Laura Tamberg

TALLINNA TEHNIKAÜLIKOOL
TALLINN UNIVERSITY OF TECHNOLOGY
TALLINN 2023

TALLINN UNIVERSITY OF TECHNOLOGY
DOCTORAL THESIS
44/2023

Studies of the Drosophila Basic Helix-loop-helix Transcription Factor Daughterless and its Mammalian Homologue Transcription Factor 4

LAURA TAMBERG

TAL TECH PRESS

TALLINN UNIVERSITY OF TECHNOLOGY
School of Science
Department of Chemistry and Biotechnology
This dissertation was accepted for the defence of the degree of Doctor of Philosophy in Chemistry and Biotechnology on 11/04/2023

Supervisor: Mari Palgi, PhD
Department of Chemistry and Biotechnology
Tallinn University of Technology
Tallinn, Estonia

Co-supervisor: Prof Tõnis Timmusk, PhD
Department of Chemistry and Biotechnology
Tallinn University of Technology
Tallinn, Estonia

Opponents: Prof Alicia Hidalgo, PhD
School of Biosciences
University of Birmingham
Birmingham, United Kingdom

Prof Osamu Shimmi, PhD
Institute of Molecular and Cell Biology
University of Tartu
Tartu, Estonia

Defence of the thesis: 11/10/2023, Tallinn

Declaration:
Hereby I declare that this doctoral thesis, my original investigation and achievement, submitted for the doctoral degree at Tallinn University of Technology has not been submitted for doctoral or equivalent academic degree.

Laura Tamberg

---
signature

European Union
European Regional
Development Fund

Investing
in your future

ISSN 2585-6898 (publication)
ISBN 978-9916-80-035-5 (publication)
ISSN 2585-6901 (PDF)
ISBN 978-9916-80-036-2 (PDF)
Printed by Koopia Niini & Rauam

TALLINNA TEHNIKAÜLIKOOL
DOKTORITÖÖ
44/2023

Äädikakärbse aluselise heeliks-ling-heeliks transkriptsioonifaktori Daughterless ja tema imetaja homoloogi Transkriptsioonifaktor 4 uuringud

LAURA TAMBERG

TAL TECH
KIRJASTUS

Contents

List of Publications

The list of author's publications, on the basis of which the thesis has been prepared:

I **Tamberg, L**., Sepp, M., Timmusk, T. and Palgi, M.
Introducing Pitt-Hopkins syndrome-associated mutations of TCF4 to *Drosophila* daughterless.
Biology Open 4, 2015; 1762–1771. doi:10.1242/bio.014696.

II **Tamberg, L.,** Jaago, M., Säälik, K., Sirp, A., Tuvikene, J., Shubina, A., Kiir, C. S., Nurm, K., Sepp, M., Timmusk, T and Palgi, M.
Daughterless, the *Drosophila* orthologue of TCF4, is required for associative learning and maintenance of the synaptic proteome.
Disease Models and Mechanisms 13, 2020; dmm042747. doi:10.1242/dmm.042747.

III Sirp, A.*, Shubina, A.*, Tuvikene, J., **Tamberg, L.,** Kiir, C. S., Kranich, L. and Timmusk, T.
Expression of alternative transcription factor 4 mRNAs and protein isoforms in the developing and adult rodent and human tissues.
Frontiers in Molecular Neuroscience 15, 2022. DOI 10.3389/fnmol.2022.1033224.

* Equal contribution.

Author's Contribution to the Publications

Contribution to the papers in this thesis are:

I The author planned part and performed majority of the experiments. The author analysed most of the experiments and wrote majority of the manuscript draft and final manuscript.

II The author was part of conceptualising the idea, planned and performed majority of the experiments and analysed all the results. The author wrote majority of the manuscript draft and final manuscript.

III The author planned and performed part of the experiments and analysed their results. The author participated in editing the manuscript.

Introduction

Transcription factor 4 (TCF4) is a class I basic helix-loop-helix (bHLH) transcription factor that is important for the development of nervous system but functions also in the adult nervous system. Mutations in *TCF4* gene cause a rare neurodevelopmental disorder – Pitt-Hopkins syndrome (PTHS), mild to moderate intellectual disability and Fuchs corneal dystrophy, and variations in the *TCF4* gene are associated with schizophrenia and other neuropsychiatric disorders. Moreover, *TCF4* is implied in a variety of cancers. Being associated with many human diseases makes *TCF4* a gene with scientifically high interest and therefore there is a need for animal models for *TCF4*-related diseases. *Drosophila melanogaster* is widely used as a model for human diseases including intellectual disabilities. The *Drosophila* models are time and cost effective, making the fruit fly a good model organism for drug screening.

In this thesis, we have generated *Drosophila melanogaster* models for PTHS and other *TCF4*-related intellectual disabilities and validated these models for drug screening. In addition, we have gained knowledge about the functions of the TCF4 fly homolog Daughterless (Da) in the adult nervous system of *Drosophila melanogaster*. Moreover, we have conducted a systematic investigation of the expression of *TCF4* gene in rodents and humans paving the way for the development of gene therapy for PTHS and other *TCF4*-related diseases. The models and knowledge generated and used in this thesis are valuable tools for finding therapies for the diseases caused by alterations in *TCF4.*

Abbreviations

Ac	Achaete
Ase	Asense
Ato	Atonal
bHLH	basic helix-loop-helix
ChIP	chromatin immunoprecipitation
CNS	central nervous system
Da	Daughterless
E(spl)	Enhancer of Split
E-box	Ephrussi box
Emc	Extramacrohaete
Ey	Eyeless
Id	Inhibition of differentiation
KCNQ1	Potassium voltage-gated channel subfamily Q members 1
Lsc	Lethal of scute
Nrx-1	Neurexin-1
PNS	peripheral nervous system
Pros	Prospero
PTHS	Pitt-Hopkins syndrome
SAHA	suberoylanilide hydroxamic acid
Sc	Scute
SCN10A	Sodium voltage-gated channel alpha subunit 10
Sens	Senseless
TCF4	Transcription factor 4

1 Review of the Literature

1.1 Transcription factor 4, a class I basic helix-loop-helix transcription factor

Transcription factor 4 (TCF4) is a class I basic helix-loop-helix (bHLH) protein that binds the Ephrussi box (E-box) motif on DNA (Corneliussen et al., 1991; reviewed in Massari and Murre, 2000). TCF4 forms homodimers or heterodimers with class II bHLH proteins to activate transcription and also heterodimerizes with Inhibition of differentiation (Id) proteins and becomes inhibited by the interaction (Figure 1) (reviewed in Massari and Murre, 2000). TCF4 has functions in many developmental processes, including the development of the nervous system, but it is also important for the functioning of the postnatal nervous system (Figure 1) (reviwed in Chen et al., 2021; Meert et al., 2022; Teixeira et al., 2021). Mutations in *TCF4* gene can be detrimental to the nervous system causing diseases ranging from psychiatric diseases to severe intellectual disability syndromes (reviwed in Chen et al., 2021; Meert et al., 2022; Teixeira et al., 2021). In the following chapters, an overview of the functioning of TCF4 and diseases related to it is given.

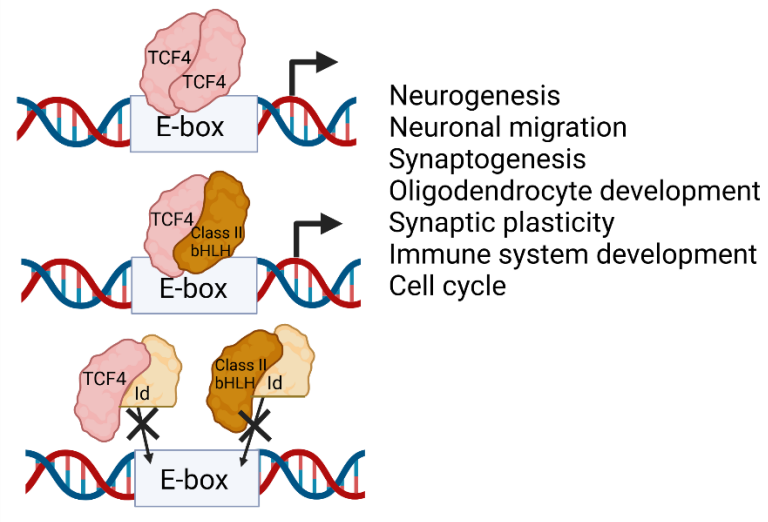


Figure 1. Functioning of TCF4. *TCF4 binds to the E-box as a homodimer or as a heterodimer with class II basic helix-loop-helix (bHLH) proteins and activates transcription. The activity is abolished by dimerization with Inhibition of differentiation (Id) proteins, since formed dimers cannot bind to E-box. TCF4 regulates neurogenesis, neuronal migration, synaptogenesis, oligodendrocyte development, synaptic plasticity, immune system and cell cycle. Figure created with BioRender.com.*

1.1.1 Functions of TCF4

TCF4 has important functions in a variety of developmental processes and in recent years a lot of advancements have been made in elucidating its functions in the development of nervous system. TCF4 has been shown to bind enhancers which regulate neurogenic transcription factor genes indicating that TCF4 is a key regulator of neurogenesis (Quevedo et al., 2019). More precisely, TCF4 regulates neuronal migration and

positioning (Chen et al., 2016; Flora et al., 2007; Zhang et al., 2021b), affects the development of cortical neurons (Mesman et al., 2020; Page et al., 2018; Rannals et al., 2016a; Rannals et al., 2016b; Schoof et al., 2020), hippocampus (Mesman et al., 2020; Schoof et al., 2020; Wang et al., 2020) and commissural formations (Mesman et al., 2020; Schoof et al., 2020; Wittmann et al., 2021). Also, neuronal morphology and synapse formation are partly regulated by TCF4 (D'Rozario et al., 2016; Phan et al., 2020; Rosato et al., 2021; Schoof et al., 2020). In addition to its roles in the development of neurons, TCF4 also regulates the development of oligodendrocytes and myelination (Fu et al., 2009; Haines et al., 2010; Phan et al., 2020; Wedel et al., 2020; Zhang et al., 2021a). In the adult nervous system, TCF4 has been shown to modulate neuronal excitability (Sarkar et al., 2021), synaptic plasticity (Crux et al., 2018), α7 nicotinic acetylcholine receptor signaling (Rex et al., 2017), adult neurogenesis in the hippocampus (Shariq et al., 2021), locomotion, cognition and learning (Li et al., 2019; Sarkar et al., 2021).

In addition to the roles of TCF4 in the nervous system, it has functions in the immune system during lymphoid cell specification and differentiation (Bouderlique et al., 2019; Qian et al., 2019), development of B and T cells (Bergqvist et al., 2000; Gloury et al., 2016; Wöhner et al., 2016; Zhuang et al., 1996) and development and maintenance of plasmacytoid dendritic cells (Cisse et al., 2008; Ghosh et al., 2010; Grajkowska et al., 2017; Onai et al., 2017). TCF4 has functions also in myogenesis (Chen and Lim, 1997; Lassar et al., 1991; Petropoulos and Skerjanc, 2000), mammary gland development (Itahana et al., 2008), testis development (Muir et al., 2006), endothelial cells development, (Yu et al., 2016a; Yu et al., 2016b), thymocyte development (Wikström et al., 2008), erythroid-megakaryocytic differentiation (In 't Hout et al., 2019), epidermis development (Kantzer et al., 2022) and melanocyte differentiation (Furumura et al., 2001). Additionally, TCF4 has been shown to regulate cell cycle (Schmidt-Edelkraut et al., 2014) and epithelial-mesenchymal transition (Cano and Portillo, 2010; Sobrado et al., 2009; Xiang et al., 2011).

1.1.2 Protein structure of TCF4

TCF4 gene expresses multiple protein isoforms, named TCF4-A, TCF4-B, etc, that are different in their N-termini (Nurm et al., 2021; Sepp et al., 2011). The expression is further diversified by internal alternative splicing, generating Δ and + or – isoforms (Sepp et al., 2011). All the isoforms contain transactivation domains AD2 and AD3, repression domain Rep and bHLH domain necessary for dimerization and binding DNA (Figure 3) (Forrest et al., 2012; Markus et al., 2002; Quong et al., 1993; Sepp et al., 2011; Sepp et al., 2012; Wy et al., 2013). Long isoforms have additional transactivation domain AD1 (Figure 3) (Massari et al., 1996; Sepp et al., 2012). Long and medium isoforms contain a nuclear localization signal which also acts as a nucleolus localization signal and which is missing in short isoforms (Figure 3) (Greb-Markiewicz et al., 2019; Sepp et al., 2012). Additionally, there are one nuclear localisation and two nuclear export signals in the bHLH domain of TCF4 (Greb-Markiewicz et al., 2019). The variety of TCF4 isoforms is probably needed for precise regulation of TCF4 activity since the isoforms differ in their capability of activating transcription (Nurm et al., 2021; Sepp et al., 2017).

1.1.3 Expression of TCF4 mRNA and protein

Tcf4 mRNA is expressed widely in most tissues of the mouse embryo (Murakami et al., 2004; Pscherer et al., 1996; Soosaar et al., 1994). More precisely, during the development of mouse embryonic central nervous system (CNS) *Tcf4* mRNA is expressed in the developing cortex, cerebellum, pons, medulla, hippocampus, olfactory bulb and spinal

cord (Chen et al., 2016; Mesman et al., 2020; Soosaar et al., 1994). *Tcf4* mRNA continues to be expressed in the adult hippocampus, cortex and cerebellum (Brzózka et al., 2010; Chiaramello et al., 1995; Wang et al., 2020) and the expression continues at low levels in the aging mouse brain (Uittenbogaard and Chiaramello, 2000).

TCF4 protein expression has been investigated in the developing mouse brain using antibodies that recognise only long TCF4 isoforms (Jung et al., 2018). TCF4 protein has been found to be expressed highly in the developing hippocampus and cortex and at low levels in other parts of the brain. In the adult mouse, TCF4 protein levels are the highest in the hippocampus, cortex, cerebellum and amygdala (D'Rozario et al., 2016; Jung et al., 2018). In addition to neurons, TCF4 protein is also expressed in astrocytes and oligodendrocytes, but not in microglial cells (Jung et al., 2018; Kim et al., 2020).

TCF4 mRNA expression has also been investigated in human embryonic development. Similar to mouse, *TCF4* is expressed at high levels in the developing CNS, especially in the hippocampus, cortex and cerebellum (de Pontual et al., 2009; Jung et al., 2018; Sepp et al., 2011; Wang et al., 2022). Additionally, during human embryonic development *TCF4* is expressed widely in other tissues like the spleen, uterus, lung, liver, pancreas, colon, developing digits, pituitary gland, gonads, kidney, thyroid and thymus (de Pontual et al., 2009; Sepp et al., 2011). In the adult human, *TCF4* mRNA is expressed in the CNS, lymphocytes, fibroblasts, gut, muscle, heart and lung (de Pontual et al., 2009; Pscherer et al., 1996).

1.1.4 Regulation of TCF4 activity and expression

Transcriptional activity of TCF4 is regulated by neuronal activity (Sepp et al., 2017). It has been shown that Protein Kinase A phosphorylates TCF4 and increases its ability to activate transcription (Sepp et al., 2017). The activity of TCF4 is also regulated by Calmodulin which is involved in the calcium signal transduction pathway (Saarikettu et al., 2004).

Additionally, *TCF4* gene expression is regulated by Wnt signalling pathway (Hennig et al., 2017; Kolligs et al., 2002). There is also evidence that E-proteins cross regulate – it has been shown that Transcription Factor 3, which is also an E-protein, binds to an E-box in the intron 2 of *TCF4* gene and regulates its transcription (Li et al., 2019). In addition, multiple micro RNAs have been shown to regulate *TCF4* gene expression (Adusumilli et al., 2020; Han et al., 2016; He et al., 2022; Kwon et al., 2013; Li et al., 2011; Xue et al., 2019). Since *TCF4* is associated with many diseases there is ongoing work in finding drugs that modulate *TCF4* expression or its function. For example, it has been shown that histone deacetylase inhibitor suberoylanilide hydroxamic acid (SAHA) upregulates *TCF4* expression (Hennig et al., 2017; Kennedy et al., 2016)

1.1.5 Target genes of TCF4

In the nervous system TCF4 target genes have been investigated *in vitro* using DNA binding experiments, reporter constructs and transcriptomics. In cultured SH-SY5Y human neuroblastoma cells TCF4 binds to genes involved in nervous system development, ion transport and signal transduction (Forrest et al., 2018). Silencing of *TCF4* in the same cells indicates that TCF4 regulates TGF-β signalling, epithelial to mesenchymal transition, apoptosis, and more specifically proneural genes *Neurogenin 2* and *Achaete-Scute Family bHLH Transcription Factor 1 (ASCL1)* (Forrest et al., 2013). Using reporter assays in HEK293 cell line, TCF4 has been shown to regulate *Neurexin 1β* and *Contactin associated protein 2* promoters (Forrest et al., 2012). In addition to cell lines, neurons differentiated from neural stem cells or neural precursors have been used

to identify the target genes for TCF4. Chromatin immunoprecipitation (ChIP) sequencing experiments have confirmed that TCF4 binds to genes that regulate cell proliferation, apoptosis, neuronal development and differentiation, axon guidance and synaptogenesis, for example *Neurotrimin* and *Discs large MAGUK scaffold protein 2, Neurexin 1α and 1β* (D'Rozario et al., 2016; D'Rozario et al., 2016; Hennig et al., 2017), and *Neurexin 1α and 1β* expression levels have been shown to be regulated by TCF4 (D'Rozario et al., 2016). *TCF4* silencing experiments in neural progenitor cell lines show that genes involved in cell cycle regulation are affected by *TCF4* expression levels (Hill et al., 2017). As useful cell models, induced pluripotent stem cells (iPSCs) generated from fibroblasts derived from Pitt-Hopkins syndrome (PTHS) patients (a disease caused by *TCF4* haploinsufficiency, Amiel et al., 2007; Brockschmidt et al., 2007; Zweier et al., 2007) and differentiated into neural progenitor cells, neurons and brain organoids have also been used (Papes et al., 2022). Genes that are expressed differentially in these models include genes involved in adhesion, neurogenesis, neuronal identity, differentiation, and regulation of neuronal excitability, for example Potassium voltage-gated channel subfamily Q members *KCNQ2* and *KCNQ3* (Papes et al., 2022). In addition, in rat primary cortical neurons TCF4 has been shown to regulate the expression of Growth arrest and DNA-damage-inducible protein coding gene *Gadd45γ* (Sepp et al., 2017) *and Brain derived neurotrophic factor* (Esvald et al., 2022; Tuvikene et al., 2021).

In vivo transcriptomics analysis of *Tcf4* knock out mouse brains indicate that genes involved in cell adhesion and neuronal functioning are regulated by TCF4 (Zhang et al., 2021b). In mouse hippocampal development *Wnt7b* is a major downstream effector of TCF4, which is in accordance with TCF4 role in adhesion (Wang et al., 2020). TCF4 has also been shown to function in neuronal migration and one important downstream effector there is Bone morphogenetic protein Bmp7 (Chen et al., 2016). Two ion channels coding genes, *Potassium voltage-gated channel subfamily Q members 1 (KCNQ1)* and *Sodium voltage-gated channel alpha subunit 10 (SCN10A),* are also regulated by TCF4 (Li et al., 2020; Rannals et al., 2016a; Rannals et al., 2016b). Analysis of transcriptomes of multiple mouse models with loss-of function of *Tcf4* have revealed that genes functioning in forebrain development, neuron projection, axon development, excitatory synapses, postsynaptic density, axon ensheathment, myelination, membrane associated processes and adhesion are misexpressed (Phan et al., 2020; Sarkar et al., 2021; Zhang et al., 2021b). Taken together, using multiple experiments and models, it has been shown that TCF4 regulates genes involved in cell cycle, adhesion, neuronal development, ion transport, myelination, synaptogenesis and neuronal excitability.

1.1.6 TCF4-related diseases

TCF4 is crucial for the development of nervous system and aberrations in its functioning lead to various neurological diseases. One of these, PTHS is a rare and severe disease caused by *de novo* heterozygous mutations in *TCF4* gene (Amiel et al., 2007; Brockschmidt et al., 2007; Zweier et al., 2007). The patients have severe mental retardation, developmental delay, microcephaly, abnormal brain waves, epilepsy, breathing problems, facial dysmorphism and other abnormalities (reviewed in Peippo and Ignatius, 2012). Anatomically, malformations of the brain include thin corpus callosum and small hippocampus (de Pontual et al., 2009). The mutations found in PTHS patients in *TCF4* locus are deletions, insertions, mutations in the bHLH region, mutations causing premature stop codon or aberrant splicing (Amiel et al., 2007; Brockschmidt et al., 2007; Sparber et al., 2020; Whalen et al., 2012; Zweier et al., 2007; Zweier et al.,

2008). Using fluorescence imaging, binding and reporter assays, it has been shown that PTHS-associated missense mutations impair the dimerization, DNA binding and the ability to activate transcription by TCF4 in an isoform and cell type specific manner (Forrest et al., 2012; Sepp et al., 2012; Sirp et al., 2021). The severity of the mutations varies from hypomorphic to dominant-negative. Mosaic mutations in *TCF4* can cause milder phenotypes and some cases of inheriting PTHS from mosaic parents have also been reported (de Pontual et al., 2009; Kousoulidou et al., 2013; Stavropoulos et al., 2010; Steinbusch et al., 2013). Mutations in *TCF4* can also cause less severe forms of intellectual disability – moderate and mild intellectual disability (Bedeschi et al., 2017; Kalscheuer et al., 2008; Kharbanda et al., 2016; Maduro et al., 2016; Masson et al., 2022).

TCF4 is also strongly associated with schizophrenia being one of the top genes linked to it and named as a master regulator of a gene network dysregulated in schizophrenia (Doostparast Torshizi et al., 2019; Ma et al., 2018). More specifically, a polymorphism in intron 4 of *TCF4* has been linked to schizophrenia in genome wide association studies (Li et al., 2010; Stefansson et al., 2009; Steinberg et al., 2011). Also, it has been found that *TCF4* mRNA levels are increased in the blood of schizophrenia psychosis patients (Wirgenes et al., 2012). In humans with schizophrenia risk allele of *TCF4* sensorimotor gating, sensory gating, verbal memory, working memory, reasoning and problem solving are affected (Albanna et al., 2014; Hall et al., 2014; Lennertz et al., 2011a; Lennertz et al., 2011b; Quednow et al., 2011; Quednow et al., 2012). In addition, chromosome inversions that affect *TCF4* gene have been associated with schizophrenia (Pickard et al., 2005).

Moreover, *TCF4* has been associated with bipolar disorder (Mokhtari et al., 2022; Pickard et al., 2005), depression (Amare et al., 2020; Cai et al., 2021; Mossakowska-Wójcik et al., 2018), autism spectrum disorders (O'Donnell et al., 2010; Talkowski et al., 2012), post-traumatic stress disorder (Gelernter et al., 2019), insomnia (Cai et al., 2021) and alcohol use disorder (Hade et al., 2021). TCF4 is also linked to a neurodegenerative disease – Huntington disease, since TCF4 expression is reduced in the hippocampus of Huntington disease mouse models and human patients (Nurm et al., 2021). In addition, Fuchs corneal dystrophy, an eye disease, is caused by lengthening of a trinucleotide CTG repeat in intron 3 of *TCF4* (Baratz et al., 2010; Foja et al., 2017; Soh et al., 2019; Vasanth et al., 2015). The repeat can cause the disease in multiple ways, one of which being changes in *TCF4* expression levels (reviewed in Fautsch et al., 2021; Okumura et al., 2019; Sirp et al., 2020).

TCF4, functioning in a variety of developmental processes and in cell cycle (Schmidt-Edelkraut et al., 2014) and in epithelial-mesenchymal transition (Cano and Portillo, 2010; Sobrado et al., 2009; Xiang et al., 2011), is also associated with many cancers (Kolligs et al., 2002). Misexpression of *TCF4* has been reported in gastric (Kim et al., 2008), colon (Herbst et al., 2009a), colorectal (Herbst et al., 2009b; Savio et al., 2016), breast (Appaiah et al., 2010; Yamamoto and Yamamoto, 2007), endometrial (Zhao et al., 2020) and prostate cancers (Yamamoto and Yamamoto, 2007). In addition, *TCF4* is associated with melanoma (Dunn et al., 2006; Luo et al., 2016), diffuse large B-cell lymphoma (Care et al., 2013; Jain et al., 2019), acute myeloid leukemia (in 't Hout et al., 2014), blastic plasmacytoid dendritic cell neoplasm (Ceribelli et al., 2016) and medulloblastoma (Wong et al., 2020). It has also been shown that TCF4 is involved in chemotherapy response and metastasis (Hur et al., 2017; Pernía et al., 2020; Xiang et al., 2011).

1.1.7 Animal models of TCF4-related diseases

Tcf4 knock-out mice die immediately after birth (Fu et al., 2009) and therefore the physiological functions of TCF4 in postnatal organism are possible to study only in *Tcf4* haploinsufficient mice. *Tcf4* haploinsufficient mice have also been used for studying diseases where TCF4 function is impaired. Anatomically they exhibit microcephaly with reduced cortical thickness and smaller hippocampi, their neuronal migration is delayed, dendrite and synapse formation is defective, adult hippocampal neurogenesis is reduced and myelination is defective (Braun et al., 2020; Jung et al., 2018; Li et al., 2019; Phan et al., 2020; Thaxton et al., 2018). *Tcf4* heterozygous knock-out mice have enhanced long-term potentiation in the hippocampus (Kennedy et al., 2016; Thaxton et al., 2018) and expression levels of the ion channel expressing genes *Kcnq1* and *Scn10a* are increased (Rannals et al., 2016b). These mice have behavioural problems – defective social interaction, ultrasonic vocalization, prepulse inhibition, spatial and associative learning and memory, reduced anxiety and breathing problems (Cleary et al., 2021; Kennedy et al., 2016; Thaxton et al., 2018). Additionally, conditional *Tcf4* knock out mice have been used for modelling of *TCF4*-related intellectual disabilities. If *Tcf4* is knocked out in the neural progenitor cells of hippocampus, then hippocampus becomes defective and small, and social memory of the animals is impaired (Wang et al., 2020). *Tcf4* gene inactivation in the neocortex causes mispositioning of neurons (Zhang et al., 2021b). Loss of TCF4 in adult and developing brain results in changes in dendrite morphology, and defects in excitability and synaptic plasticity (Crux et al., 2018; Sarkar et al., 2021; Schoof et al., 2020).

Mice with overexpressed levels of *Tcf4* can be used as schizophrenia models since there is evidence that *TCF4* levels are increased in the blood of schizophrenia patients (Wirgenes et al., 2012). These animals have impaired cognitive processes and sensorimotor gating (Brzózka and Rossner, 2013; Brzózka et al., 2010; Volkmann et al., 2020). Elevated TCF4 levels cause enhanced long-term depression, higher density of immature dendritic spines during development and elevated levels of proteins involved in synaptic function and metabolic pathways (Badowska et al., 2020).

Some potential therapeutics have been proposed using mouse models for *TCF4*-related diseases. It has been shown that histone deacetylase inhibitor SAHA rescues memory deficit in loss-of-function *Tcf4* model (Kennedy et al., 2016). The voltage-gated sodium channel $Na_v1.8$ has been proposed as a high priority target for PTHS therapy (Cleary et al., 2021; Martinowich et al., 2022) since the gene coding for $Na_v1.8$-*SCN10A* is regulated by TCF4 (Li et al., 2020; Rannals et al., 2016b). One of the potential drugs for inhibiting $Na_v1.8$ is Nicardipine. In the PTHS mouse model Nicardipine rescued behavioural deficits (Ekins et al., 2020). Another method for treating PTHS could be viral gene therapy. Using a PTHS mouse model it has been shown that reinstating *Tcf4* expression in neurons improves anxiety, activity levels, innate behaviours and memory of the mice (Kim et al., 2022). In schizophrenia model of *Tcf4* overexpressing mice spironolactone and aripiprazole improved deficits in cognition (Stephan et al., 2022).

Using mouse models for initial drug screening is time consuming and expensive. *Drosophila melanogaster* is widely used as a time and cost-efficient model for human diseases and could be used for studying *TCF4*-related diseases. The TCF4 homolog in *Drosophila melanogaster* is Daughterless (Da), which will be introduced next.

1.2 Daughterless, the *Drosophila melanogaster* homolog of TCF4

The name *daughterless (da)* was given to the gene in 1954 by Bell when an observation was made that female *Drosophila melanogaster* carrying homozygous spontaneous mutation da^1 have only male progeny (Sandler, 1972). Interestingly, this phenotype is temperature sensitive (Cline, 1976). Maternally contributed Daughterless (Da) participates in sex determination as a heterodimer with Scute (Sc) by transcriptionally regulating *Sex-lethal* gene which in the case of X chromosome to autosomal chromosome ratio 1:1 (XXAA) is activated and induces female phenotype. In the case of X chromosome to autosomal chromosome ratio 1:2 (XAA) *Sex-lethal* is inactive and embryos develop as males (Cline, 1988; Deshpande et al., 1995; Hoshijima et al., 1995).

In addition to its maternally contributed sex determination effects, Da has multiple zygotic functions (Cronmiller and Cline, 1987). In *da* mutant embryos peripheral nervous system (PNS) is completely missing, CNS has severe defects and these embryos have also defects in muscles and tracheas (Caudy et al., 1988a).

da gene was cloned in 1988 and it was found that *da* has a region similar to *myoD1* and *Achaete-Scute Complex* genes (Caudy et al., 1988b; Cronmiller et al., 1988). Soon Da protein sequence similarities to human Transcription factor 3 were found and predicted that the similar sequences form a helix-loop-helix (HLH) domain (Murre et al., 1989a). Same year, it was shown that Da forms heterodimers with Achate-Scute Complex proteins and binds E-box sequence also as a homodimer (Figure 2) (Murre et al., 1989b). In the following chapters, an overview of the functioning and molecular biology of Da will be given.

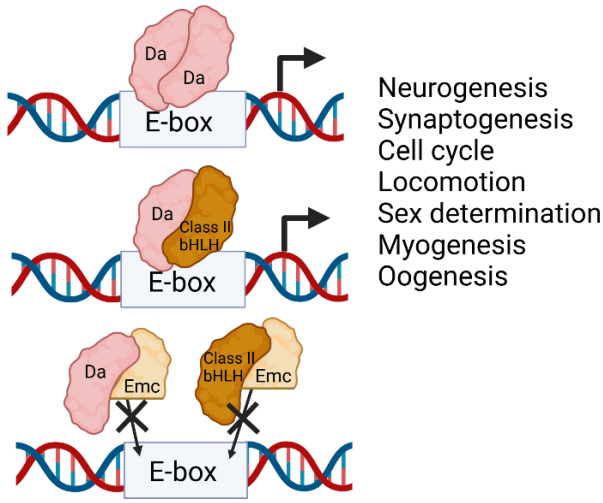


Figure 2. Functioning of Da. *Da binds to the E-box as a homodimer or as a heterodimer with class II basic helix-loop-helix (bHLH) proteins and activates transcription. The activity is abolished by dimerization with Extramacrochaete (Emc), since formed dimers cannot bind to E-box. Da regulates neurogenesis, synaptogenesis, cell cycle, locomotion, sex determination, myogenesis, oogenesis. Figure created with BioRender.com.*

1.2.1 Functions of Da

As already mentioned, Da is important for nervous system development (Figure 2) (Caudy et al., 1988a). During embryonic CNS development Da is required for determining neuronal development of neuroblasts but not for the neuroblast segregation (reviewed

in Campos-Ortega and Knust, 1990; Kovalick and Beckingham, 1992). In embryonic PNS development Da is even more important since no PNS precursor cells arise in embryos where Da is absent (Kovalick and Beckingham, 1992). In the larval brain Da is also required for neuroblast differentiation (Neumüller et al., 2011; Yasugi et al., 2008; Yasugi et al., 2014). Silencing of *da* in the larval brain causes neuroblast proliferation and formation of brain tumours (Neumüller et al., 2011; Yasugi et al., 2014). This is explained by downregulation of the Da target gene *prospero (pros)* that is a critical differentiation factor for neuroblasts (Yasugi et al., 2014).

The functions of Da in the development of the adult sensory organs have been extensively studied. One of the most investigated functions of Da is its role in the third instar larval eye imaginal discs during eye development (Bhattacharya and Baker, 2011; Brown et al., 1996; Cadigan et al., 2002; Chen and Chien, 1999; Li and Baker, 2019; Lim et al., 2008; Sukhanova et al., 2007; Tanaka-Matakatsu et al., 2014; Wang and Baker, 2015). In the eye disc Da has both proneural and anti-proneural functions depending on its expression level (Lim et al., 2008). During eye development a wave of differentiation called the morphogenetic furrow moves through the eye disc and Da is required for cell cycle preceding the wave (Brown et al., 1996). In the morphogenetic furrow Da is required for formation of R2-R5 and particularly R8 photoreceptors (Bhattacharya and Baker, 2011; Brown et al., 1996; Lim et al., 2008; Sukhanova et al., 2007). An interesting experiment where simultaneous overexpression of *da* and *eyeless (ey)* in the wing imaginal discs caused ectopic eye development highlighted the role of Da in the eye development (Tanaka-Matakatsu et al., 2014). In addition, Da also plays a crucial role during the development of wing, thorax and leg sensory organs, also called the bristles (Bhattacharya and Baker, 2011; Jafar-Nejad et al., 2006; Sukhanova et al., 2007).

Da has also functions in the mature nervous system (D'Rozario et al., 2016; Waddell et al., 2019). Like TCF4, Da functions in synaptogenesis. It has been shown that when Da levels are lower in the larval neuromuscular junctions of motoneurons then the number of synaptic boutons and axonal branching are increased, synaptic protein Bruchpilot is mis-localized and *neurexin-1 (nrx-1)* mRNA levels are increased (D'Rozario et al., 2016). On the contrary, Da overexpression causes a decrease in axonal branching, number of boutons and in neurotransmission at these synapses and decreased *nrx-1* levels. Larvae with Da overexpression or silencing were both shown to have impaired locomotion (D'Rozario et al., 2016).

In addition to being important for nervous system development and function, Da also has roles outside the nervous system (Figure 2). Maternally contributed Da functions in sex determination (Cline, 1988; Deshpande et al., 1995; Hoshijima et al., 1995). Zygotic Da is needed in endoderm for development of specific cell types (Tepass and Hartenstein, 1995), for left-right asymmetry formation in hindgut and therefore cell chirality (Ishibashi et al., 2019), for maintenance of intestinal stem cells (Bardin et al., 2010; Lan et al., 2018), during mesoderm development and myogenesis (Castanon et al., 2001; Wong et al., 2008), for corpora cardiaca precursor specification from head mesoderm (Park et al., 2011), for salivary glands development (Chandrasekaran and Beckendorf, 2003; Wang and Baker, 2015), during wing development for wing heart formation (Tögel et al., 2013), and during oogenesis (Cummings and Cronmiller, 1994; Smith et al., 2002).

Consistent with the known roles of Da in differentiation, Da has functions in cell cycle control (Andrade-Zapata and Baonza, 2014). In the wing imaginal discs, upregulation of Da causes cells to be arrested in G2 phase through Da regulating the Cdc25 phosphatase *string* (Andrade-Zapata and Baonza, 2014). In addition, Da regulates cell growth through

Salvador-Warts-Hippo pathway via regulating the expression of *expanded* (Wang and Baker, 2015). When Da is overexpressed in the eye imaginal discs then the eyes of the flies are smaller and in *expanded* null clones this growth inhibition by Da is abolished (Andrade-Zapata and Baonza, 2014).

1.2.2 Da protein

Da is the only class I bHLH protein in *Drosophila* (reviewed in Massari and Murre, 2000). It is 710 amino acid residues long and with a molecular mass of 74 kDa (Caudy et al., 1988b). Da is a very stable protein and it is phosphorylated (Kiparaki et al., 2015). Da forms dimers and binds to DNA using the bHLH domain (Figure 3) (Murre et al., 1989a; Zarifi et al., 2012). In addition it has two transactivation domains – AD1 and LH domain (Figure 3) (Zarifi et al., 2012). Experiments with expressing Da lacking either one of the transactivation domains in *da* mutant background have revealed that both activation domains are sufficient for Da activity (Zarifi et al., 2012). The AD1 also acts as a degron since when Da is in heterodimer with Sc then AD1 of Da promotes the degradation of Sc (Kiparaki et al., 2015). The LH domain is phosphorylated (Kiparaki et al., 2015). Also, Da has a repression (REP) domain (Figure 3) (Wong et al., 2008).

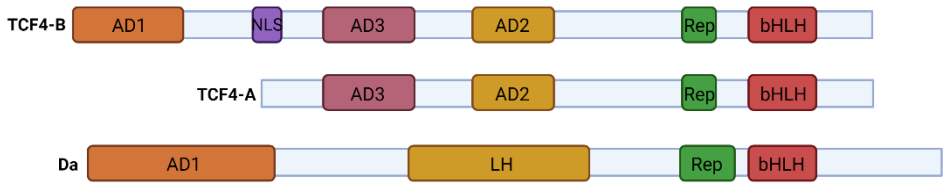


Figure 3. Schematic representation of TCF4-B, TCF4-A and Da. *TCF4 and Da have a conserved basic helix-loop-helix (bHLH) domain. Additionally, all TCF4 isoforms have a repression domain (Rep) and transactivational domains AD2 and AD3. Long and medium TCF4 isoforms have a nuclear localisation signal (NLS) domain and only long isoforms contain transactivational domain AD1. Similarly, Da contains a Rep domain and two transactivational domains AD1 and LH. The schemes are in scale. Figure created with BioRender.com.*

1.2.3 Expression of *da*

Three *da* transcripts are expressed from *da* gene - 3.2, 3.4 kb and 3.7 kb differing in the length of 3' untranslated region (UTR) (Caudy et al., 1988b; Cronmiller et al., 1988). The 3.7 kb transcript is not expressed in ovaries, the other transcripts are expressed in all tissues (Caudy et al., 1988b).

Da protein is expressed throughout embryogenesis (Cronmiller and Cummings, 1993; Vaessin et al., 1994). In preblastodermal embryos maternally contributed Da is present in all somatic cells and disappears just before blastoderm formation. Expression of *da* starts during gastrulation and peaks during stages 9-11 when neuroblasts form. During later stages Da expression is stronger in the nervous system, salivary glands, parts of the gut and muscles (Cronmiller and Cummings, 1993; Vaessin et al., 1994).

In third instar larval stage Da protein levels are high in the imaginal discs, salivary glands and CNS (Cronmiller and Cummings, 1993). In the larval CNS the level of Da protein is higher in the neuroblasts than their progeny (Vaessin et al., 1994). In the eye imaginal discs Da protein is expressed at low levels in all cells and the expression is elevated in the

morphogenic furrow where photoreceptor differentiation takes place and is highest in R8 photoreceptors (Brown et al., 1996; Vaessin et al., 1994). Similarly, in the wing and leg imaginal discs Da is expressed uniformly and increases in the sensory organ precursor (SOP) cells (Vaessin et al., 1994). So, in the larval CNS Da protein levels are higher in the neural stem cells but in the PNS Da protein levels are higher in the progenitor cells. Most recent study of Da protein expression found that Da is also expressed in postmitotic neurons in the 3rd instar larval ventral nerve cord (D'Rozario et al., 2016). This has opened up new avenues of investigating Da in mature nervous system.

In the adult *Drosophila* Da protein levels have only been investigated in the reproductive system (Cronmiller and Cummings, 1993). Both in the females and males Da protein is expressed in the somatic cells and not in the germ cells.

It is important to note that ubiquitous embryonic overexpression of Da results in lethality. It causes defective CNS condensation, and the ventral nerve cord is fragmented. However, the neuroblast patterning is normal. In addition, the pattern of PNS is disturbed. Ectopic neuronal cells are found in the gut, amnioserosa and salivary glands (Giebel et al., 1997). *Vice versa*, in *da* mutant embryos CNS has severe defects and PNS is completely missing (Caudy et al., 1988a). These embryos have also abnormal muscles and tracheas.

1.2.4 Regulation of Da expression and activity

The regulation of *da* expression has been mostly investigated using somatic clone technology in the 3rd instar larval eye imaginal discs where *da* is regulated by Wnt, Hedgehog, Decapentaplegic, Notch and EGFR signalling and also by Atonal (Ato) which is a class II bHLH protein and Da heterodimerization partner during eye development (Bhattacharya and Baker, 2011; Brown et al., 1996; Cadigan et al., 2002; Li and Baker, 2019; Lim et al., 2008). Additionally, using somatic clones in the eye disc it has been shown that Da expression is reduced in trithorax transcription factor *kismet* (Melicharek et al., 2008) and in FMR2/AF4 family protein coding gene *lilliputian* mutant clones (DiStefano et al., 2012). Using *Drosophila* S2 cell line, the regulation of *da* expression by Notch signalling has also been investigated (Mok et al., 2005; Wesley and Saez, 2000). When Notch C-terminal domain or full length Notch is over expressed in S2 cells then *da* mRNA levels are reduced (Wesley and Saez, 2000) and Delta overexpression increases *da* expression (Mok et al., 2005; Wesley and Saez, 2000).

Da has been shown to autoregulate its own expression (Bhattacharya and Baker, 2011; Smith and Cronmiller, 2001). In the ovaries *da* transcriptional autoregulation was shown by genetic and molecular experiments using a *da* allele which has an insulating element inserted in the only intron of *da* gene (Smith and Cronmiller, 2001). Initiation of *da* transcription uses an enhancer in the intron. After the initiation of expression, Da itself can positively autoregulate its own expression through E-boxes in the promoter region. There are also negative *cis*-acting sequences downstream of the insulating element that downregulate *da* expression so that Da levels do not become deleteriously high. *da* autoregulation has also been shown in the eye imaginal discs (Bhattacharya and Baker, 2011). This was demonstrated by overexpression of *da* without first non-coding exon which causes endogenous *da* mRNA levels to increase. Additionally, in the proneural regions of 3rd instar larval eye discs overexpression of class II bHLH protein Ato upregulates Da and in proneural regions of wing imaginal discs overexpression of class II bHLH protein Sc upregulates Da (Li and Baker, 2019). Since Da is an obligatory dimerization partner for class II bHLH proteins this again indicates to autoregulation. In the proneural regions of eye and wing imaginal discs *da* forms a

negative feedback loop with class V bHLH protein Emc which is the homolog for mammalian Id proteins (Bhattacharya and Baker, 2011; Li and Baker, 2018; Li and Baker, 2019). In this negative feedback loop Da regulates Emc expression positively and Emc overexpression has been shown to downregulate Da levels. Similar interaction system involving Emc also occurs in the mature neurons where silencing of *emc* increases Da expression and reduces *da* silencing phenotype and Emc overexpression reduces Da overexpression phenotype (Waddell et al., 2019).

In addition to being a negative regulator of Da expression level, Emc regulates Da function by regulating its activity post-translationally (Cabrera et al., 1994; Ellis et al., 1990; Van Doren et al., 1991). Emc as a class V bHLH protein lacks the basic region needed for the interaction with DNA but comprises the HLH domain for dimerization with Da or class II bHLH proteins (Ellis et al., 1990). *In vitro* binding assays have shown that Emc antagonizes the DNA binding ability of Da-Da homodimers or Da-class II bHLH protein heterodimers (Figure 2) (Cabrera et al., 1994; Van Doren et al., 1991).

Da is post-translationally regulated also by proteins that regulate its transactivational activity and by degradation. Transactivational capability of Da has been shown to be repressed by Escargot (Fuse et al., 1994), Deadpan (Hoshijima et al., 1995), Nervy (Wildonger and Mann, 2005), and Enhancer of Split (E(Spl)) (Oellers et al., 1994; Zarifi et al., 2012). Da has been shown to be ubiquinated by Seven in absentia which is *Drosophila* homolog for Siah-1 (Yang et al., 2010). There is also evidence that Da is targeted for degradation by E(spl) proteins by recruiting Groucho (Kiparaki et al., 2015).

1.2.5 Interactors of Da

It is commonly known that Da is an obligatory dimerization partner for proneural class II bHLH proteins. This interaction was demonstrated by electrophoretic mobility shift assays (EMSA) by Cabrera and Alonso, and Van Doren and colleagues in 1991 (Cabrera and Alonso, 1991; Van Doren et al., 1991). They found that Achaete (Ac), Sc and Lethal of scute (Lsc) cannot bind E-box sequences as homodimers or as heterodimers with each other but only as heterodimers with Da. They also demonstrated that Da-Da homodimers are able to bind DNA. Da and Sc interaction *in vivo* has been shown in the 3rd instar larval imaginal discs where even excess of Sc cannot induce neurogenesis in the absence of Da (Zarifi et al., 2012). Da also interacts with other class II bHLH proteins like Hand (Tögel et al., 2013), Twist (Castanon et al., 2001) and Sage (Chandrasekaran and Beckendorf, 2003).

As discussed above, transactivational activity of Da is negatively regulated by class V bHLH protein Emc. Da and Emc interaction has been demonstrated in yeast-two hybrid assay (Spratford and Kumar, 2015) and by proximity ligation assay in the larval motoneurons (Waddell et al., 2019). Evidence from luciferase reporter assays in a *Drosophila* cell line suggest that Emc antagonizes Da function by sequestration (Spratford and Kumar, 2015). Additionally, overexpression experiments in the 3rd instar larval eye and wing discs show that Da dramatically stabilizes Emc protein (Li and Baker, 2018). Other negative regulators of bHLH transcription factors from class VI, the E(Spl) proteins also interact with Da (Alifragis et al., 1997; Zarifi et al., 2012).

In addition to being regulated by Notch signalling (Mok et al., 2005; Wesley and Saez, 2000) it has been shown that Da is also involved in Notch signalling (Cave et al., 2005; Cave et al., 2009). Interestingly Da N-terminal domain without bHLH domain activates *E(spl)m8* promoter synergistically with Notch intracellular domain in *Drosophila* cell line and in yeast two-hybrid and pull-down assays Da N-terminal domain interacts with

Suppressor of Hairless (key transcriptional regulator of Notch signaling) (Cave et al., 2005; Cave et al., 2009).

There are a number of additional transcription factors that interact with Da. Pannier (GATA type Zn-finger transcription factor) has been shown to associate with Da through transcriptional co-factor Chip (Ramain et al., 2000), Nervy (MTG family transcriptional repressor) has been shown to interact with Da and repress its function (Wildonger and Mann, 2005), Senseless (Sens) (C2H2 type zinc finger transcription factor) has been shown to interact with Da in the wing imaginal discs for development of bristles (Jafar-Nejad et al., 2006), Escargot (Snail-type transcription factor) has been shown to interact with Da and induce its degradation (Yang et al., 2010), Ey (paired homeobox transcription factor) has been shown to interact with Da (Tanaka-Matakatsu et al., 2014), Pointed (ETS family transcription factor regulated by EGFR signalling) has been shown to activate *dacapo* enhancer synergistically with Ato-Da dimer (Sukhanova et al., 2007), FoxA (Fork Head transcription factor) has been shown to behave synergistically with Da in intestinal stem cell maintenance (Lan et al., 2018).

1.2.6 Targets of Da

As discussed above, Da is involved in sex determination through regulating *sex-lethal* (Cline, 1988; Deshpande et al., 1995; Hoshijima et al., 1995). Additionally, in early embryos *serendipity alpha* (needed for cellularization) is regulated by Da (Ibnsouda et al., 1995). During muscle development in the embryos Da has been found to regulate the expression of *myocyte enhancer factor 2* (Wong et al., 2008).

Da is well known for its role in embryonic neurogenesis (Caudy et al., 1988a). Da with its proneural dimerization partners (Ac, Sc, Lsc in embryos) is involved in the selection of sensory organ precursor cells (Kovalick and Beckingham, 1992). During this process Da-Ac, Da-Sc and Da-Lsc activate the transcription of *delta* in the neuroectodermal cell clusters (Kunisch et al., 1994). Additionally, Da regulates *sens* during this process as in *da* null embryos *sens* fails to express (Nolo et al., 2000), and *hunchback* as in *lsc* mutant embryos Hunchback expression is reduced (Cabrera and Alonso, 1991). Da-proneural protein heterodimers activate a number of neuronal precursor genes. In *da* mutant early embryos *pros*, *deadpan*, *asense (ase)*, *cyclin A, scratch* and *couch potato* expression is reduced (Bellen et al., 1992; Brand et al., 1993; Vaessin et al., 1994). *ase* expression is also dependent on the Achaete-Scute Complex genes since deletion of the genes causes loss of *ase* expression (Brand et al., 1993). Additionally, in embryos of later stages where *da* or the whole *Achaete-Scute Complex (ASC)* locus is deleted, *calmodulin* mRNA expression is disturbed (Kovalick and Beckingham, 1992). It has also been shown that Da regulates *single-minded* during CNS midline cell specification (Kasai et al., 1998).

Similarly to the embryos, in larval brains in *da* mutant clones of the optic lobes the expression of *pros* (Yasugi et al., 2014) and *deadpan* (Yasugi et al., 2008) is strongly reduced. When *da* is silenced in larval brains *nrx-1* mRNA levels are increased and when *da* is overexpressed then *nrx-1* levels are decreased (D'Rozario et al., 2016). Same was demonstrated in the neuromuscular junction for Nrx-1 protein and using a reporter. Using linked dimers it was also shown that Da mediates this at least in part as a homodimer (D'Rozario et al., 2016).

During larval stages, Da has been intensively studied in the 3rd instar imaginal discs where it regulates neurogenesis. In the 3rd instar wing imaginal discs, combining *ac* and *sc* mutant clones, fluorescence activated cell sorting and microarray analysis, the following Da-proneural target genes were identified: *charlatan, schizo, ETS-domain*

lacking, reduced ocelli, insensitive, meru, radial spoke head protein 3, malic enzyme, quail, miranda, phyllopod, rhomboid, serpin 43Aa, TNF-receptor-associated factor4, and *CG30492* (Reeves and Posakony, 2005). Using *in vivo* reporter assays, *phyllopod* has been shown to be directly regulated by Ac and Sc (Pi et al., 2004) and the regulation of *charlatan* has also been shown with Ac and Sc misexpression experiments (Escudero et al., 2005). Additionally, in the wing imaginal discs, using *da* mutant clones regulation of *ac* and *sc* by Da has been shown (Cadigan et al., 2002). In addition, Da regulates *expanded* in the wing discs and acts there as a homodimer (Wang and Baker, 2015; Wang and Baker, 2018). In the 3rd instar eye imaginal discs regulation of *ato* by Da has been extensively studied showing that Da acts there both as positive and negative regulator (Bhattacharya and Baker, 2011; Melicharek et al., 2008; Tanaka-Matakatsu et al., 2014). Additionally, in the eye discs Da regulates *dacapo* (Sukhanova et al., 2007; Tanaka-Matakatsu et al., 2014), *scabrous* (Bhattacharya and Baker, 2011)*, delta* (Bhattacharya et al., 2017) and *emc* (Li and Baker, 2019). The regulation of Cdc25 phosphatase *string* by Da has been demonstrated using both eye and wing imaginal discs as models (Andrade-Zapata and Baonza, 2014).

Outside imaginal discs, in the salivary glands, Da has been shown to bind *salivary gland secretion protein gene 4* (King-Jones et al., 1999). In addition, Da regulates *sens* in the salivary glands as a heterodimer with class II bHLH protein Sage (Chandrasekaran and Beckendorf, 2003). Additionally, using cell culture based transactivation assays *E(spl)* (Cave et al., 2009; Oellers et al., 1994) and *neuropeptide receptor NKD* (Rosay et al., 1995) have been shown to be regulated by Da. Taken together, Da regulates genes that function in sex determination, cellularization, myogenesis, neurogenesis and cell cycle.

Altogether, TCF4 and Da are homologs and they function in multiple developmental processes including neurogenesis, but also in the mature nervous system in regulating synapses. *TCF4* is related with multiple serious diseases of the nervous system and also cancer. Using *Drosophila* Da as a model could potentially give valuable knowledge of the molecular mechanisms of *TCF4*-related diseases.

2 Aims of the Study

The main aim of the thesis was to investigate Da and its mammalian homolog TCF4 to gain knowledge on *TCF4*-related diseases. More specifically the aims were as follows:

- Validate that Da and TCF4 are orthologs;
- Generate and characterize *Drosophila melanogaster* models with decreased expression levels of Da;
- Investigate the roles of Da in mature neurons;
- Validate generated models as potential tools for screening drugs for PTHS and other *TCF4*-related diseases;
- Characterise the expression of different *TCF4* transcripts and protein isoforms in developing and adult rodent and human tissues.

3 Materials and methods

The following methods, described in the indicated publications, were used in this study:

- *Drosophila melanogaster* rearing – publications I and II
- Molecular cloning – publications I, II and III
- CRISPR-Cas9 mediated tagging – publication II
- Transgenesis – publications I and II
- Growing of cell lines (HEK293, Neuro2a, SH-SY5Y) – publications I, II and III
- Growing of rat primary cortical neurons – publication II
- DNA transfection – publications I, II and III
- Luciferase reporter assay – publications I and II
- *In vivo* lacZ reporter assay – publication I
- RNA isolation, cDNA synthesis, RT-PCR or RT-qPCR – publications I, II and III
- Immunohistochemistry – publications I and II
- Laser confocal scanning microscopy – publications I and II
- Light microscopy and scanning electron microscopy – publication I
- *Drosophila* lifespan experiments – publication I
- Western blot analysis – publications II and III
- Appetitive associative learning assay of *Drosophila* larvae – publication II
- Chromatin immunoprecipitation – publication II
- Negative geotaxis assay – publication II
- Direct *Tcf4* RNA sequencing – publication III
- Tissue isolation and protein extraction – publications II and III
- *In vitro* protein translation – publication III
- Analysis of publicly available RNA-seq datasets – publication III

4 Results

Results obtained in Publication I are as follows:

- Da is the *Drosophila melanogaster* ortholog for human E-proteins;
- Introduction of PTHS-associated mutations located in the bHLH of TCF4 affects the transcriptional activity of Da;
- PTHS-associated mutations that affect arginines in the bHLH impair Da ability to activate reporter construct *in vivo* and neurogenesis in the 3rd instar larval wing imaginal discs. Two of the arginine mutations cause malformation of the thorax, indicating dominant-negative effects;
- Two of the PTHS-associated arginine mutations impair Da ability to activate neurogenesis in *Drosophila melanogaster* embryos;
- Overexpression of Da during eye development causes rough eye phenotype;
- Ubiquitous overexpression of Da in the adult flies after eclosion from the pupae causes lethality within days;
- Overexpressed human TCF4-A and TCF4-B function in *Drosophila melanogaster* similarly to Da – activate reporter construct, induce neurogenesis in the wing discs, replace Da during embryonic neurogenesis and cause rough eye phenotype. Only TCF4-B, but not TCF4-A, induces lethality when expressed ubiquitously after pupariation.

Results obtained in Publication II are as follows:

- A *Drosophila melanogaster* transgenic line was generated where Da is tagged with 3xFLAG tag, the tag does not interfere with the function of Da;
- Da is expressed in all developmental stages of *Drosophila melanogaster*. In the heads of adult flies Da expression decreases over time;
- Da is expressed widely in the 3rd instar larval brain including cells of the mushroom bodies;
- Silencing of *da* in the 3rd instar larval brain causes decrease in the levels of synaptic proteins Synapsin and Discs large 1;
- Da binds to *synapsin* promoter region and several areas in *discs large 1* gene in the adult *Drosophila* heads;
- Silencing of *da* in the 3rd instar larval brain or specifically in the mushroom bodies causes the impairment of appetitive associative learning;
- Silencing of *da* in specific areas of the CNS causes the impairment of negative geotaxis of the flies;
- Resveratrol and SAHA increase the transcriptional activity of Da, TCF4-A and TCF4-B;
- Resveratrol and SAHA alleviate the memory impairment of *da* silencing in larvae;
- SAHA rescues the locomotion impairment caused by *da* silencing.

Results obtained in Publication III are as follows:

- In the developing mouse cerebral cortex, transcripts that encode five N-terminally different Tcf4 isoforms are expressed;
- The main *Tcf4* transcripts in rodent brain code for Tcf4-A, Tcf4-B, Tcf4-D, Tcf4-C and Tcf4-I protein isoforms;
- TCF4 protein isoforms expressed in the mouse cerebral cortex can be grouped into long, medium, and short isoforms – long being TCF4-B and TCF4-C, medium being TCF4-D and short being TCF4-A and TCF4-I;

- In mouse and rat, postnatal expression of *Tcf4 i*s highest in the cerebral cortex, hippocampus and cerebellum;
- In the rodent brain, most prominent TCF4 protein isoforms are long and short isoforms. The medium isoforms are expressed at high levels in the cerebral cortex and hippocampus and at low levels in the cerebellum, olfactory bulb and pons;
- In mouse and rat non-neural tissues *Tcf4* is expressed at lower levels than in the brain, and mainly long and short isoforms are expressed;
- In human, *TCF4* mRNA expression is highest in the brain;
- The main transcripts expressed in the human brain encode TCF4-A, TCF4-D, TCF4-B, TCF4-C and TCF4-I;
- Short, medium and long TCF4 protein isoforms are expressed in human brain.

5 Discussion

5.1 *Drosophila melanogaster* is suitable for modelling PTHS

A lot of advancements have been made in understanding PTHS using mouse models of PTHS (Braun et al., 2020; Cleary et al., 2021; Jung et al., 2018; Kennedy et al., 2016; Li et al., 2019; Phan et al., 2020; Rannals et al., 2016b; Thaxton et al., 2018). *Drosophila melanogaster* models for human diseases are used widely because about 75% of human disease-related genes have homologs in the fly and in many cases using *Drosophila models* has proven to be beneficial (reviewed in Ecovoiu et al., 2022; Lasko and Lüthy, 2021; Nitta and Sugie, 2022; Yamaguchi et al., 2021). Intellectual disability syndromes have also been successfully modelled in the fruit fly (reviewed in Coll-Tané et al., 2019; van der Voet et al., 2014). It must be noted that the *Drosophila* models are very time and cost efficient, tissue specific overexpression and silencing of genes is easy since the tools are available at the stock centres, the tissues are available basically in unlimited amounts, and there are different behavioural assays available to test phenotypes. To model an intellectual disability syndrome, a fly homolog of the gene-of-interest should be identified, overexpression or silencing phenotype characterized, then the hypothesis of the disease can be tested, molecular pathways identified and drugs screened (Figure 4) (reviewed in van der Voet et al., 2014). Previously, the *Drosophila melanogaster*'s Da has been investigated in its genetic links to other intellectual disability-related genes and in habituation model (Fenckova et al., 2019; Straub et al., 2020).

In this thesis novel *Drosophila melanogaster* models for intellectual disability and PTHS were generated (Figure 4). First, we showed that human TCF4 and *Drosophila* Da are true orthologs, since TCF4 is functional in the fly. Rescuing mutant fly phenotype by expressing human disease-related gene is considered as a first step in validating a model (Ecovoiu et al., 2022). In our experiments, TCF4-A and TCF4-B were able to replace Da in the development of embryonic nervous system. Moreover, human TCF4-A and TCF4-B activated reporter expression in the larval wing discs, induced formation of excess bristles on the thorax and similarly to Da induced rough eye phenotype and lethality in adulthood.

As a second step, we introduced four PTHS-associated missense mutations into Da (Da^{PTHS}), expressed Da^{PTHS} in HEK293 cells and performed luciferase reporter assay. These mutations impaired Da transcriptional ability in the same way as the corresponding mutations in TCF4 ranging from reduced transcriptional activity to no activity at all (Sepp et al., 2012). In addition, we generated *Drosophila* lines enabling us to study Da^{PTHS} *in vivo.* Da^{PTHS} overexpression and rescue experiments showed, that the studied mutations impair Da functions *in vivo*, being hypomorphic, loss of function or dominant negative. The effects of some of the mutations were tissue specific. PTHS-associated mutations in TCF4 have also been shown to affect its function in different ways, being hypomorphic to dominant negative and the severity depends on the dimerization partner and the cell type (Forrest et al., 2012; Sepp et al., 2012; Sirp et al., 2021). Using *in vitro* reporter assays and *in vivo* overexpression and rescue experiments, we showed that TCF4 and Da function very similarly and are true orthologs and therefore *Drosophila melanogaster*'s Da can be used for modelling PTHS.

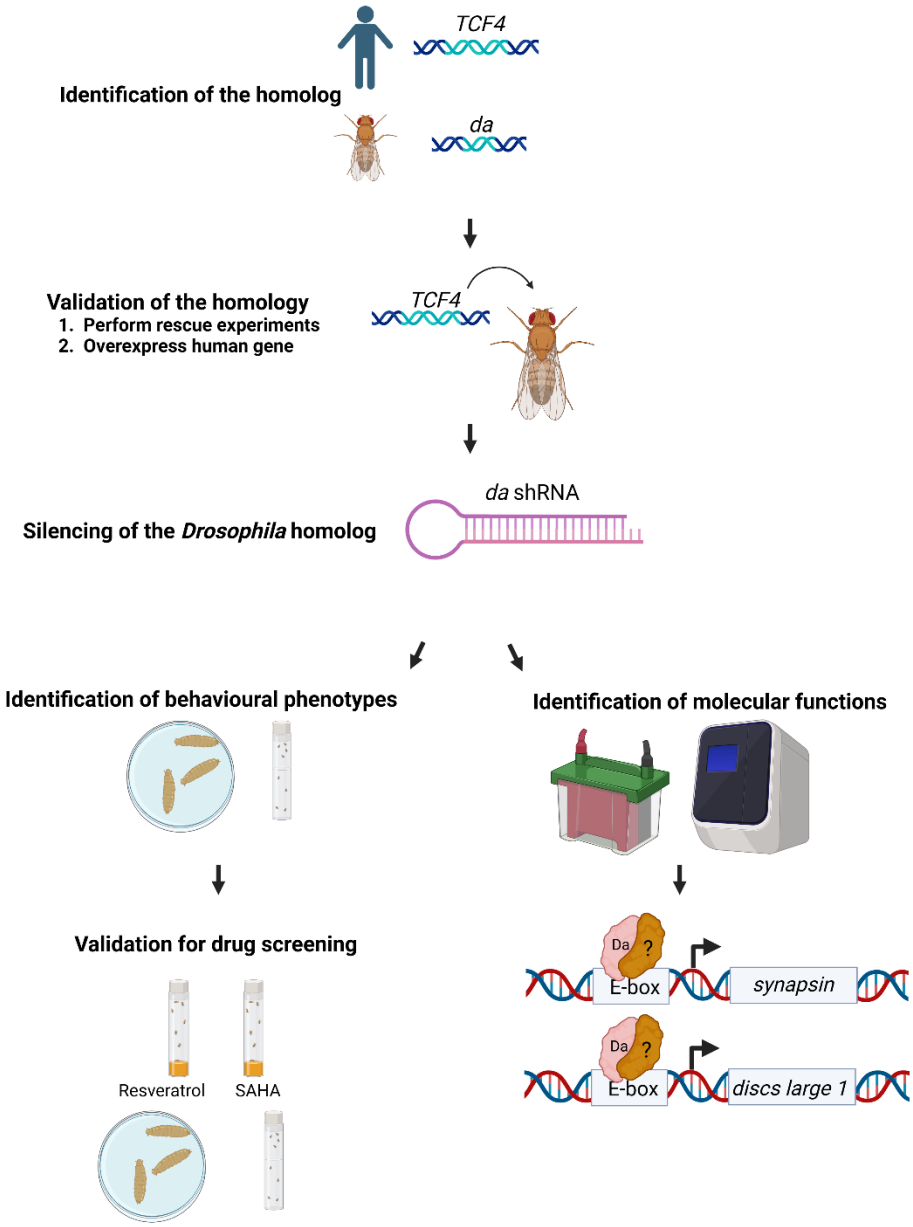


Figure 4. Generation and validation of Drosophila melanogaster models for TCF4-related diseases. *As a first step, Drosophila homolog for TCF4 was identified, then the homology was validated. Next, da silencing phenotype in behavioural paradigms was tested and the obtained models were validated for drug screening. For the identification of novel molecular functions of Da, molecular biology methods were used and Da involvement in regulating synaptic proteins was found. Figure created with BioRender.com.*

5.2 Da is important for learning and locomotion

PTHS, caused by mutations in *TCF4*, is an intellectual disability syndrome and mutations in *TCF4* cause mild to moderate intellectual disability (Bedeschi et al., 2017; Kalscheuer et al., 2008; Kharbanda et al., 2016; Maduro et al., 2016; Masson et al., 2022; reviewed in Peippo and Ignatius, 2012). It has been shown that mouse PTHS models have impaired memory (Kennedy et al., 2016; Thaxton et al., 2018). To generate the *Drosophila melanogaster* behavioural models for *TCF4*-related intellectual disabilities, we decided to use larval appetitive associative learning paradigm since no specific equipment is needed compared to adult learning paradigm. As we were unable to detect any aberrant phenotypes in the case of heterozygous *da* mutation, we decided to use wide silencing of *da* in CNS or more specifically in the mushroom bodies which are the memory centres of the fly. Using fly line where Da was tagged with 3xFLAG, wide Da expression in the larval CNS including cells in the mushroom body was found. Our *da* knock-down model was a success since silencing *da* widely in larval brain or in the mushroom bodies resulted in impaired larval appetitive associative learning.

When we allowed the larvae from the learning paradigm to develop into adulthood, we observed an interesting phenotype in one of the fly lines where *da* was silenced in the mushroom bodies but also in other cells of the CNS including numerous cells in the ventral nerve cord. These flies had defective locomotion which we measured using negative geotaxis assay. Our results are in accordance with previous results which showed that both silencing and overexpression of *da* impairs larval locomotion (D'Rozario et al., 2016). It must be noted that PTHS patient's movements are very limited (reviewed in Peippo and Ignatius, 2012). Taken together, we showed that Da silencing coupled with two behavioural assays – learning and negative geotaxis paradigms is suitable for modelling PTHS (Figure 4).

5.3 Da functions in regulating the synaptic proteome

The generated *Drosophila melanogaster* PTHS models displayed impaired learning and locomotion. We were interested what could be the mechanisms by which Da regulates these processes. Both TCF4 and Da have been shown to regulate synaptogenesis and in PTHS mouse models synaptic transmission is affected (D'Rozario et al., 2016; Phan et al., 2020). We hypothesized that the observed phenotypes could be associated with synapses. Indeed, in larval brains where d*a* was silenced, the levels of two synaptic proteins Synapsin and Discs Large 1 were decreased. In addition, Da interacted with the regulatory regions of these genes in adult *Drosophila* heads. It has been shown that the human homolog for *discs large 1* is TCF4 target gene (Hennig et al., 2017) but human homolog for *synapsin* has not yet been shown to be regulated by TCF4. Our findings support previous data that symptoms of PTHS and other neurological *TCF4*-related diseases are in part associated with dysregulation of synaptogenesis and synaptic plasticity.

5.4 *Drosophila* PTHS models can be used for drug screening

Next, we asked, if the generated *Drosophila melanogaster* PTHS models can be used for drug screening. We decided to test this using two drugs – histone deacetylase inhibitor SAHA and resveratrol. SAHA has been shown to rescue the memory impairment of PTHS mouse models (Kennedy et al., 2016) and we had unpublished data that treating rat neuronal cultures with resveratrol increases the transcriptional ability of TCF4. As a first

step, we tested whether these drugs promote an increase in the transcriptional activity of Da and TCF4 in a luciferase reporter assay in rat neuronal cultures. Indeed, treating neurons with resveratrol caused an increase in reporter activity already within 8 hours and the increase was stronger after 24-hour treatment. In addition, Da and TCF4 activity was increased after 24-hour treatment with SAHA. Based on these results we moved on to rescue experiments to validate our *Drosophila* models for drug screening.

For the appetitive associative learning experiment, *Drosophila* larvae with impaired memory caused by silencing of *da* were fed substrate supplemented with SAHA, resveratrol or just the corresponding vehicle during development. The memory of the larvae fed with SAHA or resveratrol containing substrate was moderately improved. Next, we attempted to rescue the impaired climbing phenotype caused by *da* silencing. We were interested, whether SAHA or resveratrol could improve locomotion of the flies when fed only to larvae during development or to adults after pupariation or both during development and in adulthood. To our surprise, neither of the drugs had any effect on rescuing the phenotype when fed only to larvae, but SAHA improved the climbing ability of female flies when fed after pupariation or both during development and adulthood. This was significant since it shows that increase in Da function after development is able to rescue a phenotype caused by decreased Da levels during development and possibly the same effects could be translated to diseases caused by decreased TCF4 function. In contrast to resveratrol having a moderate positive effect in the learning paradigm it had no effect on the locomotion phenotype.

The mechanisms by which resveratrol and SAHA increase the transcriptional activity of Da and TCF4 are not precisely known. It has been shown that resveratrol inhibits cAMP-degrading phosphodiesterases which leads to elevated cAMP levels (Park et al., 2012) and transcriptional activity of TCF4 is regulated by cAMP-PKA pathway-mediated phosphorylation (Sepp et al., 2017). Moreover, resveratrol is a powerful anti-inflammatory agent, it regulates apoptosis through multiple signaling pathways, it influences transcription factors associated with proliferation and stress response and it has epigenetic effects which makes it a potential therapeutic against cancer, inflammation, neurodegenerative diseases and even aging (reviewed in Vrânceanu et al., 2022). In this thesis, we propose that resveratrol could have a potential positive effect on the learning ability of PTHS and other *TCF4*-associated intellectual disability patients.

SAHA is a histone deacetylase inhibitor but the effects of SAHA are not only epigenetic since histone deacetylases also affect other proteins besides histones. Histone deacetylases also deacetylate DNA-repair proteins, cytoskeletal and adhesion proteins, cell signaling proteins and transcription factors (reviewed in Wawruszak et al., 2021). SAHA affects all these processes and this makes it a promising treatment for cancer. In our *Drosophila melanogaster* PTHS models, SAHA had positive effects on both learning and locomotion, and it has been previously shown to improve memory in PTHS mouse models. Taking all that into account shows that SAHA is a potential drug for PTHS and *TCF4*-related intellectual disability patients.

The fact that we were able to partially rescue our *da* silencing phenotypes with two drugs that positively affect transcriptional activity of Da and TCF4 shows that our models are suitable for drug screening. Other drugs could also be tested in the future, for example Nicardipine, since it has been shown to rescue behavioural deficits of PTHS mouse models (Ekins et al., 2020).

5.5 Da has functions in the mature *Drosophila melanogaster*

Da has been mostly studied in the developmental processes, where the functions include regulation of cell cycle, sex determination, endoderm and mesoderm development, myogenesis, salivary gland development, oogenesis and nervous system development (Andrade-Zapata and Baonza, 2014; Bhattacharya and Baker, 2011; Brown et al., 1996; Castanon et al., 2001; Caudy et al., 1988b; Chandrasekaran and Beckendorf, 2003; Cline, 1988; Smith et al., 2002; Tepass and Hartenstein, 1995; Wong et al., 2008). There is some evidence that Da is important for the functioning of larval nervous system where it regulates axonal branching, synaptogenesis and locomotion (D'Rozario et al., 2016; Waddell et al., 2019). In the larval brain, Da carries out the functions at least in part through regulating the expression of synaptic proteins Bruchpilot and Nrx-1 (D'Rozario et al., 2016). In this thesis we provide additional evidence that Da functions in the mature *Drosophila melanogaster.* First, we show that Da is expressed in larval and adult brains. In addition, we show that ubiquitous overexpression of Da in the adult flies after eclosion from the pupae causes lethality within days indicating that maintaining appropriate levels of Da is also crucial after development. We also demonstrate that Da has functions in mature neurons. More specifically, Da regulates memory, locomotion, and the expression of synaptic proteins Syn and Dlg1. This is in agreement with previous findings by D'Rozario and colleagues showing that in mature nervous system Da functions in maintaining the synaptic proteome.

5.6 Advancements on the expression studies of *TCF4* are beneficial for developing therapies for *TCF4*-related diseases

In this thesis, in addition to generating and validating the *Drosophila melanogaster* models for PTHS and other *TCF4*-related intellectual disabilities, we also performed a systematic characterization of the expression of *Tcf4* transcripts and TCF4 protein isoforms in mouse and rat which are commonly used as models for human diseases, including *TCF4*-related diseases (Cleary et al., 2021; Kennedy et al., 2016; Thaxton et al., 2018). *TCF4* gene structure is very complex, giving rise to numerous transcripts and protein isoforms (Nurm et al., 2021; Sepp et al., 2011). Charting the tissue specific levels of *TCF4* expression and isoform distribution is the basis for developing gene therapies for *TCF4*-associated diseases.

In accordance with previously published data, the expression of *Tcf4* was highest in the rodent hippocampus and cerebral cortex around birth and in the cerebellum in about a week after birth (Chen et al., 2016; Jung et al., 2018; Mesman et al., 2020; Soosaar et al., 1994). Although *Tcf4* gene structure allows for numerous transcripts to be expressed, in rodent brain the main transcripts were encoding isoforms TCF4-A, TCF4-B, TCF4-D, TCF4-C and TCF4-I in the decending expression level order. In the rat brain a bigger variety of transcripts were expressed than in the mouse brain. *Tcf4* was also expressed in non-neural tissues although at lower levels than in the brain. Of the analyzed non-neural tissues, the highest TCF4 protein expression was seen in the rodent skin, lung, heart, diafragm, skeletal muscle, bladder and thymus. In most mouse non-neural tissues, transcripts that code for TCF4-A and TCF4-B were prominently expressed while in the rat the variety was bigger.

We also analysed the expression of *TCF4* transcripts in human. Like in rodents, highest *TCF4* levels were expressed in the brain around birth. The main transcripts in the human brain encode TCF4-A, TCF4-D, TCF4-B, TCF4-C and TCF4-I, although more transcripts

besides these were expressed compared to the rodent brain. *TCF4* was also expressed in the human non-neural tissues although the expression levels were lower than in the brain.

Recently, a pioneering study in PTHS gene therapy was published, where the authors show that the behavioural impairments of PTHS mouse model can be alleviated by genetic reinstation of TCF4 expression in postnatal neurons (Kim et al., 2022). This ground-breaking study shows that upon early diagnosis of PTHS some of the symptoms could be reversable. These experiments are paving the way for potential gene therapy of PTHS patients but for gene therapy it is important to know which transcripts to deliver to the brain of the patient. This is important since different TCF4 isoforms have different transcriptional abilities (Nurm et al., 2021; Sepp et al., 2011; Sepp et al., 2012; Sepp et al., 2017). Our results show that in the brains of rodents and humans transcripts encoding TCF4-A, TCF4-B and TCF4-D protein isoforms are the main *TCF4* transcripts and these could be used for gene therapy of PTHS and *TCF4*-related intellectual disabilities.

The PTHS and other *TCF4*-related intellectual disability models generated and studied in this thesis have provided valuable information about these serious diseases. We have validated our *Drosophila melanogaster* models for drug screening and proposed two drugs for further studies. In addition, we have provided the research community valuable knowledge about tissue specific expression of *TCF4* transcripts and protein isoforms which is needed for developing gene therapy for PTHS.

6 Conclusions

The main findings of this thesis are as follows:

- Da is the ortholog of TCF4 since TCF4 is functional in the neurogenesis of the *Drosophila melanogaster* and can replace Da during embryonic neurogenesis;
- Silencing of *da* in the nervous system causes the impairment of larval appetitive associative learning and negative geotaxis of the adults;
- Synaptic proteins Synapsin and Discs large 1 are regulated by Da;
- *da* silencing phenotypes can be alleviated with resveratrol and SAHA, showing the *Drosophila melanogaster* PTHS models can be used for drug screening;
- In the brains of rodents and humans, transcripts that encode TCF4-A, TCF4-B and TCF4-D are the main transcripts.

References

Adusumilli, L., Facchinello, N., Teh, C., Busolin, G., Le, M. T., Yang, H., Beffagna, G., Campanaro, S., Tam, W. L., Argenton, F., et al. (2020). miR-7 Controls the Dopaminergic/Oligodendroglial Fate through Wnt/β-catenin Signaling Regulation. *Cells* 9, 711.

Albanna, A., Choudhry, Z., Harvey, P.-O., Fathalli, F., Cassidy, C., Sengupta, S. M., Iyer, S. N., Rho, A., Lepage, M., Malla, A., et al. (2014). TCF4 gene polymorphism and cognitive performance in patients with first episode psychosis. *Schizophr Res* 152, 124–129.

Alifragis, P., Poortinga, G., Parkhurst, S. M. and Delidakis, C. (1997). A network of interacting transcriptional regulators involved in Drosophila neural fate specification revealed by the yeast two-hybrid system. *Proc Natl Acad Sci U S A* 94, 13099–13104.

Amare, A. T., Vaez, A., Hsu, Y.-H., Direk, N., Kamali, Z., Howard, D. M., McIntosh, A. M., Tiemeier, H., Bültmann, U., Snieder, H., et al. (2020). Bivariate genome-wide association analyses of the broad depression phenotype combined with major depressive disorder, bipolar disorder or schizophrenia reveal eight novel genetic loci for depression. *Mol Psychiatry* 25, 1420–1429.

Amiel, J., Rio, M., de Pontual, L., Redon, R., Malan, V., Boddaert, N., Plouin, P., Carter, N. P., Lyonnet, S., Munnich, A., et al. (2007). Mutations in TCF4, encoding a class I basic helix-loop-helix transcription factor, are responsible for Pitt-Hopkins syndrome, a severe epileptic encephalopathy associated with autonomic dysfunction. *Am. J. Hum. Genet.* 80, 988–993.

Andrade-Zapata, I. and Baonza, A. (2014). The bHLH factors extramacrochaetae and daughterless control cell cycle in Drosophila imaginal discs through the transcriptional regulation of the Cdc25 phosphatase string. *PLoS Genet.* 10, e1004233.

Appaiah, H., Bhat-Nakshatri, P., Mehta, R., Thorat, M., Badve, S. and Nakshatri, H. (2010). ITF2 is a target of CXCR4 in MDA-MB-231 breast cancer cells and is associated with reduced survival in estrogen receptor-negative breast cancer. *Cancer Biol Ther* 10, 600–614.

Badowska, D. M., Brzózka, M. M., Kannaiyan, N., Thomas, C., Dibaj, P., Chowdhury, A., Steffens, H., Turck, C. W., Falkai, P., Schmitt, A., et al. (2020). Modulation of cognition and neuronal plasticity in gain- and loss-of-function mouse models of the schizophrenia risk gene Tcf4. *Transl Psychiatry* 10, 343.

Baratz, K. H., Tosakulwong, N., Ryu, E., Brown, W. L., Branham, K., Chen, W., Tran, K. D., Schmid-Kubista, K. E., Heckenlively, J. R., Swaroop, A., et al. (2010). E2-2 protein and Fuchs's corneal dystrophy. *N Engl J Med* 363, 1016–1024.

Bardin, A. J., Perdigoto, C. N., Southall, T. D., Brand, A. H. and Schweisguth, F. (2010). Transcriptional control of stem cell maintenance in the Drosophila intestine. *Development* 137, 705–714.

Bedeschi, M. F., Marangi, G., Calvello, M. R., Ricciardi, S., Leone, F. P. C., Baccarin, M., Guerneri, S., Orteschi, D., Murdolo, M., Lattante, S., et al. (2017). Impairment of different protein domains causes variable clinical presentation within Pitt-Hopkins syndrome and suggests intragenic molecular syndromology of TCF4. *Eur J Med Genet* 60, 565–571.

Bellen, H. J., Kooyer, S., D'Evelyn, D. and Pearlman, J. (1992). The Drosophila couch potato protein is expressed in nuclei of peripheral neuronal precursors and shows homology to RNA-binding proteins. *Genes Dev* 6, 2125–2136.

Bergqvist, I., Eriksson, M., Saarikettu, J., Eriksson, B., Corneliussen, B., Grundström, T. and Holmberg, D. (2000). The basic helix-loop-helix transcription factor E2-2 is involved in T lymphocyte development. *Eur J Immunol* 30, 2857–2863.

Bhattacharya, A. and Baker, N. E. (2011). A network of broadly expressed HLH genes regulates tissue-specific cell fates. *Cell* 147, 881–892.

Bhattacharya, A., Li, K., Quiquand, M., Rimesso, G. and Baker, N. E. (2017). The Notch pathway regulates the Second Mitotic Wave cell cycle independently of bHLH proteins. *Dev Biol* 431, 309–320.

Bouderlique, T., Peña-Pérez, L., Kharazi, S., Hils, M., Li, X., Krstic, A., De Paepe, A., Schachtrup, C., Gustafsson, C., Holmberg, D., et al. (2019). The Concerted Action of E2-2 and HEB Is Critical for Early Lymphoid Specification. *Front Immunol* 10, 455.

Brand, M., Jarman, A. P., Jan, L. Y. and Jan, Y. N. (1993). asense is a Drosophila neural precursor gene and is capable of initiating sense organ formation. *Development* 119, 1–17.

Braun, K., Häberle, B. M., Wittmann, M.-T. and Lie, D. C. (2020). Enriched environment ameliorates adult hippocampal neurogenesis deficits in Tcf4 haploinsufficient mice. *BMC Neurosci* 21, 50.

Brockschmidt, A., Todt, U., Ryu, S., Hoischen, A., Landwehr, C., Birnbaum, S., Frenck, W., Radlwimmer, B., Lichter, P., Engels, H., et al. (2007). Severe mental retardation with breathing abnormalities (Pitt-Hopkins syndrome) is caused by haploinsufficiency of the neuronal bHLH transcription factor TCF4. *Hum. Mol. Genet.* 16, 1488–1494.

Brown, N. L., Paddock, S. W., Sattler, C. A., Cronmiller, C., Thomas, B. J. and Carroll, S. B. (1996). daughterless is required for Drosophila photoreceptor cell determination, eye morphogenesis, and cell cycle progression. *Dev. Biol.* 179, 65–78.

Brzózka, M. M. and Rossner, M. J. (2013). Deficits in trace fear memory in a mouse model of the schizophrenia risk gene TCF4. *Behav Brain Res* 237, 348–356.

Brzózka, M. M., Radyushkin, K., Wichert, S. P., Ehrenreich, H. and Rossner, M. J. (2010). Cognitive and Sensorimotor Gating Impairments in Transgenic Mice Overexpressing the Schizophrenia Susceptibility Gene Tcf4 in the Brain. *Biological Psychiatry* 68, 33–40.

Cabrera, C. V. and Alonso, M. C. (1991). Transcriptional activation by heterodimers of the achaete-scute and daughterless gene products of Drosophila. *EMBO J* 10, 2965–2973.

Cabrera, C. V., Alonso, M. C. and Huikeshoven, H. (1994). Regulation of scute function by extramacrochaete in vitro and in vivo. *Development* 120, 3595–3603.

Cadigan, K. M., Jou, A. D. and Nusse, R. (2002). Wingless blocks bristle formation and morphogenetic furrow progression in the eye through repression of Daughterless. *Development* 129, 3393–3402.

Cai, L., Bao, Y., Fu, X., Cao, H., Baranova, A., Zhang, X., Sun, J. and Zhang, F. (2021). Causal links between major depressive disorder and insomnia: A Mendelian randomisation study. *Gene* 768, 145271.

Campos-Ortega, J. A. and Knust, E. (1990). Molecular analysis of a cellular decision during embryonic development of Drosophila melanogaster: epidermogenesis or neurogenesis. *Eur J Biochem* 190, 1–10.

Cano, A. and Portillo, F. (2010). An emerging role for class I bHLH E2-2 proteins in EMT regulation and tumor progression. *Cell Adh Migr* 4, 56–60.

Care, M. A., Barrans, S., Worrillow, L., Jack, A., Westhead, D. R. and Tooze, R. M. (2013). A Microarray Platform-Independent Classification Tool for Cell of Origin Class Allows Comparative Analysis of Gene Expression in Diffuse Large B-cell Lymphoma. *PLoS One* 8, e55895.

Castanon, I., Von Stetina, S., Kass, J. and Baylies, M. K. (2001). Dimerization partners determine the activity of the Twist bHLH protein during Drosophila mesoderm development. *Development* 128, 3145–3159.

Caudy, M., Grell, E. H., Dambly-Chaudière, C., Ghysen, A., Jan, L. Y. and Jan, Y. N. (1988a). The maternal sex determination gene daughterless has zygotic activity necessary for the formation of peripheral neurons in Drosophila. *Genes Dev.* 2, 843–852.

Caudy, M., Vässin, H., Brand, M., Tuma, R., Jan, L. Y. and Jan, Y. N. (1988b). daughterless, a Drosophila gene essential for both neurogenesis and sex determination, has sequence similarities to myc and the achaete-scute complex. *Cell* 55, 1061–1067.

Cave, J. W., Loh, F., Surpris, J. W., Xia, L. and Caudy, M. A. (2005). A DNA Transcription Code for Cell-Specific Gene Activation by Notch Signaling. *Current Biology* 15, 94–104.

Cave, J. W., Xia, L. and Caudy, M. A. (2009). The Daughterless N-terminus directly mediates synergistic interactions with Notch transcription complexes via the SPS+A DNA transcription code. *BMC Res Notes* 2, 65.

Ceribelli, M., Hou, Z. E., Kelly, P. N., Huang, D. W., Wright, G., Ganapathi, K., Evbuomwan, M. O., Pittaluga, S., Shaffer, A. L., Marcucci, G., et al. (2016). A Druggable TCF4- and BRD4-Dependent Transcriptional Network Sustains Malignancy in Blastic Plasmacytoid Dendritic Cell Neoplasm. *Cancer Cell* 30, 764–778.

Chandrasekaran, V. and Beckendorf, S. K. (2003). senseless is necessary for the survival of embryonic salivary glands in Drosophila. *Development* 130, 4719–4728.

Chen, C.-K. and Chien, C.-T. (1999). Negative regulation of atonal in proneural cluster formation of Drosophila R8 photoreceptors. *Proc Natl Acad Sci U S A* 96, 5055–5060.

Chen, B. and Lim, R. W. (1997). Physical and functional interactions between the transcriptional inhibitors Id3 and ITF-2b. Evidence toward a novel mechanism regulating muscle-specific gene expression. *J Biol Chem* 272, 2459–2463.

Chen, T., Wu, Q., Zhang, Y., Lu, T., Yue, W. and Zhang, D. (2016). Tcf4 Controls Neuronal Migration of the Cerebral Cortex through Regulation of Bmp7. *Front Mol Neurosci* 9, 94.

Chen, H.-Y., Bohlen, J. F. and Maher, B. J. (2021). Molecular and Cellular Function of Transcription Factor 4 in Pitt-Hopkins Syndrome. *Dev Neurosci* 43, 159–167.

Chiaramello, A., Soosaar, A., Neuman, T. and Zuber, M. X. (1995). Differential expression and distinct DNA-binding specificity of ME1a and ME2 suggest a unique role during differentiation and neuronal plasticity. *Molecular Brain Research* 29, 107–118.

Cisse, B., Caton, M. L., Lehner, M., Maeda, T., Scheu, S., Locksley, R., Holmberg, D., Zweier, C., den Hollander, N. S., Kant, S. G., et al. (2008). Transcription factor E2-2 is an essential and specific regulator of plasmacytoid dendritic cell development. *Cell* 135, 37–48.

Cleary, C. M., James, S., Maher, B. J. and Mulkey, D. K. (2021). Disordered breathing in a Pitt-Hopkins syndrome model involves Phox2b-expressing parafacial neurons and aberrant Nav1.8 expression. *Nat Commun* 12, 5962.

Cline, T. W. (1976). A Sex-Specific, Temperature-Sensitive Maternal Effect of the Daughterless Mutation of Drosophila Melanogaster. *Genetics* 84, 723–742.

Cline, T. W. (1988). Evidence that sisterless-a and sisterless-b are two of several discrete "numerator elements" of the X/A sex determination signal in Drosophila that switch Sxl between two alternative stable expression states. *Genetics* 119, 829–862.

Coll-Tané, M., Krebbers, A., Castells-Nobau, A., Zweier, C. and Schenck, A. (2019). Intellectual disability and autism spectrum disorders 'on the fly': insights from Drosophila. *Disease Models & Mechanisms* 12, dmm039180.

Corneliussen, B., Thornell, A., Hallberg, B. and Grundström, T. (1991). Helix-loop-helix transcriptional activators bind to a sequence in glucocorticoid response elements of retrovirus enhancers. *J Virol* 65, 6084–6093.

Cronmiller, C. and Cline, T. W. (1987). The Drosophila sex determination gene daughterless has different functions in the germ line versus the soma. *Cell* 48, 479–487.

Cronmiller, C. and Cummings, C. A. (1993). The daughterless gene product in Drosophila is a nuclear protein that is broadly expressed throughout the organism during development. *Mechanisms of Development* 42, 159–169.

Cronmiller, C., Schedl, P. and Cline, T. W. (1988). Molecular characterization of daughterless, a Drosophila sex determination gene with multiple roles in development. *Genes Dev* 2, 1666–1676.

Crux, S., Herms, J. and Dorostkar, M. M. (2018). Tcf4 regulates dendritic spine density and morphology in the adult brain. *PLoS One* 13,.

Cummings, C. A. and Cronmiller, C. (1994). The daughterless gene functions together with Notch and Delta in the control of ovarian follicle development in Drosophila. *Development* 120, 381–394.

de Pontual, L., Mathieu, Y., Golzio, C., Rio, M., Malan, V., Boddaert, N., Soufflet, C., Picard, C., Durandy, A., Dobbie, A., et al. (2009). Mutational, functional, and expression studies of the TCF4 gene in Pitt-Hopkins syndrome. *Human Mutation* 30, 669–676.

Deshpande, G., Stukey, J. and Schedl, P. (1995). scute (sis-b) function in Drosophila sex determination. *Mol Cell Biol* 15, 4430–4440.

DiStefano, G. M., Gangemi, A. J., Khandelwal, P. J., Saunders, A. J. and Marenda, D. R. (2012). Drosophila lilliputian is required for proneural gene expression in retinal development. *Dev Dyn* 241, 553–562.

Doostparast Torshizi, A., Armoskus, C., Zhang, H., Forrest, M. P., Zhang, S., Souaiaia, T., Evgrafov, O. V., Knowles, J. A., Duan, J. and Wang, K. (2019). Deconvolution of transcriptional networks identifies TCF4 as a master regulator in schizophrenia. *Sci Adv* 5, eaau4139.

D'Rozario, M., Zhang, T., Waddell, E. A., Zhang, Y., Sahin, C., Sharoni, M., Hu, T., Nayal, M., Kutty, K., Liebl, F., et al. (2016). Type I bHLH proteins Daughterless and TCF4 restrict neurite branching and synapse formation by repressing Neurexin in postmitotic neurons. *Cell Rep* 15, 386–397.

Dunn, L. L., Sekyere, E. O., Suryo Rahmanto, Y. and Richardson, D. R. (2006). The function of melanotransferrin: a role in melanoma cell proliferation and tumorigenesis. *Carcinogenesis* 27, 2157–2169.

Ecovoiu, A. Al., Ratiu, A. C., Micheu, M. M. and Chifiriuc, M. C. (2022). Inter-Species Rescue of Mutant Phenotype—The Standard for Genetic Analysis of Human Genetic Disorders in Drosophila melanogaster Model. *Int J Mol Sci* 23, 2613.

Ekins, S., Puhl, A. C. and Davidow, A. (2020). Repurposing the Dihydropyridine Calcium Channel Inhibitor Nicardipine as a Nav1.8 Inhibitor In Vivo for Pitt Hopkins Syndrome. *Pharm Res* 37, 127.

Ellis, H. M., Spann, D. R. and Posakony, J. W. (1990). extramacrochaetae, a negative regulator of sensory organ development in Drosophila, defines a new class of helix-loop-helix proteins. *Cell* 61, 27–38.

Escudero, L. M., Caminero, E., Schulze, K. L., Bellen, H. J. and Modolell, J. (2005). Charlatan, a Zn-finger transcription factor, establishes a novel level of regulation of the proneural achaete/scute genes of Drosophila. *Development* 132, 1211–1222.

Esvald, E.-E., Tuvikene, J., Moistus, A., Rannaste, K., Kõomägi, S. and Timmusk, T. (2022). Differential regulation of the BDNF gene in cortical and hippocampal neurons. *J Neurosci* JN-RM-2535-21.

Fautsch, M. P., Wieben, E. D., Baratz, K. H., Bhattacharyya, N., Sadan, A. N., Hafford-Tear, N. J., Tuft, S. J. and Davidson, A. E. (2021). TCF4-mediated Fuchs endothelial corneal dystrophy: Insights into a common trinucleotide repeat-associated disease. *Prog Retin Eye Res* 81, 100883.

Fenckova, M., Blok, L. E. R., Asztalos, L., Goodman, D. P., Cizek, P., Singgih, E. L., Glennon, J. C., IntHout, J., Zweier, C., Eichler, E. E., et al. (2019). Habituation Learning Is a Widely Affected Mechanism in Drosophila Models of Intellectual Disability and Autism Spectrum Disorders. *Biological Psychiatry* 86, 294–305.

Flora, A., Garcia, J. J., Thaller, C. and Zoghbi, H. Y. (2007). The E-protein Tcf4 interacts with Math1 to regulate differentiation of a specific subset of neuronal progenitors. *Proc Natl Acad Sci U S A* 104, 15382–15387.

Foja, S., Luther, M., Hoffmann, K., Rupprecht, A. and Gruenauer-Kloevekorn, C. (2017). CTG18.1 repeat expansion may reduce TCF4 gene expression in corneal endothelial cells of German patients with Fuchs' dystrophy. *Graefes Arch Clin Exp Ophthalmol* 255, 1621–1631.

Forrest, M., Chapman, R. M., Doyle, A. M., Tinsley, C. L., Waite, A. and Blake, D. J. (2012). Functional analysis of TCF4 missense mutations that cause Pitt-Hopkins syndrome. *Hum. Mutat.* 33, 1676–1686.

Forrest, M. P., Waite, A. J., Martin-Rendon, E. and Blake, D. J. (2013). Knockdown of human TCF4 affects multiple signaling pathways involved in cell survival, epithelial to mesenchymal transition and neuronal differentiation. *PLoS ONE* 8, e73169.

Forrest, M. P., Hill, M. J., Kavanagh, D. H., Tansey, K. E., Waite, A. J. and Blake, D. J. (2018). The Psychiatric Risk Gene Transcription Factor 4 (TCF4) Regulates Neurodevelopmental Pathways Associated With Schizophrenia, Autism, and Intellectual Disability. *Schizophr Bull* 44, 1100–1110.

Fu, H., Cai, J., Clevers, H., Fast, E., Gray, S., Greenberg, R., Jain, M. K., Ma, Q., Qiu, M., Rowitch, D. H., et al. (2009). A Genome-Wide Screen for Spatially Restricted Expression Patterns Identifies Transcription Factors That Regulate Glial Development. *J Neurosci* 29, 11399–11408.

Furumura, M., Potterf, S. B., Toyofuku, K., Matsunaga, J., Muller, J. and Hearing, V. J. (2001). Involvement of ITF2 in the transcriptional regulation of melanogenic genes. *J Biol Chem* 276, 28147–28154.

Fuse, N., Hirose, S. and Hayashi, S. (1994). Diploidy of Drosophila imaginal cells is maintained by a transcriptional repressor encoded by escargot. *Genes Dev* 8, 2270–2281.

Gelernter, J., Sun, N., Polimanti, R., Pietrzak, R., Levey, D. F., Bryois, J., Lu, Q., Hu, Y., Li, B., Radhakrishnan, K., et al. (2019). Genome-wide association study of post-traumatic stress disorder reexperiencing symptoms in >165,000 US veterans. *Nat Neurosci* 22, 1394–1401.

Ghosh, H. S., Cisse, B., Bunin, A., Lewis, K. L. and Reizis, B. (2010). Continuous expression of the transcription factor e2-2 maintains the cell fate of mature plasmacytoid dendritic cells. *Immunity* 33, 905–916.

Giebel, B., Stüttem, I., Hinz, U. and Campos-Ortega, J. A. (1997). Lethal of Scute requires overexpression of Daughterless to elicit ectopic neuronal development during embryogenesis in Drosophila. *Mechanisms of Development* 63, 75–87.

Gloury, R., Zotos, D., Zuidscherwoude, M., Masson, F., Liao, Y., Hasbold, J., Corcoran, L. M., Hodgkin, P. D., Belz, G. T., Shi, W., et al. (2016). Dynamic changes in Id3 and E-protein activity orchestrate germinal center and plasma cell development. *J Exp Med* 213, 1095–1111.

Grajkowska, L. T., Ceribelli, M., Lau, C. M., Warren, M. E., Tiniakou, I., Higa, S. N., Bunin, A., Haecker, H., Mirny, L. A., Staudt, L. M., et al. (2017). Isoform-specific expression and feedback regulation of E protein TCF4 control dendritic cell lineage specification. *Immunity* 46, 65–77.

Greb-Markiewicz, B., Kazana, W., Zarębski, M. and Ożyhar, A. (2019). The subcellular localization of bHLH transcription factor TCF4 is mediated by multiple nuclear localization and nuclear export signals. *Sci Rep* 9, 15629.

Hade, A.-C., Philips, M.-A., Reimann, E., Jagomäe, T., Eskla, K.-L., Traks, T., Prans, E., Kõks, S., Vasar, E. and Väli, M. (2021). Chronic Alcohol Use Induces Molecular Genetic Changes in the Dorsomedial Thalamus of People with Alcohol-Related Disorders. *Brain Sci* 11, 435.

Haines, J. D., Fang, J., Mushynski, W. E. and Almazan, G. (2010). Mitogen-activated protein kinase activated protein kinase 2 (MK2) participates in p38 MAPK regulated control of oligodendrocyte differentiation. *Glia* 58, 1384–1393.

Hall, M.-H., Levy, D. L., Salisbury, D. F., Haddad, S., Gallagher, P., Lohan, M., Cohen, B., Ongür, D. and Smoller, J. W. (2014). Neurophysiologic effect of GWAS derived schizophrenia and bipolar risk variants. *Am J Med Genet B Neuropsychiatr Genet* 165B, 9–18.

Han, S. M., Na, H. Y., Ham, O., Choi, W., Sohn, M., Ryu, S. H., In, H., Hwang, K.-C. and Park, C. G. (2016). TCF4-Targeting miR-124 is Differentially Expressed amongst Dendritic Cell Subsets. *Immune Netw* 16, 61–74.

He, Q., Liu, L., Wang, Y., Xu, C., Xu, M., Fu, J., Zhu, J., Zhao, B., Ni, C., Yao, M., et al. (2022). miR-155-5p in the spinal cord regulates hypersensitivity in a rat model of bone cancer pain. *Mol Pain* 18, 17448069221127812.

Hennig, K. M., Fass, D. M., Zhao, W.-N., Sheridan, S. D., Fu, T., Erdin, S., Stortchevoi, A., Lucente, D., Cody, J. D., Sweetser, D., et al. (2017). WNT/β-Catenin Pathway and Epigenetic Mechanisms Regulate the Pitt-Hopkins Syndrome and Schizophrenia Risk Gene TCF4. *Mol Neuropsychiatry* 3, 53–71.

Herbst, A., Bommer, G. T., Kriegl, L., Jung, A., Behrens, A., Csanadi, E., Gerhard, M., Bolz, C., Riesenberg, R., Zimmermann, W., et al. (2009a). ITF-2 Is Disrupted via Allelic Loss of Chromosome 18q21, and ITF-2B Expression Is Lost at the Adenoma-Carcinoma Transition. *Gastroenterology* 137, 639-648.e9.

Herbst, A., Helferich, S., Behrens, A., Göke, B. and Kolligs, F. T. (2009b). The transcription factor ITF-2A induces cell cycle arrest via p21(Cip1). *Biochem Biophys Res Commun* 387, 736–740.

Hill, M. J., Killick, R., Navarrete, K., Maruszak, A., McLaughlin, G. M., Williams, B. P. and Bray, N. J. (2017). Knockdown of the schizophrenia susceptibility gene TCF4 alters gene expression and proliferation of progenitor cells from the developing human neocortex. *J Psychiatry Neurosci* 42, 181–188.

Hoshijima, K., Kohyama, A., Watakabe, I., Inoue, K., Sakamoto, H. and Shimura, Y. (1995). Transcriptional regulation of the Sex-lethal gene by helix-loop-helix proteins. *Nucleic Acids Res* 23, 3441–3448.

Hur, E.-H., Goo, B.-K., Moon, J., Choi, Y., Hwang, J. J., Kim, C.-S., Bae, K. S., Choi, J., Cho, S. Y., Yang, S.-H., et al. (2017). Induction of immunoglobulin transcription factor 2 and resistance to MEK inhibitor in melanoma cells. *Oncotarget* 8, 41387–41400.

Ibnsouda, S., Schweisguth, F., Jullien, D., Kücherer, C., Lepesant, J. A. and Vincent, A. (1995). Evolutionarily conserved positive and negative cis-acting elements control the blastoderm-specific expression of the Drosophila serendipity alpha cellularisation gene. *Mech Dev* 49, 71–82.

in ’t Hout, F. E. M., van der Reijden, B. A., Monteferrario, D., Jansen, J. H. and Huls, G. (2014). High expression of transcription factor 4 (TCF4) is an independent adverse prognostic factor in acute myeloid leukemia that could guide treatment decisions. *Haematologica* 99, e257–e259.

In ’t Hout, F. E. M., van Duren, J., Monteferrario, D., Brinkhuis, E., Mariani, N., Westers, T. M., Chitu, D., Nikoloski, G., van de Loosdrecht, A. A., van der Reijden, B. A., et al. (2019). TCF4 promotes erythroid development. *Exp Hematol* 69, 17-21.e1.

Ishibashi, T., Hatori, R., Maeda, R., Nakamura, M., Taguchi, T., Matsuyama, Y. and Matsuno, K. (2019). E and ID proteins regulate cell chirality and left–right asymmetric development in Drosophila. *Genes to Cells* 24, 214–230.

Itahana, Y., Piens, M., Fong, S., Singh, J., Sumida, T. and Desprez, P.-Y. (2008). Expression of Id and ITF-2 genes in the mammary gland during pregnancy. *Biochem Biophys Res Commun* 372, 826–830.

Jafar-Nejad, H., Tien, A.-C., Acar, M. and Bellen, H. J. (2006). Senseless and Daughterless confer neuronal identity to epithelial cells in the Drosophila wing margin. *Development* 133, 1683–1692.

Jain, N., Hartert, K., Tadros, S., Fiskus, W., Havranek, O., Ma, M. C. J., Bouska, A., Heavican, T., Kumar, D., Deng, Q., et al. (2019). Targetable genetic alterations of TCF4 (E2-2) drive immunoglobulin expression in diffuse large B cell lymphoma. *Sci Transl Med* 11, eaav5599.

Jung, M., Häberle, B. M., Tschaikowsky, T., Wittmann, M.-T., Balta, E.-A., Stadler, V.-C., Zweier, C., Dörfler, A., Gloeckner, C. J. and Lie, D. C. (2018). Analysis of the expression pattern of the schizophrenia-risk and intellectual disability gene TCF4 in the developing and adult brain suggests a role in development and plasticity of cortical and hippocampal neurons. *Mol Autism* 9,.

Kalscheuer, V. M., Feenstra, I., Van Ravenswaaij-Arts, C. M. A., Smeets, D. F. C. M., Menzel, C., Ullmann, R., Musante, L. and Ropers, H.-H. (2008). Disruption of the TCF4 gene in a girl with mental retardation but without the classical Pitt-Hopkins syndrome. *Am J Med Genet A* 146A, 2053–2059.

Kantzer, C. G., Yang, W., Grommisch, D., Patil, K. V., Mak, K. H.-M., Shirokova, V. and Genander, M. (2022). ID1 and CEBPA coordinate epidermal progenitor cell differentiation. *Development* dev.201262.

Kasai, Y., Stahl, S. and Crews, S. (1998). Specification of the Drosophila CNS midline cell lineage: direct control of single-minded transcription by dorsal/ventral patterning genes. *Gene Expr* 7, 171–189.

Kennedy, A. J., Rahn, E. J., Paulukaitis, B. S., Savell, K. E., Kordasiewicz, H. B., Wang, J., Lewis, J. W., Posey, J., Strange, S. K., Guzman-Karlsson, M. C., et al. (2016). Tcf4 Regulates Synaptic Plasticity, DNA Methylation, and Memory Function. *Cell Rep* 16, 2666–2685.

Kharbanda, M., Kannike, K., Lampe, A., Berg, J., Timmusk, T. and Sepp, M. (2016). Partial deletion of TCF4 in three generation family with non-syndromic intellectual disability, without features of Pitt-Hopkins syndrome. *Eur J Med Genet* 59, 310–314.

Kim, S.-K., Jang, H.-R., Kim, J.-H., Kim, M., Noh, S.-M., Song, K.-S., Kang, G. H., Kim, H. J., Kim, S.-Y., Yoo, H.-S., et al. (2008). CpG methylation in exon 1 of transcription factor 4 increases with age in normal gastric mucosa and is associated with gene silencing in intestinal-type gastric cancers. *Carcinogenesis* 29, 1623–1631.

Kim, H., Berens, N. C., Ochandarena, N. E. and Philpot, B. D. (2020). Region and Cell Type Distribution of TCF4 in the Postnatal Mouse Brain. *Front Neuroanat* 14, 42.

Kim, H., Gao, E. B., Draper, A., Berens, N. C., Vihma, H., Zhang, X., Higashi-Howard, A., Ritola, K. D., Simon, J. M., Kennedy, A. J., et al. (2022). Rescue of behavioral and electrophysiological phenotypes in a Pitt-Hopkins syndrome mouse model by genetic restoration of Tcf4 expression. *Elife* 11, e72290.

King-Jones, K., Korge, G. and Lehmann, M. (1999). The helix-loop-helix proteins dAP-4 and daughterless bind both in vitro and in vivo to SEBP3 sites required for transcriptional activation of the Drosophila gene Sgs-4. *J. Mol. Biol.* 291, 71–82.

Kiparaki, M., Zarifi, I. and Delidakis, C. (2015). bHLH proteins involved in Drosophila neurogenesis are mutually regulated at the level of stability. *Nucleic Acids Res* 43, 2543–2559.

Kolligs, F. T., Nieman, M. T., Winer, I., Hu, G., Van Mater, D., Feng, Y., Smith, I. M., Wu, R., Zhai, Y., Cho, K. R., et al. (2002). ITF-2, a downstream target of the Wnt/TCF pathway, is activated in human cancers with beta-catenin defects and promotes neoplastic transformation. *Cancer Cell* 1, 145–155.

Kousoulidou, L., Tanteles, G., Moutafi, M., Sismani, C., Patsalis, P. C. and Anastasiadou, V. (2013). 263.4 kb deletion within the TCF4 gene consistent with Pitt-Hopkins syndrome, inherited from a mosaic parent with normal phenotype. *Eur J Med Genet* 56, 314–318.

Kovalick, G. E. and Beckingham, K. (1992). Calmodulin transcription is limited to the nervous system during Drosophila embryogenesis. *Dev Biol* 150, 33–46.

Kunisch, M., Haenlin, M. and Campos-Ortega, J. A. (1994). Lateral inhibition mediated by the Drosophila neurogenic gene delta is enhanced by proneural proteins. *Proc Natl Acad Sci U S A* 91, 10139–10143.

Kwon, E., Wang, W. and Tsai, L.-H. (2013). Validation of schizophrenia-associated genes CSMD1, C10orf26, CACNA1C and TCF4 as miR-137 targets. *Mol Psychiatry* 18, 11–12.

Lan, Q., Cao, M., Kollipara, R. K., Rosa, J. B., Kittler, R. and Jiang, H. (2018). FoxA transcription factor Fork head maintains the intestinal stem/progenitor cell identities in Drosophila. *Developmental Biology* 433, 324–343.

Lasko, P. and Lüthy, K. (2021). Investigating rare and ultrarare epilepsy syndromes with Drosophila models. *Fac Rev* 10, 10.

Lassar, A. B., Davis, R. L., Wright, W. E., Kadesch, T., Murre, C., Voronova, A., Baltimore, D. and Weintraub, H. (1991). Functional activity of myogenic HLH proteins requires hetero-oligomerization with E12/E47-like proteins in vivo. *Cell* 66, 305–315.

Lennertz, L., Rujescu, D., Wagner, M., Frommann, I., Schulze-Rauschenbach, S., Schuhmacher, A., Landsberg, M. W., Franke, P., Möller, H.-J., Wölwer, W., et al. (2011a). Novel schizophrenia risk gene TCF4 influences verbal learning and memory functioning in schizophrenia patients. *Neuropsychobiology* 63, 131–136.

Lennertz, L., Quednow, B. B., Benninghoff, J., Wagner, M., Maier, W. and Mössner, R. (2011b). Impact of TCF4 on the genetics of schizophrenia. *Eur Arch Psychiatry Clin Neurosci* 261 Suppl 2, S161-165.

Li, K. and Baker, N. E. (2018). Regulation of the Drosophila ID protein Extra macrochaetae by proneural dimerization partners. *eLife*.

Li, K. and Baker, N. E. (2019). Transcriptional and post-transcriptional regulation of extra macrochaetae during Drosophila adult peripheral neurogenesis. *Developmental Biology* 449, 41–51.

Li, T., Li, Z., Chen, P., Zhao, Q., Wang, T., Huang, K., Li, J., Li, Y., Liu, J., Zeng, Z., et al. (2010). Common variants in major histocompatibility complex region and TCF4 gene are significantly associated with schizophrenia in Han Chinese. *Biol Psychiatry* 68, 671–673.

Li, G., Luna, C., Qiu, J., Epstein, D. L. and Gonzalez, P. (2011). Role of miR-204 in the Regulation of Apoptosis, Endoplasmic Reticulum Stress Response, and Inflammation in Human Trabecular Meshwork Cells. *Invest Ophthalmol Vis Sci* 52, 2999–3007.

Li, H., Zhu, Y., Morozov, Y. M., Chen, X., Page, S. C., Rannals, M. D., Maher, B. J. and Rakic, P. (2019). Disruption of TCF4 regulatory networks leads to abnormal cortical development and mental disabilities. *Mol. Psychiatry*.

Li, N., Liu, B., Wu, W., Hong, Y., Zhang, J., Liu, Y., Zhang, M., Zhang, X. and Duan, G. (2020). Upregulation of transcription factor 4 downregulates NaV1.8 expression in DRG neurons and prevents the development of rat inflammatory and neuropathic hypersensitivity. *Exp Neurol* 327, 113240.

Lim, J., Jafar-Nejad, H., Hsu, Y.-C. and Choi, K.-W. (2008). Novel function of the class I bHLH protein Daughterless in the negative regulation of proneural gene expression in the Drosophila eye. *EMBO Rep* 9, 1128–1133.

Luo, C., Lim, J.-H., Lee, Y., Granter, S. R., Thomas, A., Vazquez, F., Widlund, H. R. and Puigserver, P. (2016). A PGC1α-mediated transcriptional axis suppresses melanoma metastasis. *Nature* 537, 422–426.

Ma, C., Gu, C., Huo, Y., Li, X. and Luo, X.-J. (2018). The integrated landscape of causal genes and pathways in schizophrenia. *Transl Psychiatry* 8, 67.

Maduro, V., Pusey, B. N., Cherukuri, P. F., Atkins, P., du Souich, C., Rupps, R., Limbos, M., Adams, D. R., Bhatt, S. S., Eydoux, P., et al. (2016). Complex translocation disrupting TCF4 and altering TCF4 isoform expression segregates as mild autosomal dominant intellectual disability. *Orphanet J Rare Dis* 11, 62.

Markus, M., Du, Z. and Benezra, R. (2002). Enhancer-specific Modulation of E Protein Activity *. *Journal of Biological Chemistry* 277, 6469–6477.

Martinowich, K., Das, D., Sripathy, S. R., Mai, Y., Kenney, R. F. and Maher, B. J. (2022). Evaluation of Nav1.8 as a therapeutic target for Pitt Hopkins Syndrome. *Mol Psychiatry*.

Massari, M. E. and Murre, C. (2000). Helix-Loop-Helix Proteins: Regulators of Transcription in Eucaryotic Organisms. *Mol Cell Biol* 20, 429–440.

Massari, M. E., Jennings, P. A. and Murre, C. (1996). The AD1 transactivation domain of E2A contains a highly conserved helix which is required for its activity in both Saccharomyces cerevisiae and mammalian cells. *Mol Cell Biol* 16, 121–129.

Masson, J., Pons, L., Busa, T., Missirian, C., Lines, M., Tevissen, H., Diguet, F., Rollat-Farnier, P.-A., Lesca, G., Sanlaville, D., et al. (2022). Disruption and deletion of the proximal part of TCF4 are associated with mild intellectual disability: About three new patients. *Eur J Med Genet* 65, 104458.

Meert, L., Birkhoff, J. C., Conidi, A., Poot, R. A. and Huylebroeck, D. (2022). Different E-box binding transcription factors, similar neuro-developmental defects: ZEB2 (Mowat-Wilson syndrome) and TCF4 (Pitt-Hopkins syndrome). *Rare Disease and Orphan Drugs Journal* 1, 8.

Melicharek, D., Shah, A., DiStefano, G., Gangemi, A. J., Orapallo, A., Vrailas-Mortimer, A. D. and Marenda, D. R. (2008). Identification of Novel Regulators of atonal Expression in the Developing Drosophila Retina. *Genetics* 180, 2095–2110.

Mesman, S., Bakker, R. and Smidt, M. P. (2020). Tcf4 is required for correct brain development during embryogenesis. *Mol Cell Neurosci* 106, 103502.

Mok, L.-P., Qin, T., Bardot, B., LeComte, M., Homayouni, A., Ahimou, F. and Wesley, C. (2005). Delta activity independent of its activity as a ligand of Notch. *BMC Dev Biol* 5, 6.

Mokhtari, M. A., Sargazi, S., Saravani, R., Heidari Nia, M., Mirinejad, S., Hadzsiev, K., Bene, J. and Shakiba, M. (2022). Genetic Polymorphisms in miR-137 and Its Target Genes, TCF4 and CACNA1C, Contribute to the Risk of Bipolar Disorder: A Preliminary Case-Control Study and Bioinformatics Analysis. *Dis Markers* 2022, 1886658.

Mossakowska-Wójcik, J., Orzechowska, A., Talarowska, M., Szemraj, J. and Gałecki, P. (2018). The importance of TCF4 gene in the etiology of recurrent depressive disorders. *Prog Neuropsychopharmacol Biol Psychiatry* 80, 304–308.

Muir, T., Sadler-Riggleman, I., Stevens, J. D. and Skinner, M. K. (2006). Role of the basic helix-loop-helix protein ITF2 in the hormonal regulation of Sertoli cell differentiation. *Mol Reprod Dev* 73, 491–500.

Murakami, M., Kataoka, K., Tominaga, J., Nakagawa, O. and Kurihara, H. (2004). Differential cooperation between dHAND and three different E-proteins. *Biochem Biophys Res Commun* 323, 168–174.

Murre, C., McCaw, P. S. and Baltimore, D. (1989a). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* 56, 777–783.

Murre, C., McCaw, P. S., Vaessin, H., Caudy, M., Jan, L. Y., Jan, Y. N., Cabrera, C. V., Buskin, J. N., Hauschka, S. D., Lassar, A. B., et al. (1989b). Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* 58, 537–544.

Neumüller, R. A., Richter, C., Fischer, A., Novatchkova, M., Neumüller, K. G. and Knoblich, J. A. (2011). Genome-wide analysis of self-renewal in Drosophila neural stem cells by transgenic RNAi. *Cell Stem Cell* 8, 580–593.

Nitta, Y. and Sugie, A. (2022). Studies of neurodegenerative diseases using Drosophila and the development of novel approaches for their analysis. *Fly (Austin)* 16, 275–298.

Nolo, R., Abbott, L. A. and Bellen, H. J. (2000). Senseless, a Zn Finger Transcription Factor, Is Necessary and Sufficient for Sensory Organ Development in Drosophila. *Cell* 102, 349–362.

Nurm, K., Sepp, M., Castany-Pladevall, C., Creus-Muncunill, J., Tuvikene, J., Sirp, A., Vihma, H., Blake, D. J., Perez-Navarro, E. and Timmusk, T. (2021). Isoform-Specific Reduction of the Basic Helix-Loop-Helix Transcription Factor TCF4 Levels in Huntington's Disease. *eNeuro* 8, ENEURO.0197-21.2021.

O'Donnell, L., Soileau, B., Heard, P., Carter, E., Sebold, C., Gelfond, J., Hale, D. E. and Cody, J. D. (2010). Genetic determinants of autism in individuals with deletions of 18q. *Hum Genet* 128, 155–164.

Oellers, N., Dehio, M. and Knust, E. (1994). bHLH proteins encoded by the Enhancer of split complex of Drosophila negatively interfere with transcriptional activation mediated by proneural genes. *Mol Gen Genet* 244, 465–473.

Okumura, N., Hayashi, R., Nakano, M., Yoshii, K., Tashiro, K., Sato, T., Blake, D. J., Aleff, R., Butz, M., Highsmith, E. W., et al. (2019). Effect of Trinucleotide Repeat Expansion on the Expression of TCF4 mRNA in Fuchs' Endothelial Corneal Dystrophy. *Invest Ophthalmol Vis Sci* 60, 779–786.

Onai, N., Asano, J., Kurosaki, R., Kuroda, S. and Ohteki, T. (2017). Flexible fate commitment of E2-2high common DC progenitors implies tuning in tissue microenvironments. *Int Immunol* 29, 443–456.

Page, S. C., Hamersky, G. R., Gallo, R. A., Rannals, M. D., Calcaterra, N. E., Campbell, M. N., Mayfield, B., Briley, A., Phan, B. N., Jaffe, A. E., et al. (2018). The schizophrenia- and autism-associated gene, transcription factor 4 regulates the columnar distribution of layer 2/3 prefrontal pyramidal neurons in an activity-dependent manner. *Mol. Psychiatry* 23, 304–315.

Papes, F., Camargo, A. P., de Souza, J. S., Carvalho, V. M. A., Szeto, R. A., LaMontagne, E., Teixeira, J. R., Avansini, S. H., Sánchez-Sánchez, S. M., Nakahara, T. S., et al. (2022). Transcription Factor 4 loss-of-function is associated with deficits in progenitor proliferation and cortical neuron content. *Nat Commun* 13, 2387.

Park, S., Bustamante, E. L., Antonova, J., McLean, G. W. and Kim, S. K. (2011). Specification of Drosophila Corpora Cardiaca Neuroendocrine Cells from Mesoderm Is Regulated by Notch Signaling. *PLoS Genet* 7, e1002241.

Park, S.-J., Ahmad, F., Philp, A., Baar, K., Williams, T., Luo, H., Ke, H., Rehmann, H., Taussig, R., Brown, A. L., et al. (2012). Resveratrol ameliorates aging-related metabolic phenotypes by inhibiting cAMP phosphodiesterases. *Cell* 148, 421–433.

Peippo, M. and Ignatius, J. (2012). Pitt-Hopkins Syndrome. *Mol Syndromol* 2, 171–180.

Pernía, O., Sastre-Perona, A., Rodriguez-Antolín, C., García-Guede, A., Palomares-Bralo, M., Rosas, R., Sanchez-Cabrero, D., Cruz, P., Rodriguez, C., Diestro, Md., et al. (2020). A Novel Role for the Tumor Suppressor Gene ITF2 in Tumorigenesis and Chemotherapy Response. *Cancers (Basel)* 12, E786.

Petropoulos, H. and Skerjanc, I. S. (2000). Analysis of the inhibition of MyoD activity by ITF-2B and full-length E12/E47. *J Biol Chem* 275, 25095–25101.

Phan, B. N., Bohlen, J. F., Davis, B. A., Ye, Z., Chen, H.-Y., Mayfield, B., Sripathy, S. R., Page, S. C., Campbell, M. N., Smith, H. L., et al. (2020). A myelin-related transcriptomic profile is shared between Pitt Hopkins syndrome models and human autism spectrum disorder. *Nat Neurosci* 23, 375–385.

Pi, H., Huang, S.-K., Tang, C.-Y., Sun, Y. H. and Chien, C.-T. (2004). phyllopod is a target gene of proneural proteins in Drosophila external sensory organ development. *PNAS* 101, 8378–8383.

Pickard, B. S., Malloy, M. P., Clark, L., Lehellard, S., Ewald, H. L., Mors, O., Porteous, D. J., Blackwood, D. H. R. and Muir, W. J. (2005). Candidate psychiatric illness genes identified in patients with pericentric inversions of chromosome 18. *Psychiatr Genet* 15, 37–44.

Pscherer, A., Dörflinger, U., Kirfel, J., Gawlas, K., Rüschoff, J., Buettner, R. and Schüle, R. (1996). The helix-loop-helix transcription factor SEF-2 regulates the activity of a novel initiator element in the promoter of the human somatostatin receptor II gene. *EMBO J* 15, 6680–6690.

Qian, L., Bajana, S., Georgescu, C., Peng, V., Wang, H.-C., Adrianto, I., Colonna, M., Alberola-Ila, J., Wren, J. D. and Sun, X.-H. (2019). Suppression of ILC2 differentiation from committed T cell precursors by E protein transcription factors. *J Exp Med* 216, 884–899.

Quednow, B. B., Ettinger, U., Mössner, R., Rujescu, D., Giegling, I., Collier, D. A., Schmechtig, A., Kühn, K.-U., Möller, H.-J., Maier, W., et al. (2011). The schizophrenia risk allele C of the TCF4 rs9960767 polymorphism disrupts sensorimotor gating in schizophrenia spectrum and healthy volunteers. *J. Neurosci.* 31, 6684–6691.

Quednow, B. B., Brinkmeyer, J., Mobascher, A., Nothnagel, M., Musso, F., Gründer, G., Savary, N., Petrovsky, N., Frommann, I., Lennertz, L., et al. (2012). Schizophrenia risk polymorphisms in the TCF4 gene interact with smoking in the modulation of auditory sensory gating. *Proc Natl Acad Sci U S A* 109, 6271–6276.

Quevedo, M., Meert, L., Dekker, M. R., Dekkers, D. H. W., Brandsma, J. H., van den Berg, D. L. C., Ozgür, Z., van IJcken, W. F. J., Demmers, J., Fornerod, M., et al. (2019). Mediator complex interaction partners organize the transcriptional network that defines neural stem cells. *Nat Commun* 10, 2669.

Quong, M. W., Massari, M. E., Zwart, R. and Murre, C. (1993). A new transcriptional-activation motif restricted to a class of helix-loop-helix proteins is functionally conserved in both yeast and mammalian cells. *Mol Cell Biol* 13, 792–800.

Ramain, P., Khechumian, R., Khechumian, K., Arbogast, N., Ackermann, C. and Heitzler, P. (2000). Interactions between Chip and the Achaete/Scute–Daughterless Heterodimers Are Required for Pannier-Driven Proneural Patterning. *Molecular Cell* 6, 781–790.

Rannals, M. D., Hamersky, G. R., Page, S. C., Campbell, M. N., Briley, A., Gallo, R. A., Phan, B. N., Hyde, T. M., Kleinman, J. E., Shin, J. H., et al. (2016a). Psychiatric Risk Gene Transcription Factor 4 Regulates Intrinsic Excitability of Prefrontal Neurons via Repression of SCN10a and KCNQ1. *Neuron* 90, 43–55.

Rannals, M. D., Page, S. C., Campbell, M. N., Gallo, R. A., Mayfield, B. and Maher, B. J. (2016b). Neurodevelopmental models of transcription factor 4 deficiency converge on a common ion channel as a potential therapeutic target for Pitt Hopkins syndrome. *Rare Dis* 4,.

Reeves, N. and Posakony, J. W. (2005). Genetic Programs Activated by Proneural Proteins in the Developing Drosophila PNS. *Developmental Cell* 8, 413–425.

Rex, E. B., Shukla, N., Gu, S., Bredt, D. and DiSepio, D. (2017). A Genome-Wide Arrayed cDNA Screen to Identify Functional Modulators of α7 Nicotinic Acetylcholine Receptors. *SLAS Discov* 22, 155–165.

Rosato, M., Stringer, S., Gebuis, T., Paliukhovich, I., Li, K. W., Posthuma, D., Sullivan, P. F., Smit, A. B. and van Kesteren, R. E. (2021). Combined cellomics and proteomics analysis reveals shared neuronal morphology and molecular pathway phenotypes for multiple schizophrenia risk genes. *Mol Psychiatry* 26, 784–799.

Rosay, P., Colas, J. F. and Maroteaux, L. (1995). Dual organisation of the Drosophila neuropeptide receptor NKD gene promoter. *Mech Dev* 51, 329–339.

Saarikettu, J., Sveshnikova, N. and Grundström, T. (2004). Calcium/calmodulin inhibition of transcriptional activity of E-proteins by prevention of their binding to DNA. *J Biol Chem* 279, 41004–41011.

Sandler, L. (1972). On the Genetic Control of Genes Located in the Sex-Chromosome Heterochromatin of DROSOPHILA MELANOGASTER. *Genetics* 70, 261–274.

Sarkar, D., Shariq, M., Dwivedi, D., Krishnan, N., Naumann, R., Bhalla, U. S. and Ghosh, H. S. (2021). Adult brain neurons require continual expression of the schizophrenia-risk gene Tcf4 for structural and functional integrity. *Transl Psychiatry* 11, 494.

Savio, A. J., Daftary, D., Dicks, E., Buchanan, D. D., Parfrey, P. S., Young, J. P., Weisenberger, D., Green, R. C., Gallinger, S., McLaughlin, J. R., et al. (2016). Promoter methylation of ITF2, but not APC, is associated with microsatellite instability in two populations of colorectal cancer patients. *BMC Cancer* 16, 113.

Schmidt-Edelkraut, U., Daniel, G., Hoffmann, A. and Spengler, D. (2014). Zac1 regulates cell cycle arrest in neuronal progenitors via Tcf4. *Mol Cell Biol* 34, 1020–1030.

Schoof, M., Hellwig, M., Harrison, L., Holdhof, D., Lauffer, M. C., Niesen, J., Virdi, S., Indenbirken, D. and Schüller, U. (2020). The basic helix-loop-helix transcription factor TCF4 impacts brain architecture as well as neuronal morphology and differentiation. *Eur J Neurosci* 51, 2219–2235.

Sepp, M., Kannike, K., Eesmaa, A., Urb, M. and Timmusk, T. (2011). Functional diversity of human basic helix-loop-helix transcription factor TCF4 isoforms generated by alternative 5' exon usage and splicing. *PLoS ONE* 6, e22138.

Sepp, M., Pruunsild, P. and Timmusk, T. (2012). Pitt-Hopkins syndrome-associated mutations in TCF4 lead to variable impairment of the transcription factor function ranging from hypomorphic to dominant-negative effects. *Hum. Mol. Genet.* 21, 2873–2888.

Sepp, M., Vihma, H., Nurm, K., Urb, M., Page, S. C., Roots, K., Hark, A., Maher, B. J., Pruunsild, P. and Timmusk, T. (2017). The Intellectual Disability and Schizophrenia Associated Transcription Factor TCF4 Is Regulated by Neuronal Activity and Protein Kinase A. *J Neurosci* 37, 10516–10527.

Shariq, M., Sahasrabuddhe, V., Krishna, S., Radha, S., Nruthyathi, Bellampalli, R., Dwivedi, A., Cheramangalam, R., Reizis, B., Hébert, J., et al. (2021). Adult neural stem cells have latent inflammatory potential that is kept suppressed by Tcf4 to facilitate adult neurogenesis. *Sci Adv* 7, eabf5606.

Sirp, A., Leite, K., Tuvikene, J., Nurm, K., Sepp, M. and Timmusk, T. (2020). The Fuchs corneal dystrophy-associated CTG repeat expansion in the TCF4 gene affects transcription from its alternative promoters. *Sci Rep* 10, 18424.

Sirp, A., Roots, K., Nurm, K., Tuvikene, J., Sepp, M. and Timmusk, T. (2021). Functional consequences of TCF4 missense substitutions associated with Pitt-Hopkins syndrome, mild intellectual disability, and schizophrenia. *Journal of Biological Chemistry* 297, 101381.

Smith, J. E. and Cronmiller, C. (2001). The Drosophila daughterless gene autoregulates and is controlled by both positive and negative cis regulation. *Development* 128, 4705–4714.

Smith, J. E., Cummings, C. A. and Cronmiller, C. (2002). Daughterless coordinates somatic cell proliferation, differentiation and germline cyst survival during follicle formation in Drosophila. *Development* 129, 3255–3267.

Sobrado, V. R., Moreno-Bueno, G., Cubillo, E., Holt, L. J., Nieto, M. A., Portillo, F. and Cano, A. (2009). The class I bHLH factors E2-2A and E2-2B regulate EMT. *J Cell Sci* 122, 1014–1024.

Soh, Y. Q., Peh Swee Lim, G., Htoon, H. M., Gong, X., Mootha, V. V., Vithana, E. N., Kocaba, V. and Mehta, J. S. (2019). Trinucleotide repeat expansion length as a predictor of the clinical progression of Fuchs' Endothelial Corneal Dystrophy. *PLoS One* 14, e0210996.

Soosaar, A., Chiaramello, A., Zuber, M. X. and Neuman, T. (1994). Expression of basic-helix-loop-helix transcription factor ME2 during brain development and in the regions of neuronal plasticity in the adult brain. *Molecular Brain Research* 25, 176–180.

Sparber, P., Filatova, A., Anisimova, I., Markova, T., Voinova, V., Chuhrova, A., Tabakov, V. and Skoblov, M. (2020). Various haploinsufficiency mechanisms in Pitt-Hopkins syndrome. *Eur J Med Genet* 63, 104088.

Spratford, C. M. and Kumar, J. P. (2015). Inhibition of Daughterless by Extramacrochaetae mediates Notch-induced cell proliferation. *Development* 142, 2058–2068.

Stavropoulos, D. J., MacGregor, D. L. and Yoon, G. (2010). Mosaic microdeletion 18q21 as a cause of mental retardation. *Eur J Med Genet* 53, 396–399.

Stefansson, H., Ophoff, R. A., Steinberg, S., Andreassen, O. A., Cichon, S., Rujescu, D., Werge, T., Pietiläinen, O. P. H., Mors, O., Mortensen, P. B., et al. (2009). Common variants conferring risk of schizophrenia. *Nature* 460, 744–747.

Steinberg, S., de Jong, S., Irish Schizophrenia Genomics Consortium, Andreassen, O. A., Werge, T., Børglum, A. D., Mors, O., Mortensen, P. B., Gustafsson, O., Costas, J., et al. (2011). Common variants at VRK2 and TCF4 conferring risk of schizophrenia. *Hum Mol Genet* 20, 4076–4081.

Steinbusch, C. V. M., van Roozendaal, K. E. P., Tserpelis, D., Smeets, E. E. J., Kranenburg-de Koning, T. J., de Waal, K. H., Zweier, C., Rauch, A., Hennekam, R. C. M., Blok, M. J., et al. (2013). Somatic mosaicism in a mother of two children with Pitt-Hopkins syndrome. *Clin Genet* 83, 73–77.

Stephan, M., Schoeller, J., Raabe, F. J., Schmitt, A., Hasan, A., Falkai, P., Jensen, N. and Rossner, M. J. (2022). Spironolactone alleviates schizophrenia-related reversal learning in Tcf4 transgenic mice subjected to social defeat. *Schizophrenia (Heidelb)* 8, 77.

Straub, J., Gregor, A., Sauerer, T., Fliedner, A., Distel, L., Suchy, C., Ekici, A. B., Ferrazzi, F. and Zweier, C. (2020). Genetic interaction screen for severe neurodevelopmental disorders reveals a functional link between Ube3a and Mef2 in Drosophila melanogaster. *Sci Rep* 10, 1204.

Sukhanova, M. J., Deb, D. K., Gordon, G. M., Matakatsu, M. T. and Du, W. (2007). Proneural Basic Helix-Loop-Helix Proteins and Epidermal Growth Factor Receptor Signaling Coordinately Regulate Cell Type Specification and cdk Inhibitor Expression during Development. *Mol Cell Biol* 27, 2987–2996.

Talkowski, M. E., Rosenfeld, J. A., Blumenthal, I., Pillalamarri, V., Chiang, C., Heilbut, A., Ernst, C., Hanscom, C., Rossin, E., Lindgren, A. M., et al. (2012). Sequencing chromosomal abnormalities reveals neurodevelopmental loci that confer risk across diagnostic boundaries. *Cell* 149, 525–537.

Tanaka-Matakatsu, M., Miller, J., Borger, D., Tang, W.-J. and Du, W. (2014). Daughterless homodimer synergizes with Eyeless to induce Atonal expression and retinal neuron differentiation. *Dev Biol* 392, 256–265.

Teixeira, J. R., Szeto, R. A., Carvalho, V. M. A., Muotri, A. R. and Papes, F. (2021). Transcription factor 4 and its association with psychiatric disorders. *Transl Psychiatry* 11, 1–12.

Tepass, U. and Hartenstein, V. (1995). Neurogenic and proneural genes control cell fate specification in the Drosophila endoderm. *Development* 121, 393–405.

Thaxton, C., Kloth, A. D., Clark, E. P., Moy, S. S., Chitwood, R. A. and Philpot, B. D. (2018). Common Pathophysiology in Multiple Mouse Models of Pitt–Hopkins Syndrome. *J Neurosci* 38, 918–936.

Tögel, M., Meyer, H., Lehmacher, C., Heinisch, J. J., Pass, G. and Paululat, A. (2013). The bHLH transcription factor hand is required for proper wing heart formation in Drosophila. *Developmental Biology* 381, 446–459.

Tuvikene, J., Esvald, E.-E., Rähni, A., Uustalu, K., Zhuravskaya, A., Avarlaid, A., Makeyev, E. V. and Timmusk, T. (2021). Intronic enhancer region governs transcript-specific Bdnf expression in rodent neurons. *Elife* 10, e65161.

Uittenbogaard, M. and Chiaramello, A. (2000). Differential expression patterns of the basic helix-loop-helix transcription factors during aging of the murine brain. *Neurosci Lett* 280, 95–98.

Vaessin, H., Brand, M., Jan, L. Y. and Jan, Y. N. (1994). daughterless is essential for neuronal precursor differentiation but not for initiation of neuronal precursor formation in Drosophila embryo. *Development* 120, 935–945.

van der Voet, M., Nijhof, B., Oortveld, M. A. W. and Schenck, A. (2014). Drosophila models of early onset cognitive disorders and their clinical applications. *Neuroscience & Biobehavioral Reviews* 46, 326–342.

Van Doren, M., Ellis, H. M. and Posakony, J. W. (1991). The Drosophila extramacrochaetae protein antagonizes sequence-specific DNA binding by daughterless/achaete-scute protein complexes. *Development* 113, 245–255.

Vasanth, S., Eghrari, A. O., Gapsis, B. C., Wang, J., Haller, N. F., Stark, W. J., Katsanis, N., Riazuddin, S. A. and Gottsch, J. D. (2015). Expansion of CTG18.1 Trinucleotide Repeat in TCF4 Is a Potent Driver of Fuchs' Corneal Dystrophy. *Invest Ophthalmol Vis Sci* 56, 4531–4536.

Volkmann, P., Stephan, M., Krackow, S., Jensen, N. and Rossner, M. J. (2020). PsyCoP - A Platform for Systematic Semi-Automated Behavioral and Cognitive Profiling Reveals Gene and Environment Dependent Impairments of Tcf4 Transgenic Mice Subjected to Social Defeat. *Front Behav Neurosci* 14, 618180.

Vrânceanu, M., Galimberti, D., Banc, R., Dragoş, O., Cozma-Petruţ, A., Hegheş, S.-C., Voştinaru, O., Cuciureanu, M., Stroia, C. M., Miere, D., et al. (2022). The Anticancer Potential of Plant-Derived Nutraceuticals via the Modulation of Gene Expression. *Plants (Basel)* 11, 2524.

Waddell, E. A., Viveiros, J. M., Robinson, E. L., Sharoni, M. A., Latcheva, N. K. and Marenda, D. R. (2019). Extramacrochaetae promotes branch and bouton number via the sequestration of daughterless in the cytoplasm of neurons. *Dev Neurobiol* 79, 805–818.

Wang, L.-H. and Baker, N. E. (2015). Salvador-Warts-Hippo pathway in a developmental checkpoint monitoring helix-loop-helix proteins. *Dev Cell* 32, 191–202.

Wang, L.-H. and Baker, N. E. (2018). Spatial regulation of expanded transcription in the Drosophila wing imaginal disc. *PLoS One* 13,.

Wang, Y., Lu, Z., Zhang, Y., Cai, Y., Yun, D., Tang, T., Cai, Z., Wang, C., Zhang, Y., Fang, F., et al. (2020). Transcription Factor 4 Safeguards Hippocampal Dentate Gyrus Development by Regulating Neural Progenitor Migration. *Cerebral Cortex* 30, 3102–3115.

Wang, Y., Liu, L. and Lin, M. (2022). Psychiatric risk gene transcription factor 4 preferentially regulates cortical interneuron neurogenesis during early brain development. *J Biomed Res* 36, 242–254.

Wawruszak, A., Borkiewicz, L., Okon, E., Kukula-Koch, W., Afshan, S. and Halasa, M. (2021). Vorinostat (SAHA) and Breast Cancer: An Overview. *Cancers (Basel)* 13, 4700.

Wedel, M., Fröb, F., Elsesser, O., Wittmann, M.-T., Lie, D. C., Reis, A. and Wegner, M. (2020). Transcription factor Tcf4 is the preferred heterodimerization partner for Olig2 in oligodendrocytes and required for differentiation. *Nucleic Acids Res* 48, 4839–4857.

Wesley, C. S. and Saez, L. (2000). Analysis of Notch Lacking the Carboxyl Terminus Identified in Drosophila Embryos. *J Cell Biol* 149, 683–696.

Whalen, S., Héron, D., Gaillon, T., Moldovan, O., Rossi, M., Devillard, F., Giuliano, F., Soares, G., Mathieu-Dramard, M., Afenjar, A., et al. (2012). Novel comprehensive diagnostic strategy in Pitt-Hopkins syndrome: clinical score and further delineation of the TCF4 mutational spectrum. *Hum Mutat* 33, 64–72.

Wikström, I., Forssell, J., Penha-Goncalves, M. N., Bergqvist, I. and Holmberg, D. (2008). A role for E2-2 at the DN3 stage of early thymopoiesis. *Mol Immunol* 45, 3302–3311.

Wirgenes, K. V., Sønderby, I. E., Haukvik, U. K., Mattingsdal, M., Tesli, M., Athanasiu, L., Sundet, K., Røssberg, J. I., Dale, A. M., Brown, A. A., et al. (2012). TCF4 sequence variants and mRNA levels are associated with neurodevelopmental characteristics in psychotic disorders. *Transl Psychiatry* 2, e112.

Wittmann, M.-T., Katada, S., Sock, E., Kirchner, P., Ekici, A. B., Wegner, M., Nakashima, K., Lie, D. C. and Reis, A. (2021). scRNA sequencing uncovers a TCF4-dependent transcription factor network regulating commissure development in mouse. *Development* 148, dev196022.

Wöhner, M., Tagoh, H., Bilic, I., Jaritz, M., Poliakova, D. K., Fischer, M. and Busslinger, M. (2016). Molecular functions of the transcription factors E2A and E2-2 in controlling germinal center B cell and plasma cell development. *J Exp Med* 213, 1201–1221.

Wong, M.-C., Castanon, I. and Baylies, M. K. (2008). Daughterless dictates Twist activity in a context-dependent manner during somatic myogenesis. *Dev. Biol.* 317, 417–429.

Wong, G. C.-H., Li, K. K.-W., Wang, W.-W., Liu, A. P.-Y., Huang, Q. J., Chan, A. K.-Y., Poon, M. F.-M., Chung, N. Y.-F., Wong, Q. H.-W., Chen, H., et al. (2020). Clinical and mutational profiles of adult medulloblastoma groups. *Acta Neuropathol Commun* 8, 191.

Wy, C., J, Z., H, G., Z, D., T, N. and Rg, R. (2013). A TAF4 coactivator function for E proteins that involves enhanced TFIID binding. *Genes & development* 27,.

Xiang, X., Zhuang, X., Jiang, H., Ju, S., Mu, J., Zhang, L., Miller, D., Grizzle, W. and Zhang, H.-G. (2011). miR-155 promotes macroscopic tumor formation yet inhibits tumor dissemination from mammary fat pads to the lung by preventing EMT. *Oncogene* 30, 3440–3453.

Xue, H., Tu, Y., Ma, T., Wen, T., Yang, T., Xue, L., Cai, M., Wang, F. and Guan, M. (2019). miR-93-5p attenuates IL-1β-induced chondrocyte apoptosis and cartilage degradation in osteoarthritis partially by targeting TCF4. *Bone* 123, 129–136.

Yamaguchi, M., Lee, I.-S., Jantrapirom, S., Suda, K. and Yoshida, H. (2021). Drosophila models to study causative genes for human rare intractable neurological diseases. *Exp Cell Res* 403, 112584.

Yamamoto, F. and Yamamoto, M. (2007). Scanning copy number and gene expression on the 18q21-qter chromosomal region by the systematic multiplex PCR and reverse transcription-PCR methods. *Electrophoresis* 28, 1882–1895.

Yang, D.-J., Chung, J.-Y., Lee, S.-J., Park, S.-Y., Pyo, J.-H., Ha, N.-C., Yoo, M.-A. and Park, B.-J. (2010). Slug, mammalian homologue gene of Drosophila escargot, promotes neuronal-differentiation through suppression of HEB/daughterless. *Cell Cycle* 9, 2789–2802.

Yasugi, T., Umetsu, D., Murakami, S., Sato, M. and Tabata, T. (2008). Drosophila optic lobe neuroblasts triggered by a wave of proneural gene expression that is negatively regulated by JAK/STAT. *Development* 135, 1471–1480.

Yasugi, T., Fischer, A., Jiang, Y., Reichert, H. and Knoblich, J. A. (2014). A Regulatory Transcriptional Loop Controls Proliferation and Differentiation in Drosophila Neural Stem Cells. *PLoS One* 9,.

Yu, Y., Liang, Y., Liu, X., Yang, H., Su, Y., Xia, X. and Wang, H. (2016a). Id1 modulates endothelial progenitor cells function through relieving the E2-2-mediated repression of FGFR1 and VEGFR2 in vitro. *Mol Cell Biochem* 411, 289–298.

Yu, Y., Liang, Y., Yin, C., Liu, X., Su, Y., Zhang, L. and Wang, H. (2016b). Inhibitor of DNA-binding 1 promotes endothelial progenitor cell proliferation and migration by suppressing E2-2 through the helix-loop-helix domain. *Int J Mol Med* 38, 1549–1557.

Zarifi, I., Kiparaki, M., Koumbanakis, K. A., Giagtzoglou, N., Zacharioudaki, E., Alexiadis, A., Livadaras, I. and Delidakis, C. (2012). Essential Roles of Da Transactivation Domains in Neurogenesis and in E(spl)-Mediated Repression. *Mol Cell Biol* 32, 4534–4548.

Zhang, Y., Cai, Y., Wang, Y., Deng, X., Zhao, Y., Zhang, Y. and Xie, Y. (2021a). Survival control of oligodendrocyte progenitor cells requires the transcription factor 4 during olfactory bulb development. *Cell Death Dis* 12, 91.

Zhang, Y., Cai, Z., Hu, G., Hu, S., Wang, Y., Li, N., Chen, S., Liu, Q., Zeng, L., Tang, T., et al. (2021b). Transcription factor 4 controls positioning of cortical projection neurons through regulation of cell adhesion. *Mol Psychiatry* 26, 6562–6577.

Zhao, H., Jiang, A., Yu, M. and Bao, H. (2020). Identification of biomarkers correlated with diagnosis and prognosis of endometrial cancer using bioinformatics analysis. *Journal of Cellular Biochemistry* 121, 4908–4921.

Zhuang, Y., Cheng, P. and Weintraub, H. (1996). B-lymphocyte development is regulated by the combined dosage of three basic helix-loop-helix genes, E2A, E2-2, and HEB. *Mol Cell Biol* 16, 2898–2905.

Zweier, C., Peippo, M. M., Hoyer, J., Sousa, S., Bottani, A., Clayton-Smith, J., Reardon, W., Saraiva, J., Cabral, A., Gohring, I., et al. (2007). Haploinsufficiency of TCF4 causes syndromal mental retardation with intermittent hyperventilation (Pitt-Hopkins syndrome). *Am. J. Hum. Genet.* 80, 994–1001.

Zweier, C., Sticht, H., Bijlsma, E. K., Clayton-Smith, J., Boonen, S. E., Fryer, A., Greally, M. T., Hoffmann, L., Hollander, N. S. den, Jongmans, M., et al. (2008). Further delineation of Pitt–Hopkins syndrome: phenotypic and genotypic description of 16 novel patients. *Journal of Medical Genetics* 45, 738–744.

Acknowledgements

First, I want to thank my supervisors Mari Palgi and Tõnis Timmusk. I am very grateful to Mari for introducing me to the "fly world", for guiding me through Master's and Doctoral studies and encouraging me along the way. I am thankful to Tõnis for taking me into a great research group both in scientific sense and work-environmental sense. It has really been a wonderful place to carry out my studies.

I thank all the co-authors for their contributions to the publications and all the Neurobiology laboratory members. Without them this work would have never been done. I am especially thankful to Alex, Annela, Anastassia and Carl Sander who were my companions to share the good times and my rocks at the tough times. I am very grateful to Jürgen for huge help with many experiments and for all the discussions, I have learned a lot from him. My special thanks go to Epp and Florencia for being helpful, warm and friendly.

My sincere gratitude goes to my parents Žanna and Artur, my sister Karina and my son Ragnar. I am really lucky to have such a supporting team behind my back whom I can always count on.

I thank all the funding agencies that supported this work: Estonian Research Council (grants MJD341, IUT19-18 and PRG805), National R&D program "Health" (grant AR12098), Estonian Academy of Sciences, the European Union through the European Regional Development Fund (Project No. 2014-2020.4.01.15-0012) and H2020-MSCA-RISE-2016 (EU734791), the Pitt Hopkins Research Foundation (grants 8 and 21), the Million Dollar Bike Ride Pilot Grant Program for Rare Disease Research at UPenn Orphan Disease Center (grants MDBR-16-122-PHP and MDBR-17-127-Pitt Hopkins). I thank the "TUT Institutional Development Program for 2016−2022" Graduate School in Clinical Medicine, which received funding from the European Regional Development Fund under program ASTRA 2014-2020.4.01.16-0032 in Estonia.

Abstract

Studies of the *Drosophila* basic helix-loop-helix transcription factor Daughterless and its mammalian homologue Transcription factor 4

Transcription factor 4 (TCF4) is a class I basic helix-loop-helix (bHLH) transcription factor that carries out a variety of functions. Most investigated are the roles of TCF4 in the development of nervous system, the functioning of adult nervous system, development and functioning of the immune system and regulation of cell cycle. The importance of TCF4 in the nervous system is emphasised by the diseases associated with it – mutations in *TCF4* gene cause rare but serious mental retardation named Pitt-Hopkins syndrome (PTHS) and mild to moderate intellectual disability. Moreover, variations of *TCF4* gene are associated with schizophrenia and other psychiatric disorders. In addition, repeat expansions in *TCF4* gene cause Fuchs corneal dystrophy and *TCF4* is implied in cancer. This makes *TCF4* as an important gene for research and animal models are needed to facilitate the research.

Currently, there are different mouse models for *TCF4*-related intellectual disabilities available, including *TCF4* haploinsufficiency and conditional knock outs. Using these models, great advancements in understanding *TCF4*-related neurological diseases have been made, but screening for drugs in the mouse models is not optimal. *Drosophila melanogaster* disease models are suitable and efficient for drug screening and modeling *TCF4*-related diseases in the fly has great potential. The TCF4 homolog in *Drosophila melanogaster* is Daughterless (Da). Da has many roles during development including embryogenesis, sex determination, oogenesis, cell cycle, but it is mostly known for its roles in the development of the nervous system. In addition, Da functions in the mature nervous system by regulating synaptogenesis.

In this thesis, we set out to generate and validate *Drosophila melanogaster* models for PTHS and other *TCF4*-related intellectual disabilities. In addition, we characterize the tissue specific expression of *TCF4* transcripts and protein isoforms in human and the most used animal models – mouse and rat. This knowledge is important for laying the basis for developing gene therapy applications for PTHS.

First, we showed that TCF4 and Da are true orthologs, since human TCF4 is able to activate the development of the nervous system of *Drosophila melanogaster* and replace it during embryonic neurogenesis. Moreover, PTHS-associated mutations impair the functioning of Da similarly to TCF4 - corresponding mutations cause hypomorphic, loss of function or dominant negative effects in both proteins and some of the effects are context specific. This affirmed that *Drosophila*'s *da* is suitable for studying *TCF4*-related diseases.

Second, to generate models for PTHS and other TCF4-related intellectual disabilities, we used RNA interference-based silencing of *da* in the *Drosophila melanogaster* nervous system and two behavioral assays – larval appetitive associative learning and negative geotaxis of adult flies. Our results showed that decreased Da expression in the nervous system impairs the memory of the larvae and locomotion of the adults. In addition, we found that Da functions in regulating the synaptic proteome. Next, we validated the *Drosophila melanogaster* models for drug screening. The phenotypes of our *TCF4*-related intellectual disability models were improved by two drugs – suberoylanilide hydroxamic

acid (SAHA) and resveratrol. Moreover, the transcriptional ability of both Da and TCF4 was increased by these drugs.

Third, we set out to systematically characterise the expression of *Tcf4* transcripts and protein isoforms in two most extensively used animal models – mouse and rat. In addition, we investigated the expression of *TCF4* in human. We found that even though *TCF4* gene structure is complicated and allows for a variety of transcripts to be expressed, the main transcripts in both rodents and humans encode TCF4-A, TCF4-D, TCF4-B, TCF4-C and TCF4-I protein isoforms. *TCF4* expression is highest in the brain but it is also expressed in numerous non-neural tissues.

Collectively, the *TCF4*-related disease models generated and characterised by us provide tools for developing therapeutics for PTHS and other diseases associated with *TCF4.* The *Drosophila melanogaster* models have great potential in drug screening and discovery and the knowledge of isoform and tissue specific *TCF4* expression is crucial for development of gene therapy.

Lühikokkuvõte

Äädikakärbse aluselise heeliks-ling-heeliks transkriptsioonifaktori Daughterless ja tema imetaja homoloogi Transkriptsioonifaktor 4 uuringud

Transkriptsioonifaktor 4 (TCF4) on klass I aluseline heeliks-ling-heeliks transkriptsioonifaktor, millel on organismis mitmeid funktsioone. Enim on uuritud TCF4 rolli närvisüsteemi arengus, täiskasvanud närvisüsteemi talitluses, immuunsüsteemi arengus ja toimimises ning rakutsükli reguleerimises. TCF4 olulisust närvisüsteemis rõhutab asjaolu, et see on seotud raskete haigustega – mutatsioonid *TCF4* geenis põhjustavad haruldast kuid tõsis vaimset alaarengut nimega Pitt-Hopkinsi sündroom (PTHS) ja kerget kuni mõõdukat intellektipuuet. Enamgi veel, *TCF4* geeni variatsioonid on seotud skisofreenia ja teiste psühhiaatriliste häiretega. Lisaks põhjustavad spetsiifilised kordused *TCF4* geenis Fuchsi sarvkesta düstroofiat ja *TCF4* on seotud vähiga. See muudab *TCF4* teaduse jaoks oluliseks geeniks ja uurimistöö hõlbustamiseks on vaja loommudeleid.

TCF4-seoseliste intellektipuude haiguste jaoks on loodud erinevaid hiiremudeleid, sealhulgas *Tcf4* heterosügoodid ja konditsionaalsed *knock out'*id. Neid mudeleid kasutades on tehtud suuri edusamme *TCF4*-ga seotud neuroloogiliste haiguste mõistmisel, kuid ravimite sõeluuringute läbiviimine hiiremudelites ei ole optimaalne. *Drosophila melanogaster'*i haiguste mudelid on sobivad ja tõhusad ravimite sõeluuringuteks ning *TCF4*-ga seotud haiguste modelleerimisel kärbses on suur potentsiaal. *Drosophila melanogaster*'i TCF4 homoloog on Daughterless (Da). Da-l on arengu ajal palju rolle, sealhulgas embrüogenees, soo määramine, oogenees, rakutsükkel, kuid see on peamiselt tuntud oma rollide poolest närvisüsteemi arengus. Lisaks toimib Da täiskasvanud närvisüsteemis, reguleerides sünaptogeneesi.

Käesolevas väitekirjas lõime ja valideerisime PTHS-i ja teiste *TCF4*-ga seotud intellektipuuete *Drosophila melanogaster*'i haigusmudelid. Lisaks iseloomustasime *TCF4* isovormide koespetsiifilist ekspressiooni enimkasutatud loommudelites – hiirtes ja rottides ning lisaks ka inimeses. Need teadmised on oluliseks aluseks geeniteraapia loomisel.

Esiteks tõestasime, et TCF4 ja Da on tõelised ortoloogid, kuna inimese TCF4 on võimeline aktiveerima *Drosophila melanogaster'*i närvisüsteemi arengut ja asendama seda embrüonaalse neurogeneesi ajal. Veelgi enam, PTHS-iga seotud mutatsioonid mõjutavad Da toimimist sarnaselt TCF4-ga – vastavad mutatsioonid põhjustavad mõlema valgu hüpomorfset, funktsiooni kaotust põhjustavat või domineerivat negatiivset mõju ning mõned nendest mõjudest on kontekstispetsiifilised. Need tulemused kinnitasid, et *Drosophila da* sobib *TCF4*-ga seotud haiguste uurimiseks.

Teiseks kasutasime PTHS-i ja muude *TCF4*-ga seotud intellektipuuete mudelite loomiseks RNA interferentsipõhist *da* vaigistamist *Drosophila melanogaster*'i närvisüsteemis ja kahte käitumuslikku analüüsi – vastsete isu-seoselist õppimist ja täiskasvanud kärbeste negatiivset geotaksist. Meie tulemused näitasid, et vähenenud Da ekspressioon närvisüsteemis kahjustab vastsete mälu ja täiskasvanute liikumisvõimet. Lisaks leidsime, et Da osaleb sünaptilise proteoomi reguleerimisel. Järgmisena valideerisime, et *Drosophila melanogaster'*i mudelid sobivad ravimite sõeluuringuteks. Meie *TCF4*-ga seotud intellektipuude mudelite fenotüüpe parandasid kaks ravimit –

suberoüülaniliidhüdroksaamhape (SAHA) ja resveratrool. Lisaks suurendasid need ravimid nii Da kui ka TCF4 transaktivatsioonivõimet.

Kolmandaks iseloomustasime süstemaatiliselt *Tcf4* transkriptide ja valgu isovormide ekspressiooni kahes kõige laialdasemalt kasutatavas loommudelis – hiires ja rotis. Lisaks uurisime *TCF4* ekspressiooni inimeses. Leidsime, et kuigi *TCF4* geenistruktuur on keeruline ja võimaldab ekspresseerida palju erinevaid transkripte, kodeerivad nii närilistel kui ka inimestel peamised transkriptid TCF4-A, TCF4-D, TCF4-B, TCF4-C ja TCF4-I valgu isovorme. *TCF4* ekspressioon on kõrgeim ajus, kuid see ekspresseerub ka paljudes mitteneuraalsetes kudedes.

Käesolevas väitekirjas loodud ja iseloomustatud *TCF4*-ga seotud haiguste mudeleid saab kasutada PTHS-i ja muude *TCF4*-ga seotud haiguste ravi leidmiseks. *Drosophila melanogaster'*i mudelitel on suur potentsiaal ravimite sõeluurimisel ja avastamisel ning teadmised isovormi ja koespetsiifilise *TCF4* ekspressiooni kohta on geeniteraapia arendamiseks üliolulised.

Appendix

Publication I

Tamberg, L., Sepp, M., Timmusk, T. and Palgi, M.
Introducing Pitt-Hopkins syndrome-associated mutations of TCF4 to *Drosophila* daughterless.
Biology Open 4, 2015; 1762–1771. doi:10.1242/bio.014696.

© 2015. Published by The Company of Biologists Ltd | Biology Open (2015) 4, 1762-1771 doi:10.1242/bio.014696

RESEARCH ARTICLE

Introducing Pitt-Hopkins syndrome-associated mutations of *TCF4* to *Drosophila daughterless*

Laura Tamberg, Mari Sepp, Tõnis Timmusk and Mari Palgi*

ABSTRACT

Pitt-Hopkins syndrome (PTHS) is caused by haploinsufficiency of *Transcription factor 4 (TCF4)*, one of the three human class I basic helix-loop-helix transcription factors called E-proteins. *Drosophila* has a single E-protein, Daughterless (Da), homologous to all three mammalian counterparts. Here we show that human TCF4 can rescue Da deficiency during fruit fly nervous system development. Overexpression of Da or TCF4 specifically in adult flies significantly decreases their survival rates, indicating that these factors are crucial even after development has been completed. We generated *da* transgenic fruit fly strains with corresponding missense mutations R578H, R580W, R582P and A614V found in TCF4 of PTHS patients and studied the impact of these mutations *in vivo*. Overexpression of wild type Da as well as human TCF4 in progenitor tissues induced ectopic sensory bristles and the rough eye phenotype. By contrast, overexpression of Da^{R580W} and Da^{R582P} that disrupt DNA binding reduced the number of bristles and induced the rough eye phenotype with partial lack of pigmentation, indicating that these act dominant negatively. Compared to the wild type, Da^{R578H} and Da^{A614V} were less potent in induction of ectopic bristles and the rough eye phenotype, respectively, suggesting that these are hypomorphic. All studied PTHS-associated mutations that we introduced into Da led to similar effects *in vivo* as the same mutations in TCF4 *in vitro*. Consequently, our *Drosophila* models of PTHS are applicable for further studies aiming to unravel the molecular mechanisms of this disorder.

KEY WORDS: Pitt-Hopkins syndrome, *Drosophila melanogaster*, Intellectual disability, Daughterless, bHLH, Nervous system

INTRODUCTION

Pitt-Hopkins syndrome (PTHS, OMIM #610954) is a rare human disorder characterised by severe developmental delay, autistic behaviours, absence of speech, distinct facial features, epilepsy, constipation and hyperventilation (Pitt and Hopkins, 1978; Whalen et al., 2012). PTHS is caused by haploinsufficiency of the *Transcription factor 4* (*TCF4*, located at 18q21.1, OMIM #602272) (Amiel et al., 2007; Brockschmidt et al., 2007; Zweier et al., 2007). Large chromosomal deletions, partial gene deletions, frame shift, nonsense, splice site or missense mutations in the *TCF4* gene have been found in PTHS patients. These mutations are usually sporadic, but in some cases children have inherited the mutant allele from a mosaic parent (reviewed in Sweatt, 2013). *In vitro,* PTHS-associated missense mutations result in hypomorphic, non-functional or dominant-negative *TCF4* alleles (Sepp et al., 2012). It is unclear whether mutations causing PTHS impair development of the nervous system or functioning of the adult central nervous system (CNS), or both. In addition to PTHS, *TCF4* is associated with several other human diseases such as schizophrenia, Fuchs' corneal endothelial dystrophy and primary sclerosing cholangitis (reviewed by Forrest et al., 2014).

TCF4 (previously also known as ITF2, SEF2 or E2-2) belongs to the family of class I basic helix-loop-helix (bHLH) transcription factors (Massari and Murre, 2000) and should be distinguished from T-cell factor 4 (TCF4/TCF7L2) involved in the Wnt signalling pathway. The bHLH transcription factors form a large evolutionarily conserved family with important roles in numerous developmental processes including neurogenesis, myogenesis, haematopoiesis, and sex determination. The highly conserved bHLH region mediates interaction with other bHLH proteins and specific binding to DNA. Class I bHLH proteins, also called the E-proteins, comprise the mammalian TCF3/E2A, TCF4, TCF12/HEB and the *Drosophila* Daughterless (Da) (Massari and Murre, 2000). They are widely expressed and form homo- or heterodimers with class II bHLH proteins to bind DNA at the Ephrussi box (E-box) sequence, CANNTG. Class II bHLH proteins (Achaete-Scute complex proteins, MyoD, Myogenin, Atonal family etc.) are expressed in a tissue-specific manner but are not capable of activating transcription without E-proteins. In this study, we used the only E-protein in *Drosophila*, Da, to study PTHS-related mutations in the fruit fly.

Being the sole E-protein in *Drosophila*, Da has multiple roles in development – in sex determination and oogenesis (Cronmiller and Cummings, 1993; Smith et al., 2002), neurogenesis (Caudy et al., 1988; Hassan and Vaessin, 1997; Vaessin et al., 1994), eye development (Bhattacharya and Baker, 2011; Lim et al., 2008; Tanaka-Matakatsu et al., 2014), intestine stem cell maintenance (Bardin et al., 2010), and mesoderm development (Castanon et al., 2001). Da can both homodimerise or form heterodimers with class II bHLH proneural proteins and regulate the establishment of neural precursors (Cabrera and Alonso, 1991; Murre et al., 1989; Smith and Cronmiller, 2001; Tanaka-Matakatsu et al., 2014; Troost et al., 2015; Zarifi et al., 2012). Both lack of Da and ubiquitous overexpression of Da result in embryonic lethality, showing the importance of proper dosage (Cronmiller and Cline, 1987; Giebel et al., 1997).

Previously, *Drosophila melanogaster* has been successfully used to model human neurodegenerative diseases (Sang and Jackson, 2005). Recently efforts have been made towards exploiting fruit fly to model neuropsychiatric diseases and intellectual disability disorders (O'Kane, 2011; van der Voet et al., 2014). However, so far there are no fruit fly models of PTHS. Being the only E-protein

Laboratory of Molecular Neurobiology, Department of Gene Technology, Tallinn University of Technology, Akadeemia Rd.15, Tallinn 12618, Estonia.

*Author for correspondence (mari.palgi@ttu.ee)

Received 3 September 2015; Accepted 3 November 2015

Biology Open (2015) 4, 1762-1771 doi:10.1242/bio.014696

in *Drosophila,* Da is probably a functional orthologue of all mammalian E-proteins. Here we prove that TCF4 is a true functional orthologue of Da capable of mediating neuronal development in *Drosophila*.

To recapitulate the PTHS in *Drosophila* we introduced four PTHS-associated mutations of TCF4 – R580W, R578H, R582P and A614V – into Da. Their transcriptional activation capability was compared and analysed *in vivo*. All these mutations caused similar defects as their counterparts in TCF4 *in vitro* ranging from hypomorphic to dominant-negative effects.

Finally we show that activation of *da* transgenes with and without PTHS mutations in adult *Drosophila* leads to a reduced lifespan ranging from a few days to couple of weeks. These results implicate that in addition to their roles in development, expression of E-proteins has to be regulated in a spatially and temporally restricted manner also in adult flies.

RESULTS

Daughterless is the only orthologue of human E-proteins in *Drosophila*

Amino acid sequence analysis showed that Da is about 35% identical to human E-proteins TCF3, TCF4, and TCF12 and has the highest identity with TCF4 (35.54%). Even though the entire amino acid sequence homology between human E-proteins and Da is below 50%, the amino acid identity of bHLH domains between Da and human E-proteins reaches 75% (Fig. 1) which allows extrapolation of mutations found in bHLH of TCF4 of PTHS patients into Da.

Seven out of nine PTHS-associated missense mutations in TCF4 are found in the bHLH region (Sepp et al., 2012). From these bHLH positioned conserved mutations we selected four R578H, R580W, R582P and A614V for introduction into Da (Fig. 1). We named all mutations after PTHS-associated mutations in TCF4, although the numeral positions in Da differ by 14 amino acids (R564H, R566W, R568P, and A600V respectively). Two additional mutations were generated in order to study the importance of amino acid position and specificity. The first of these is D515G (D501G in Da) in nonconserved region close to the mutation D535G found in a PTHS patient. The second is R580L (R566L in Da) where leucine replaces arginine 580 originally found to be mutated into tryptophan in a patient of PTHS. The obtained transgenic fly strains with PTHS mutations introduced to *da* (da^{PTHS}) were crossed to the da^{G32}-GAL4 line and verified to produce mutated transcripts (Fig. S1).

Mutated Daughterless proteins have variable transactivation capabilities in HEK293 cells in comparison to wild type protein

We and others have previously demonstrated that several PTHS-associated missense mutations impair the functions of TCF4 homodimers and/or heterodimers with ASCL1 *in vitro* (Zweier et al., 2007; de Pontual et al., 2009; Sepp et al., 2012; Forrest et al., 2012). Particularly, PTHS-associated mutations R578H, R580W and R582P abolish the DNA-binding and transactivational capacity of TCF4 homodimers and TCF4:ASCL1 heterodimers, whereas A614V mutation impairs the functions of TCF4 homodimer, but retains the activity of the TCF4:ASCL1 heterodimer (Sepp et al., 2012). To test the transactivation capability of Da proteins carrying the same PTHS-associated mutations and two additional mutations (D515G and R580L), we used luciferase reporter assay in human embryonic kidney-derived cell line (HEK293). Three out of seven constructs tested – Da^{wt}, Da^{D515G}, and Da^{A614V} – were capable of activating E-box controlled luciferase gene transcription (Fig. 2). Da^{wt} did not activate transcription from the reporter construct without E-boxes indicating that the transcriptional activation is E-box specific. The following mutations abolished reporter gene expression: R578H, R580W, R580L and R582P. In the case of Da^{A614V} the luciferase signal was lower compared to Da^{wt} whereas the control mutation D515G showed no effect on Da transactivation capability. These results indicate that arginines R578 and R580 in the basic region and R582 in the beginning of the first helix of Da bHLH domain are essential for activating E-box controlled transcription. The mutation A614V in the second helix shows diminished transactivation of transcription (Fig. 2). The above results are completely consistent with the results obtained with TCF4 proteins carrying the same mutations (Sepp et al., 2012).

TCF4 and Da, but not Da^{R578H}, Da^{R580W}, Da^{R580L} or Da^{R582P}, activate E-box controlled reporter gene expression in wing disc and induce ectopic thoracic bristles in *Drosophila*

To analyse the functional consequences of Da mutations *in vivo*, we took the following approaches. First, we analysed the capacity of the overexpressed mutants to activate E-box dependent transcription in wing disc. Second, we tested the impact of the mutations on formation of ectopic sensory bristles induced by Da overexpression (Zarifi et al., 2012). To this end, transgenic *Drosophila* strains were generated expressing Da^{PTHS} under GAL4 control element UAS (Brand and Perrimon, 1993). Additionally, we generated flies with two most widely expressed alternative

```
TCF4        PNQVPVPQLPVQSATSPDLN-PPQDPYRGMPPGLQGQSVSSGSSEIKSDD-EGDENLQDT
TCF12       SSTV--------TTSSTDLNHKTQENYRG---GLQSQSGTVVTTEIKTENKEKDENLHEP
TCF3        HNHAA---LPSQPGTLPDLSRPP-DSYSG----LGRAGATAAASEIKREEKEDEENTSAA
Da          S-------LKLDRSASTSLPKQTKKRKEHTAISNSVPAGVSTTSSLTSLDISDTKPTSSI
                          : ..*   . .            .    ::.:.  : .  :

TCF4        KSSEDKKLDDD--KKDIK-----SITSNNDDEDLTPEQKAEREKERRMANNARERLRVRD
TCF12       PSSDDMKSDDESSQKDIKVSSRGRTSSTNEDEDLNPEQKIEREKERRMANNARERLRVRD
TCF3        DHSEEEK-------KELKAPRARTSPDEDEDDLLPPEQKAEREKERRVANNARERLRVRD
Da          ESSNS-GLQQHSQGKGTKRPR-RYCSSADEDDDAEPAVKAIREKERRQANNARERIRIRD
              *:.         *  *          . *:      *  *::*** *******:*:**

TCF4        INEAFKELGRMVQLHLKSDKPQTKLLILHQAVAVILSLEQQVRERNLNPKAACLKRREEE
TCF12       INEAFKELGRMCQLHLKSEKPQTKLLILHQAVAVILSLEQQVRERNLNPKAACLKRREEE
TCF3        INEAFKELGRMCQLHLNSEKPQTKLLILHQAVSVILNLEQQVRERNLNPKAACLKRREEE
Da          INEALKELGRMCMTHLKSDKPQTKLGILNMAVEVIMTLEQQVRERNLNPKAACLKRREEE
            ****::*****   **:*:*.**** **: ** **: **********************
```

Fig. 1. Human E-proteins and Drosophila Da show high amino acid sequence conservation of bHLH domains. The amino acid sequences of all three human E-proteins TCF3, TCF4, and TCF12 together with *Drosophila* Da bHLH domains with surrounding sequences are presented. The basic and loop regions are marked as light grey while the helices are coloured dark grey. Amino acid residues marked in red in TCF4 sequence are found mutated in PTHS patients and amino acid residues marked in red in Da sequence where mutations were introduced in this study.

Biology Open (2015) 4, 1762-1771 doi:10.1242/bio.014696

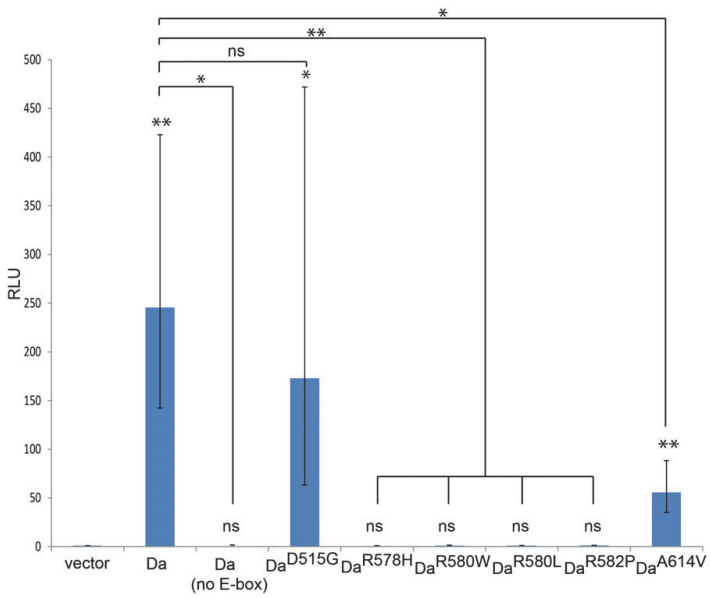


Fig. 2. E-box controlled luciferase reporter transcription activation capabilities of Da^{wt} and Da^{PTHS} *in vitro*. HEK293 cells were co-transfected with wt or mutated *da* construct, firefly luciferase construct carrying 12 µE5 boxes and a minimal promoter, or just a minimal promoter (no E-box), and *Renilla* luciferase construct with minimal promoter for normalisation. Luciferase activities were measured and data are presented as fold induced levels compared to the signals obtained from cells transfected with an empty vector. The mean results from three independent transfection experiments performed in duplicates are shown. Error bars show standard deviations. Statistical significance is shown with asterisks relative to empty-vector transfected cells, or between the groups connected with lines. *P<0.05, **P<0.01, ***P<0.001, ns, not significant, Student t-test; RLU, relative luciferase unit.

splice forms of *TCF4* – shorter *TCF4-A* isoform and the longer *TCF4-B* isoform (Sepp et al., 2011). *In vivo* lacZ reporter assay was performed using the transgenic flies with four E-boxes CATCTG upstream of lacZ reporter region as previously described (Culi, Modolell, 1998; Zarifi et al., 2012). UAS-transgenes were ectopically expressed under *pannier*-GAL4 (*pnr*-GAL4) in wing disc notum region (Fig. 3). Notum showed small spot-like areas of transgenic E-box activation by endogenous Da. Expression of GFP was used as a control showing no activation of reporter lacZ and mimicking transgenic E-box activation by endogenous Da (Fig. 3A). Importantly, this control did not induce ectopic bristles or thorax defects (Fig. 3F). Da^{wt} (Fig. 3B), Da^{D515G} (Fig. 3C), Da^{A614V} (Fig. 3M), and to a lesser extent TCF4-A (Fig. 3N) and TCF4-B (Fig. 3O) were able to activate reporter transcription from the E-box *in vivo*. Additionally, the same transgenes induced ectopic bristle formation on thorax (Fig. 3G,H,R-T). Arginine mutations R578H (Fig. 3D), R580W (Fig. 3E), R580L (Fig. 3K) and R582P (Fig. 3L) abolished the ability of Da to activate reporter expression in the wing disc. Interestingly, Da^{R580W} (Fig. 3J), Da^{R580L} (Fig. 3P) and Da^{R582P} (Fig. 3Q) reduced the number and size of thoracic bristles and caused malformation of the whole adult thorax indicating that these mutations share dominant-negative effects. Overexpression of Da^{R578H} (Fig. 3I) resulted in no major defects on the thorax, but on several occasions formation of some ectopic bristles was observed. Taken together, these results demonstrate that TCF4 isoforms TCF4-A and TCF4-B, Da^{wt}, Da^{D515G} and Da^{A614V} activate E-box controlled transcription in *Drosophila*. The arginine mutations R580W, R580L and R582P cause dominant negative effects while the mutation R578H considerably reduces the ability of Da to induce ectopic bristles.

Da, Da^{D515G}, Da^{R578H}, Da^{A614V}, TCF4-A and TCF4-B are capable of rescuing *da* null embryonic neuronal phenotype

Next we asked whether Da^{wt} or any of the mutants used is able to rescue *da* null lethality or severe embryonic nervous system phenotype with total lack of peripheral nervous system (PNS) and disrupted CNS (Caudy et al., 1988). For this experiment transgenic fruit fly strains expressing Da^{wt}, Da^{PTHS}, TCF4-A or TCF4-B under GAL4 responsive element were crossed to the GAL4 lines in *da* null background. First, we performed the rescue experiments using nervous system specific *GMR12B08*-GAL4 made of the only intron of *da* gene fused with *Drosophila* core synthetic promoter followed by GAL4 coding region (Pfeiffer et al., 2008). *GMR12B08*-GAL4 drives expression specifically in the nervous system in all developmental stages. This driver failed to rescue *da* null embryonic lethality with all our transgenes. The result obtained is consistent with known functions of Da outside the nervous system, for example in the mesoderm and muscle development (Castanon et al., 2001; Gonzalez-Crespo and Levine, 1993; Wong et al., 2008).

Subsequently, we repeated the rescue experiment with Da^{wt} under ubiquitous da^{G32}-GAL4 in *da* null background that led to embryonic lethality (Giebel et al., 1997; Smith and Cronmiller, 2001). Despite embryonic lethality, the severe nervous system phenotype i.e. peripheral nervous system absence was rescued. The nervous system of embryos from the rescue crosses with da^{G32}-GAL4 was visualised by immunohistochemistry using neuronal marker Futsch (*Drosophila* homologue to mammalian Microtubule associated protein 1B). Embryos homozygous for null mutant allele da^{10} lack the entire PNS and have defects in CNS (Fig. 4B) compared to wt embryos (Fig. 4A). Expressing Da^{wt} in *da* null background rescued the neuronal phenotype as reported before (Giebel et al., 1997; Smith and Cronmiller, 2001) (Fig. 4C). In embryos expressing Da^{D515G} (Fig. 4D), Da^{A614V} (Fig. 4I), TCF4-A (Fig. 4J) or TCF4-B (Fig. 4K) in *da* null background the neuronal phenotype was rescued as well. Interestingly, Da^{R578H} that was unable to activate E-box controlled reporter expression *in vitro* (Fig. 2) and in the wing disc (Fig. 3D), rescued the development of the embryonic nervous system (Fig. 4E). Da^{R580W} (Fig. 4F), Da^{R580L} (Fig. 4G) and Da^{R582P} (Fig. 4H) were unable to rescue the neuronal phenotype of *da* null mutants, which is consistent with their inability to activate transcription from E-box.

Biology Open (2015) 4, 1762-1771 doi:10.1242/bio.014696

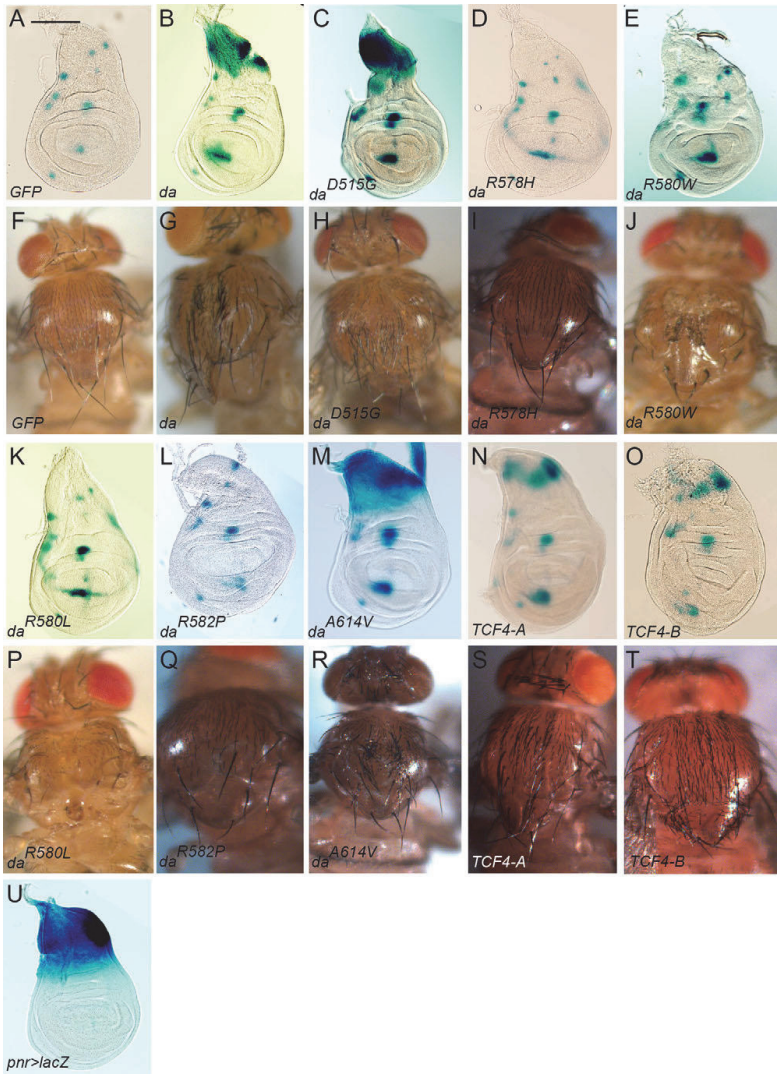


Fig. 3. *In vivo E-box*-lacZ reporter assay in larval wing discs and activation of bristle formation by overexpression of Da^{wt}, Da^{PTHS}, TCF4-A and TCF4-B. Expression of GFP under *pnr*-GAL4 served as a control in wing disc resembling endogenous activation of *E-box*-lacZ reporter expression (A) and shows no deviation from wild type bristles and thorax (F). Da^{wt} (B), Da^{D515G} (C), Da^{A614V} (M), TCF4-A (N) and TCF4-B (O) were able to activate E-box controlled lacZ expression in *pnr*-GAL4 pattern and caused ectopic bristle formation on adult thorax (G,H,R-T). Da^{R578H} (D), Da^{R580W} (E), Da^{R580L} (K) and Da^{R582P} (L) were unable to activate reporter expression. Da^{R580W} (J), Da^{R580L} (P) and Da^{R582P} (Q) caused thorax malformation, Da^{R578H} (I) had no effect on thorax but induced occasional extra bristles. The expression pattern of *pnr*-GAL4 is shown in blue on panel U. Scale bar on A represents 100 µm and is applicable for A-E,K-O and U.

Because rescue of *da* null mutants with expression of Da^{wt} using da^{G32}-GAL4 driver failed, other GAL4 strains with broad expression like *tub*-GAL4, *69B*-GAL4 and *ubi*-GAL4 were used for rescue experiments. However, the rescue of *da* null embryonic lethality using these drivers was unsuccessful (data not shown). These results demonstrate that the exact dosage and spatial and temporal regulation of Da protein expression are highly important for *Drosophila* viability.

Overexpression of Da, Da^{PTHS}, TCF4-A, or TCF4-B by *GMR12B08*-GAL4 results in the rough eye phenotype

Overexpression of Da^{wt}, Da^{PTHS}, TCF4-A or TCF4-B under ubiquitous da^{G32}-GAL4 driver resulted in embryonic lethality. Overexpressing these transgenes under the control of the nervous-system specific *GMR12B08*-GAL4 resulted in viable flies with the rough eye phenotype (Fig. 5). During eye development, this driver line is weakly expressed in larval eye discs and strongly in larval optic lobes (Pfeiffer et al., 2008 and our unpublished results). Overexpression of Da^{wt} (Fig. 5B,B′), Da^{D515G} (Fig. 5C,C′), Da^{R578H} (Fig. 5D,D′), TCF4-A, (Fig. 5M,M′) or TCF4-B (Fig. 5N,N′) resulted in the rough eye phenotype only when *GMR12B08*-GAL4 driver was homozygous. Exceptionally, Da^{A614V} overexpression under *GMR12B08*-GAL4 led to the rough eye phenotype only in double homozygous state (Fig. 5K,K′,L,L′). *TCF4* and other *da* transgenes except da^{R580W} remained heterozygous since double homozygous flies never survived into adulthood. Additionally, overexpression of Da^{R580W} (Fig. 5E,E′), Da^{R580L} (Fig. 5G,G′) and Da^{R582P} (Fig. 5I,I′) by the same driver resulted in partial loss of eye pigmentation. This phenotype was strongest in the case of Da^{R580W} female flies, with insertion of the

Biology Open (2015) 4, 1762-1771 doi:10.1242/bio.014696

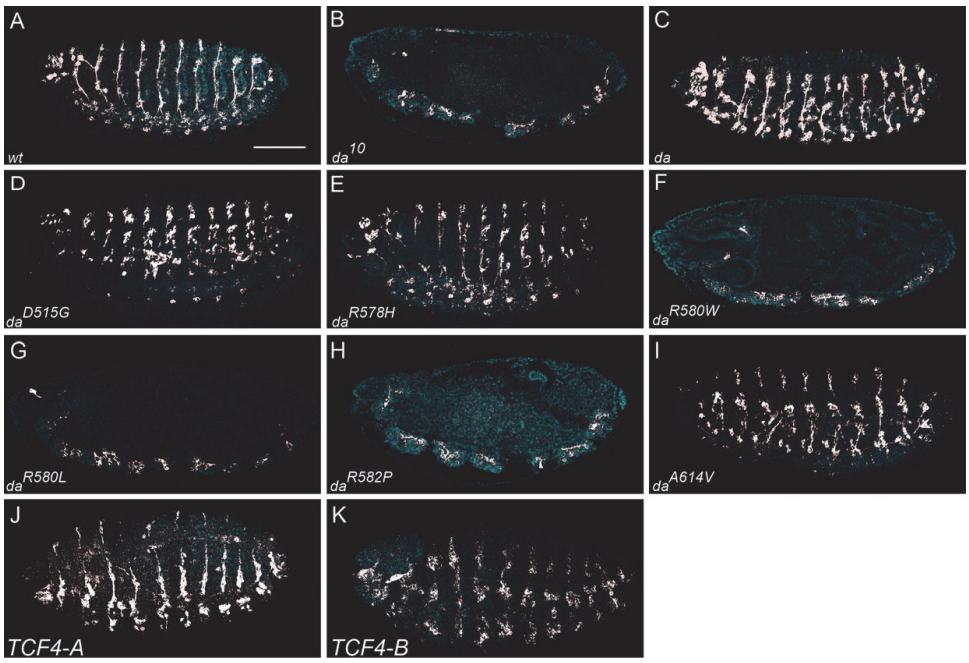


Fig. 4. Rescue of *da* null embryonic neuronal phenotype by expressing Da^{wt}, Da^{PTHS}, TCF4-A or TCF4-B under da^{G32}-GAL4. Neuronal marker Futsch expression is shown in white. DAPI, in light blue, marks nuclei. Compared to *wt* embryos (A), homozygous da^{10} embryos lack PNS and CNS is discontinuous (B). Overexpression of Da^{wt} (C), Da^{D515G} (D), Da^{R578H} (E), Da^{A614V} (I), TCF4-A (J) and TCF4-B (K) under da^{G32}-GAL4 rescued peripheral nervous system. All three arginine mutations Da^{R580W} (F), Da^{R580L} (G) and Da^{R582P} (H) failed to rescue *da* null nervous system phenotype as the CNS is still fragmented and PNS missing. Anterior side is to the left and dorsal side is up. Scale bar represents 100 µm.

transgene in the X-chromosome making them double homozygous. Interestingly, the rough eye phenotype was also weaker in the males of other transgenic lines compared to females carrying arginine mutations (R580L and R582P) with transgene insertions in the second chromosome (Fig. 5H,H′,J,J′). These results demonstrate high sensitivity of eye development to *da* transgene dosage. A614V mutation yielded the most subtle rough eye phenotype and all three arginine mutations R580W, R580L and R582P produced the strongest rough eye phenotype. Our results also showed that Da^{R578H}, Da^{D515G}, TCF4-A and TCF4-B behave similarly to Da^{wt} when overexpressed under *GMR12B08*-GAL4, indicating that all the mentioned proteins have transactivation activity during fruit fly eye development.

Overexpression of Da, TCF4-B and Da^{PTHS} in young adult flies results in significantly altered survivorship

Next, we wanted to evaluate the impact of Da in adult flies. For this the temperature sensitive repressor of GAL4, $GAL80^{ts}$ was used. The overexpression of Da under the control of the ubiquitous da^{G32}-GAL4 driver is lethal during embryogenesis (Giebel et al., 1997; Smith and Cronmiller, 2001). In order to overcome the lethality during development, we repressed GAL4 by $GAL80^{ts}$. After the eclosion of adults, the collected virgin females were kept at restrictive temperature to inactivate $GAL80^{ts}$. Activated overexpression of Da^{wt} or Da^{D515G} or *TCF4* long isoform B lead to lethality in 2-3 days (Fig. 6). The flies initially lost their flight ability and most of locomotor activity and died soon afterwards. Also, the activation of Da carrying PTHS-related arginine mutations unable to bind DNA (R580W, R580L, and R582P) lead to lethality in median 3-4 days (Fig. 6). Flies overexpressing the Da with weaker mutations, Da^{R578H} and Da^{A614V}, resulted in median survival of 10-11 days while the control group flies overexpressing GFP survived generally 40 days (Fig. 6; Table S1). Surprisingly, the flies overexpressing the shorter TCF4-A isoform survived significantly longer than the flies overexpressing TCF4-B and were closer to the control group, with median survival of 30 days (Fig. 6; Table S1). All survivorship curves obtained were statistically significant by log-rank as compared to the GFP control curve ($P<0.0001$, Mantel–Cox test).

DISCUSSION

In this study we show that Da, the only E-protein in *Drosophila* with highly conserved bHLH domain, functions as human TCF4 orthologue. As the overall identity of a protein sequence between *Drosophila* and mammals is usually around 40% between homologues and 80-90% within conserved functional domains (Pandey and Nichols, 2011), Da can be considered the orthologue for all three human E-proteins. In all experiments conducted in this study, TCF4 acted in very similarly as Da (Table 1), proving the possibility of modelling PTHS in the fruit fly. The two human TCF4 isoforms, TCF4-A and TCF4-B, were able to activate E-box dependent lacZ expression in *Drosophila*, and more importantly, to induce ectopic bristle formation in the adult thorax, to rescue embryonic nervous system development in *da* null embryos, and to induce the rough eye phenotype when overexpressed in the nervous system identically to Da. Altogether these results show that TCF4 has comparable activity in the fruit fly as Da.

Biology Open (2015) 4, 1762-1771 doi:10.1242/bio.014696

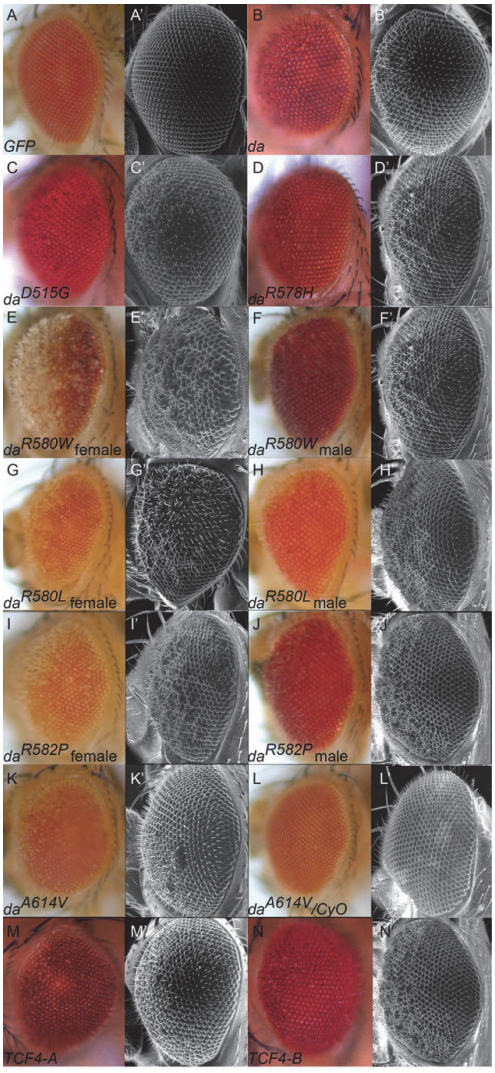


Fig. 5. Expression of Da^{wt}, Da^{PTHS}, TCF4-A and TCF4-B under neuronal driver *GMR12B08*-GAL4 results in rough eye phenotype. Rough eye phenotype was observed when *GMR12B08*-GAL4 was homozygous. Overexpression of GFP served as a control showing no deviation from wild type eye (A,A′). Da^{wt} (B,B′), Da^{D515G} (C,C′), Da^{R578H} (D,D′), TCF4-A (M,M′) and TCF4-B (N,N′) caused rough eye phenotype with irregularly placed and occasionally fused ommatidia closer to the head midline half of the eye. Da^{A614V} caused similar phenotype only in double homozygous state (K,K′), one copy of *UAS-da*^{A614V} was insufficient to induce the rough eye (L,L′). Da^{R580W} (E,E′), Da^{R580L} (G,G′) and Da^{R582P} (I,I′) resulted in eye pigmentation defects in addition to rough eye phenotype and the phenotype was enhanced in all females compared to males (F,F′,H,H′,J and J′).

To further study the PTHS-associated mutations in *Drosophila,* we introduced four mutations found in TCF4 (R580W, R578H, R582P and A614V) and two control mutations (D515G and R580L) into Da. We analysed the mutants by luciferase assay in a mammalian cell line in order to compare the results of Da directly to our previous results obtained with human TCF4 (Sepp et al., 2012). Subsequently we studied PTHS-associated Da mutants *in vivo* in E-box lacZ reporter assay, and in both rescue and overexpression experiments.

PTHS-associated arginine mutations R580W, R580L, and R582P abolished Da transactivation capability in luciferase reporter assays in HEK293 cells. Da^{R580W}, Da^{R580L} and Da^{R582P} behaved similarly to each other in both overexpression and rescue experiments in the fruit fly. The rescue by *da*^{G32}-GAL4 driver of *da* null embryonic nervous system phenotype failed when using Da proteins with these arginine mutations. When Da carrying one of above mentioned mutations was overexpressed in flies under the control of the nervous system specific driver *GMR12B08*-GAL4, the strongest eye phenotype was observed. These flies had rough and partially unpigmented eyes with fused ommatidia consistent with Da having an important role in *Drosophila* eye development (Brown et al., 1996). In addition, overexpression of these arginine mutants under *pnr*-GAL4 caused malformation of the thorax. Altogether, these results indicate that mutations R580W, R580L and R582P abolish the Da transactivation capability resulting in dominant-negative effects. This is in line with the previous data about the corresponding mutations in TCF4 having dominant-negative effects *in vitro* (Sepp et al., 2012).

R578H differed from the other three arginine mutations (R580W, R580L and R582P) in *in vivo* experiments. Although Da^{R578H} was unable to activate reporter gene expression in luciferase assay carried out in mammalian cell line HEK293 and in lacZ assay *in vivo*, it caused rough eye phenotype similar to Da^{wt} when overexpressed by *GMR12B08*-GAL4. Furthermore, Da^{R578H} rescued *da* null embryonic neuronal phenotype when expressed using *da*^{G32}-GAL4. Also Da^{R578H} showed weak induction of ectopic bristles. Taken together these results indicate that transactivation capability of Da^{R578H} probably depends on its dimerisation partners, which could be lacking in mammalian cell line and weakly presented in the wing disc notum. Similarly, we have previously found that while TCF4 carrying the R578H mutation is unable to bind to E-box *in vitro* as a homodimer or in complex with either ASCL1 or NEUROD2, it does not act in dominant negative manner in reporter assays in mammalian cells (Sepp et al., 2012).

The A614V mutation positioned in the second helix of the bHLH domain showed the mildest effects. Da^{A614V} was able to activate E-box-specific transcription *in vitro* and *in vivo*. Expressing Da^{A614V} using *da*^{G32}-GAL4 rescued *da* null embryonic neuronal phenotype. Overexpression using *GMR12B08*-GAL4 resulted in the rough eye phenotype only when both of the transgenes were homozygous, indicating that this mutation causes hypomorphic effects. This is consistent with our recent study which showed that the A614V mutation leads to lower levels of TCF4 because of reduced protein stability (Sepp et al., 2012).

The control mutation generated by us, D515G, did not reduce Da transactivation capability *in vitro* and behaved similarly to Da^{wt} *in vivo*. This shows that D515 positioned outside of the conserved bHLH is not required for Da transcriptional activity. The other control mutation generated by us, R580L, where the same arginine was mutated as in Da^{R580W}, led to dominant-negative effects *in vivo* similarly to R580W. At least in the case of R580, the mutation specificity, whether it was mutated into tryptophan or leucine, made no difference in our study.

In rescue experiments with tested driver strains (*69B*-GAL4, *tub*-GAL4, *ubi*-GAL4, *GMR12B08*-GAL4, *da*^{G32}-GAL4) all Da transgenes failed to rescue *da* null embryonic lethality. Apparently

Biology Open (2015) 4, 1762-1771 doi:10.1242/bio.014696

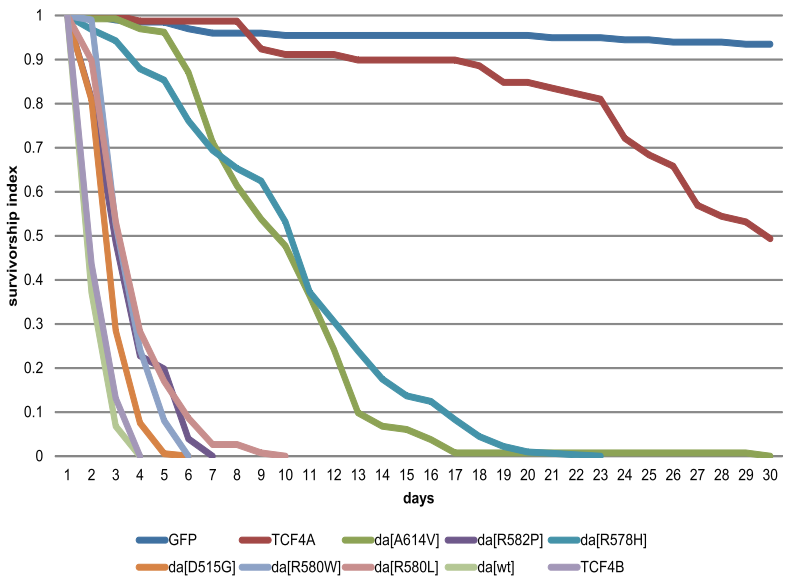


Fig. 6. Survivorship of adults after activation of Da^{wt}, Da^{PTHS}, or TCF4-B overexpression under da^{G32}-GAL4 is significantly altered. Using *tub*-$GAL80^{ts}$ repressor during development allows activation of UAS-*da* transgenes expression under da^{G32}-GAL4 only after eclosion of adults. The *x*-axis represents days from eclosion on day 0 to day 30, the *y*-axis represents survivorship index. Different UAS-transgenes are colour-coded. *tub*-$GAL80^{ts}$/+; da^{G32}-GAL4/UAS-*m*CD8-GFP flies were used as a control group.

the successful rescue of *da* null lethality closely mimics the endogenous Da expression. da^{G32}-GAL4, comprising of 3.2 kb of *da* gene covering the promoter, the first intron, and the upstream noncoding region (Wodarz et al., 1995), is widely used as a ubiquitous driver line. Most probably the expression of this driver line is far too strong compared to the native expression of *da* gene as Da has been shown to positively autoregulate its own expression via a transcriptional feedback loop (Bhattacharya and Baker, 2011; Smith and Cronmiller, 2001). If da^{G32}-GAL4 expression is regulated by Da itself, then Da overexpression might drive even stronger GAL4 expression, resulting in a positive feedback loop. Furthermore, it has been hypothesised that da^{G32}-GAL4 lacks putative regulatory repressor elements since using a 15 kb genomic *da* transgene that has an additional 12 kb of downstream sequence rescues *da* null embryonic lethality (Smith and Cronmiller, 2001).

Little is known about the role of E proteins in adult nervous system. Here we show that exact temporal and spatial expression of Da/TCF4 remains vitally important during adulthood of fruit flies. We show that overexpression of Da/TCF4 in adults leads to lethality within 2-3 days. Surprisingly, TCF4 isoforms A and B lead to strikingly different outcomes when overexpressed in adult fruit flies. While the long isoform TCF4-B behaved identically to Da, TCF4-A affected the survival only slightly compared to the control group. This could be related to the lack of interaction capability of much shorter N terminus of isoform A in fruit fly or different regulation of subcellular location and dimerisation of the alternative TCF4 isoforms (Sepp et al., 2011). Analysis of survival divided the PTHS related mutations into severe (R580W, R580L and R582P) and milder (R578H and A614V) according to survivorship. The severe mutants led to lethality within 3-4 days and the milder ones in

Table 1. Functional effects of transgenes used in this study

UAS construct	*In vitro* activation of transcription	*In vivo* activation of transcription in wing discs	Bristle induction by pnr-GAL4	Eye phenotype by homozygous *GMR12B08*-GAL4	Rescue of PNS of *da* null embryos	Adult median survival in days
da	yes	yes	yes	rough	yes	2
da^{D515G}	yes	yes	yes	rough	yes	3
da^{R578H}	no	no	weak	rough	yes	11
da^{R580W}	no	no	reduced bristles	rough, partly discoloured	no	4
da^{R580L}	no	no	reduced bristles	rough, partly discoloured	no	4
da^{R582P}	no	no	reduced bristles	rough, partly discoloured	no	4
da^{A614V}	weaker	yes	yes	rough only when homozygous	yes	10
TCF4-A	yes	weaker	yes	rough	yes	30
TCF4-B	yes	weaker	yes	rough	yes	2

Da^{wt}, Da^{D515G}, TCF4-A and TCF4-B share similar properties in most experiments conducted. Da^{R580W}, Da^{R580L} and Da^{R582P} are incapable of transactivation but result in additional phenotypes in eye and thorax development, also the survivorship of adults is severely reduced. Both Da^{R578H} Da^{A614V} have milder effects in survivorship assay and capable of rescuing embryonic nervous system development. The activity of Da^{R578H} may depend on the dimerisation partner present in the developing eye and which is weakly presented in developing thorax Da^{A614V} is a hypomorphic mutation allowing homodimerisation.

Biology Open (2015) 4, 1762-1771 doi:10.1242/bio.014696

10-11 days. Further experiments with cell type or tissue-specific drivers would help to understand the role of E-proteins during adulthood in more details.

The fact that overexpression of wild type as well as dominant negative forms of Da causes comparable reduction in survival and induction of the rough eye phenotype raises the possibility that overexpression of wt protein is also eliciting dominant negative effects as suggested earlier (Sweatt, 2013). One explanation for this phenomenon could be that excess homodimers outcompete transcriptionally more potent heterodimers at various promoter sites. Intriguingly, recent studies suggest that in addition to TCF4 haploinsufficiency, increased TCF4 dose is also a risk factor for disturbed cognitive development as a TCF4 duplication has been described in a patient with developmental delay (Talkowski et al., 2012) and a partial duplication in a patient with major depressive disorder (Ye et al., 2012). Nevertheless, in case of induction of ectopic bristles we observed opposite effects for Da^{wt} and dominant negative Da mutants, indicating that in addition to its dominant negative effects, excess wt protein also has specific effects during development.

In patients with PTHS just one copy of *TCF4* is mutated or deleted. Seemingly the most relevant way to model PTHS in animal models would be to use the appropriate heterozygotes of the orthologous protein. However, in *Drosophila* there is a sole E-protein Da corresponding to all three mammalian E-proteins. In a way the heterozygous Da null mutation corresponds to the heterozygous deletion of all three E-proteins in mammals. Accordingly, Da as the only binding partner of class II bHLH proteins has a large variety of roles outside nervous system. As TCF4 is highly expressed in the nervous system we have chosen here the approach to overexpress the mutated alleles specifically in the nervous system in a wild type background. Overexpression of Da^{PTHS} under the nervous system specific *GMR12B08*-GAL4 led to viable flies and we were able to create stocks with each mutation generated in this study. An alternative tactic to model PTHS and to mimic dosage loss by TCF4 deletions would be to slightly downregulate Da expression nervous system specifically by RNAi. Additional studies are needed to generate and compare different PTHS models and to perform behavioural tests that would give valuable information about cognition and social behaviour of the PTHS model flies.

In conclusion, this study is the first where experiments with PTHS-associated mutations have been performed *in vivo*. We have verified Da as a functional TCF4 homologue, described similarities between Da and TCF4 carrying the same mutations, and obtained insights how PTHS-associated mutated Da genes could affect *Drosophila* embryonic nervous system development and thoracic bristle formation. The similarities between the effects of PTHS-associated mutations on Da and TCF4, ranging from hypomorphic to dominant-negative, prove that these proteins have similar functions and Da can be used for modelling of PTHS in *Drosophila melanogaster*. Our novel models of PTHS in *Drosophila* allow the design of further studies addressing the molecular mechanisms and treatment of PTHS.

MATERIALS AND METHODS

Drosophila stocks

All *Drosophila* stocks and crosses were kept on malt and semolina based food with 12 h light and dark daily rhythms at 25°C with 60% humidity unless otherwise noted. *Drosophila* strains used in this study were da^{G32}-GAL4 (Wodarz et al., 1995), *69B*-GAL4 (Brand and Perrimon, 1993), and *tub*-GAL4 provided by Riitta Lindström, *ubi*-GAL4 provided by Mari Teesalu, *EE4-lacZ;pnr-GAL4/T(2,3)SM6-TM6B* kindly provided by Christos Delidakis, *GMR12B08*-GAL4 (Pfeiffer et al., 2008), da^{10},*FRT40A* (BL#5531), *UAS-*da^{G} (BL#37291), *UAS-mCD8-GFP* (BL#5137), and *tub*-$GAL80^{ts}$ (BL#7019) from Bloomington Stock Center at Indiana University, USA. The following transgenic lines were generated in this study: UAS-da^{D515G}, UAS-da^{R578H}, UAS-da^{R580W}, UAS-da^{R580L}, UAS-da^{R582P}, UAS-da^{A614V}, UAS-*TCF4-B*, and UAS-*TCF4-A*.

Mutagenesis, cloning and transgenesis

The amino acid sequences were aligned and homology of human E-proteins and Da was calculated using Clustal Omega 2.1 (EMBL-EBI, Cambridge, UK). Site-directed mutagenesis was performed using the partial *da* cDNA construct GH10651-pOT_2 as a template (Drosophila Genomics Resource Center, Bloomington, IN, USA). Primers were designed with that *de novo* restriction sites created next to the mutation with no change in amino acid sequence. The primer sequences are listed in Table S2. The constructs obtained by PCR were sequenced and subcloned into full-length *da* cDNA construct LP14713-pOT_2 (Drosophila Genomics Resource Center) using SphI and NruI restriction sites. Full-length *da* carrying appropriate PTHS-associated mutation together with TCF4-A(-) and TCF4-B(-) cDNAs from the respective pCDNA3.1 vectors (Sepp et al., 2011) were then subcloned into pUAST vector. Generation of transgenic flies with random insertions was ordered from Fly Facility (Clermont-Ferrand Cedex, France).

DNA transfection and luciferase assay

The transfection and luciferase assay was performed as described before (Sepp et al., 2011). Briefly, HEK293 cells obtained from ATCC (LGC Standards GmbH, Wesel, Germany) and routinely tested for contamination were transfected using LipoD293™ (SignaGen Laboratories, Gaithersburg, MD, USA) with pCDNA3.1 based *TCF4* or *da* constructs and firefly luciferase construct pGL4.29[luc2P/12µE5/Hygro] or pGL4.29[luc2P/min/Hygro] and *Renilla* luciferase construct pGL4.29[hRlucP/min/Hygro] for normalisation. Transfections were carried out as duplicates on a 48-well plate. After 24 h cells were lysed with 50 µl Passive Lysis Buffer (Promega, Madison, Wisconsin, USA). Dual-Glo Luciferase assay (Promega) was performed following manufacturer's protocol and luminescence was measured with GENios Pro Multifunction Microplate Reader (Tecan Group, Männedorf, Switzerland). For data analysis, background signals from untransfected cells were subtracted and firefly luciferase signals were normalised to *Renilla* luciferase signals. The data was then log-transformed, auto-scaled, means and standard deviations were calculated and Student *t*-tests were performed. The data was back-transformed for graphical representation.

In vivo lacZ reporter assay

For expression of transgene in wing discs, ubiquitous temperature-sensitive *tub*-$GAL80^{ts}$ repressor was used to avoid larval lethality caused by *pnr*-GAL4>UAS-*da*. Vials with 3rd instar larvae were transferred from permissive temperature at 18°C to restrictive temperature at 30°C for 24 h before dissection of wing discs and afterwards moved to 25°C until adults emerged. Each *da* mutant strain under UAS was crossed to *EE4-lacZ; pnr-GAL4/T(2,3)SM6-TM6B*. 3rd instar wing discs of their progeny were dissected and the X-gal staining was performed. Expression of transgene in wing discs and X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) histochemistry was performed as described before (Zarifi et al., 2012). For imaging, Olympus BX61 microscope (Tokyo, Japan) with UPlanSApo 20×/0.75 objective was used.

RNA isolation and RT-PCR

Total RNA was isolated from embryos using RNeasy Micro Kit (Qiagen, Hilden, Germany) and treated with TURBO DNase (Ambion, Thermo Fisher Scientific, Waltham, MA, USA). First strand cDNA was reverse transcribed from 1 µg of RNA using oligo$(dT)_{20}$ primer and Superscript III Reverse Transcriptase (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA). PCR was performed using FirePol DNA polymerase (Solis Biodyne, Tartu, Estonia). Primer annealing temperature was 55°C and

primer sequences are presented in Table S2. After PCR, restriction analysis was performed using restriction sites created during mutagenesis (Fig. S1).

Immunohistochemical staining of embryos

The following primary antibodies and dilutions were used: rabbit EGFP antiserum (provided by Andres Merits, Tartu University), 1:2000; mouse monoclonal 22C10 anti-Futsch antibody (deposited by Benzer, Seymour/Colley, Nansi, obtained from the Developmental Studies Hybridoma Bank, University of Iowa, IA, USA), 1:20. The following secondary antibodies and dilutions were used: goat anti-mouse Alexa 594, 1:1000; and goat anti-rabbit Alexa 488 (both ImmunoResearch Laboratories, Inc., West Grove, PA, USA), 1:1000. *Drosophila* embryos were dechorionated using 2% sodium hypochlorite (Sigma-Aldrich, St. Louis, MO, USA) and fixed using 4% formaldehyde (AppliChem GmbH, Darmstadt, Germany) in PEM buffer (100 mM PIPES, 1 mM EGTA, 2 mM $MgSO_4$, pH 7.4) and stored in methanol at −20°C until used. Primary antibody labelling was performed overnight on overhead rotator at 4°C. Secondary antibodies were preadsorbed to wt embryos before use. Incubation with secondary antibodies was performed for 3 h on overhead rotator at room temperature. All staining procedures and washes were performed with 0.1% PBS-Triton X-100. Fluorescently labelled embryos were mounted in Vectashield mounting medium with DAPI (Vector Laboratories, Burlingame, CA, USA). For image collection, Zeiss LSM 510 Meta confocal microscope with Pln Apo 20×/0.8 DICII objective (Carl Zeiss Microscopy GmbH, Jena, Germany) was used. Suitable layers were selected using Zeiss LSM Image Browser (Carl Zeiss Microscopy). Homozygous da^{10} embryos were selected by the lack of GFP marker expression present in balancer chromosome.

Light microscopy and imaging of adult flies

Flies were euthanised using chloroform (Sigma-Aldrich). For capturing eye and thorax images, Zeiss Stereo Discovery V8 microscope and Zeiss Axiocam MRc camera were used. For scanning electron microscopy, Zeiss Evo LS15 (Carl Zeiss Microscopy) was used.

Evaluating the lifespan of *da* transgenic flies

w^{-}; *tub*-GAL80^{ts}; da^{G32}-GAL4 flies were crossed to *da* transgenic flies with insertions in second chromosome (lines UAS-da^{D515G1}, UAS-da^{R578H3}, UAS-da^{R580L4}, UAS-da^{R582P2}, UAS-da^{A614V3}, UAS-*TCF4-B*^{A}) or in third chromosome (lines UAS-da^{R580W3}, UAS-da^{G}, UAS-*mCD8*-GFP, UAS-*TCF4-A*^{2}) and raised in permissive temperature 18°C where GAL80^{ts} represses the activity of GAL4. The progeny – eclosed virgin females – were collected and transferred to restrictive temperature 30°C where GAL80^{ts} is inactive and GAL4 is produced. Altogether, on average 200 flies (from at least 79 up to 314) were collected from each cross, maintained in uncrowded vials, counted on daily bases, and changed to new vials at three-day intervals. Survivorship index was calculated by dividing the survivors with the starting number of flies. Comparison of survival curves and *P*-value calculations were performed using Mantel–Cox log rank method with Prism 6.0 GraphPad Software (La Jolla, CA, USA).

Acknowledgements

We thank Christos Delidakis (University of Crete, Greece), Riitta Lindström and Mari Teesalu (University of Helsinki, Finland) for sharing fly stocks. We thank Heiti Paves for the help with TEM and imaging, and Epp Väli for technical assistance. We are grateful to Jaan Palgi, Jürgen Tuvikene, Indrek Koppel, and Richard Tamme for critical comments on manuscript. Other stocks used in this study were obtained from the Bloomington Drosophila Stock Center (NIH P40OD018537). We acknowledge the Drosophila Genomics Resource Center (NIH 2P40OD010949-10A1) and Developmental Studies Hybridoma Bank (created by the NICHD of the NIH and maintained at University of Iowa, Department of Biology, Iowa City, IA, USA).

Competing interests

The authors declare no competing or financial interests.

Author contributions

L.T., M.S., M.P. and T.T. conceived and designed the experiments, M.P. and T.T. provided research tools; L.T. and M.P. performed the experiments, L.T. and M.P. analysed the data, L.T. and M.P. wrote the paper. All authors commented and edited the manuscript.

Funding

This work was supported by Estonian Research Council [MJD341 to M.P.; IUT19-18 to T.T.], National R&D program 'Health' [AR12098 to T.T.], Pitt Hopkins Research Foundation (T.T. and M.S.), Estonian Academy of Sciences (T.T.).

Supplementary information

Supplementary information available online at
http://bio.biologists.org/lookup/suppl/doi:10.1242/bio.014696/-/DC1

References

Amiel, J., Rio, M., de Pontual, L., Redon, R., Malan, V., Boddaert, N., Plouin, P., Carter, N. P., Lyonnet, S., Munnich, A. et al. (2007). Mutations in TCF4, encoding a class I basic helix-loop-helix transcription factor, are responsible for Pitt-Hopkins syndrome, a severe epileptic encephalopathy associated with autonomic dysfunction. *Am. J. Hum. Genet.* **80**, 988-993.

Bardin, A. J., Perdigoto, C. N., Southall, T. D., Brand, A. H. and Schweisguth, F. (2010). Transcriptional control of stem cell maintenance in the Drosophila intestine. *Development* **137**, 705-714.

Bhattacharya, A. and Baker, N. E. (2011). A network of broadly-expressed HLH genes regulates tissue-specific cell fates. *Cell* **147**, 881-892.

Brand, A. H. and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401-415.

Brockschmidt, A., Todt, U., Ryu, S., Hoischen, A., Landwehr, C., Birnbaum, S., Frenck, W., Radlwimmer, B., Lichter, P., Engels, H. et al. (2007). Severe mental retardation with breathing abnormalities (Pitt–Hopkins syndrome) is caused by haploinsufficiency of the neuronal bHLH transcription factor TCF4. *Hum. Mol. Genet.* **16**, 1488-1494.

Brown, N. L., Paddock, S. W., Sattler, C. A., Cronmiller, C., Thomas, B. J. and Carroll, S. B. (1996). daughterless is required for Drosophila photoreceptor cell determination, eye morphogenesis, and cell cycle progression. *Dev. Biol.* **179**, 65-78.

Cabrera, C. V. and Alonso, M. C. (1991). Transcriptional activation by heterodimers of the achaete-scute and daughterless gene products of Drosophila. *EMBO J.* **10**, 2965-2973.

Castanon, I., Stetina, S. V., Kass, J. and Baylies, M. K. (2001). Dimerization partners determine the activity of the Twist bHLH protein during Drosophila mesoderm. *Development* **128**, 3145-3159.

Caudy, M., Grell, E. H., Dambly-Chaudiere, C., Ghysen, A., Jan, L. Y. and Jan, Y. N. (1988). The maternal sex determination gene daughterless has zygotic activity necessary for the formation of peripheral neurons in Drosophila. *Genes Dev.* **2**, 843-852.

Cronmiller, C. and Cline, T. W. (1987). The Drosophila sex determination gene daughterless has different functions in the germ line versus the soma. *Cell* **48**, 479-487.

Cronmiller, C. and Cummings, C. A. (1993). The daughterless gene product in Drosophila is a nuclear protein that is broadly expressed throughout the organism during development. *Mech. Dev.* **42**, 159-169.

Culi, J. and Modolell, J. (1998). Proneural gene self-stimulation in neural precursors: an essential mechanism for sense organ development that is regulated by Notch signaling. *Genes Dev.* **12**, 2036-2047.

de Pontual, L., Mathieu, Y., Golzio, C., Rio, M., Malan, V., Boddaert, N., Soufflet, C., Picard, C., Durandy, A., Dobbie, A. et al. (2009). Mutational, functional, and expression studies of the TCF4 gene in Pitt-Hopkins syndrome. *Hum. Mutat.* **30**, 669-676.

Forrest, M. P., Chapman, R. M., Doyle, A. M., Tinsley, C. L., Waite, A. and Blake, D. J. (2012). Functional analysis of TCF4 missense mutations that cause Pitt-Hopkins syndrome. *Hum. Mutat.* **33**, 1676-1686.

Forrest, M. P., Hill, M. J., Quantock, A. J., Martin-Rendon, E. and Blake, D. J. (2014). The emerging roles of TCF4 in disease and development. *Trends Mol. Med.* **20**, 322-331.

Giebel, B., Stüttem, I., Hinz, U. and Campos-Ortega, J. A. (1997). Lethal of scute requires overexpression of daughterless to elicit ectopic neuronal development during embryogenesis in Drosophila. *Mech. Dev.* **63**, 75-87.

Gonzalez-Crespo, S. and Levine, M. (1993). Interactions between dorsal and helix-loop-helix proteins initiate the differentiation of the embryonic mesoderm and neuroectoderm in Drosophila. *Genes Dev.* **7**, 1703-1713.

Hassan, B. and Vaessin, H. (1997). Daughterless is required for the expression of cell cycle genes in peripheral nervous system precursors of Drosophila embryos. *Dev. Genet.* **21**, 117-122.

Lim, J., Jafar-Nejad, H., Hsu, Y.-C. and Choi, K.-W. (2008). Novel function of the class I bHLH protein Daughterless in the negative regulation of proneural gene expression in the Drosophila eye. *EMBO Rep.* **9**, 1128-1133.

Massari, M. E. and Murre, C. (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol. Cell. Biol.* **20**, 429-440.

Murre, C., McCaw, P. S. and Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* **56**, 777-783.

O'Kane, C. J. (2011). Drosophila as a model organism for the study of neuropsychiatric disorders. *Curr. Top. Behav. Neurosci.* **7**, 37-60.
Pandey, U. B. and Nichols, C. D. (2011). Human disease models in Drosophila melanogaster and the role of the fly in therapeutic drug discovery. *Pharmacol. Rev.* **63**, 411-436.
Pfeiffer, B. D., Jenett, A., Hammonds, A. S., Ngo, T.-T. B., Misra, S., Murphy, C., Scully, A., Carlson, J. W., Wan, K. H., Laverty, T. R. et al. (2008). Tools for neuroanatomy and neurogenetics in Drosophila. *Proc. Natl. Acad. Sci. USA* **105**, 9715-9720.
Pitt, D. and Hopkins, I. (1978). A syndrome of mental retardation, wide mouth and intermittent overbreathing. *Aust. Paediatr. J.* **14**, 182-184.
Sang, T.-K. and Jackson, G. R. (2005). Drosophila models of neurodegenerative disease. *NeuroRx J. Am. Soc. Exp. Neurother.* **2**, 438-446.
Sepp, M., Kannike, K., Eesmaa, A., Urb, M. and Timmusk, T. (2011). Functional diversity of human basic Helix-Loop-Helix Transcription Factor TCF4 isoforms generated by alternative 5′ exon usage and splicing. *PLoS ONE* **6**, e22138.
Sepp, M., Pruunsild, P. and Timmusk, T. (2012). Pitt-Hopkins syndrome-associated mutations in TCF4 lead to variable impairment of the transcription factor function ranging from hypomorphic to dominant-negative effects. *Hum. Mol. Genet.* **21**, 2873-2888.
Smith, J. E., III and Cronmiller, C. (2001). The Drosophila daughterless gene autoregulates and is controlled by both positive and negative cis regulation. *Development* **128**, 4705-4714.
Smith, J. E., III, Cummings, C. A. and Cronmiller, C. (2002). Daughterless coordinates somatic cell proliferation, differentiation and germline cyst survival during follicle formation in Drosophila. *Development* **129**, 3255-3267.
Sweatt, J. D. (2013). Pitt-Hopkins Syndrome: intellectual disability due to loss of TCF4-regulated gene transcription. *Exp. Mol. Med.* **45**, e21.
Talkowski, M. E., Rosenfeld, J. A., Blumenthal, I., Pillalamarri, V., Chiang, C., Heilbut, A., Ernst, C., Hanscom, C., Rossin, E., Lindgren, A. M. et al. (2012). Sequencing chromosomal abnormalities reveals neurodevelopmental loci that confer risk across diagnostic boundaries. *Cell* **149**, 525-537.
Tanaka-Matakatsu, M., Miller, J., Borger, D., Tang, W.-J. and Du, W. (2014). Daughterless homodimer synergizes with Eyeless to induce Atonal expression and retinal neuron differentiation. *Dev. Biol.* **392**, 256-265.
Troost, T., Schneider, M. and Klein, T. (2015). A re-examination of the selection of the sensory organ precursor of the bristle sensilla of Drosophila melanogaster. *PLoS Genet.* **11**, e1004911.
Vaessin, H., Brand, M., Jan, L. Y. and Jan, Y. N. (1994). daughterless is essential for neuronal precursor differentiation but not for initiation of neuronal precursor formation in Drosophila embryo. *Development* **120**, 935-945.
van der Voet, M., Nijhof, B., Oortveld, M. A. W. and Schenck, A. (2014). Drosophila models of early onset cognitive disorders and their clinical applications. *Neurosci. Biobehav. Rev.* **46**, 326-342.
Whalen, S., Héron, D., Gaillon, T., Moldovan, O., Rossi, M., Devillard, F., Giuliano, F., Soares, G., Mathieu-Dramard, M., Afenjar, A. et al. (2012). Novel comprehensive diagnostic strategy in Pitt-Hopkins syndrome: clinical score and further delineation of the TCF4 mutational spectrum. *Hum. Mutat.* **33**, 64-72.
Wodarz, A., Hinz, U., Engelbert, M. and Knust, E. (1995). Expression of crumbs confers apical character on plasma membrane domains of ectodermal epithelia of Drosophila. *Cell* **82**, 67-76.
Wong, M.-C., Castanon, I. and Baylies, M. K. (2008). Daughterless dictates Twist activity in a context-dependent manner during somatic myogenesis. *Dev. Biol.* **317**, 417-429.
Ye, T., Lipska, B. K., Tao, R., Hyde, T. M., Wang, L., Li, C., Choi, K. H., Straub, R. E., Kleinman, J. E. and Weinberger, D. R. (2012). Analysis of copy number variations in brain DNA from patients with schizophrenia and other psychiatric disorders. *Biol. Psychiatry* **72**, 651-654.
Zarifi, I., Kiparaki, M., Koumbanakis, K. A., Giagtzoglou, N., Zacharioudaki, E., Alexiadis, A., Livadaras, I. and Delidakis, C. (2012). Essential roles of Da transactivation domains in neurogenesis and in E(spl)-mediated repression. *Mol. Cell. Biol.* **32**, 4534-4548.
Zweier, C., Peippo, M. M., Hoyer, J., Sousa, S., Bottani, A., Clayton-Smith, J., Reardon, W., Saraiva, J., Cabral, A., Göhring, I. et al. (2007). Haploinsufficiency of TCF4 causes syndromal mental retardation with intermittent hyperventilation (Pitt-Hopkins syndrome). *Am. J. Hum. Genet.* **80**, 994-1001.

Publication II

Tamberg, L., Jaago, M., Säälik, K., Sirp, A., Tuvikene, J., Shubina, A., Kiir, C. S., Nurm, K., Sepp, M., Timmusk, T and Palgi, M.
Daughterless, the *Drosophila* orthologue of TCF4, is required for associative learning and maintenance of the synaptic proteome.
Disease Models and Mechanisms 13, 2020; dmm042747. doi:10.1242/dmm.042747.

© 2020. Published by The Company of Biologists Ltd | Disease Models & Mechanisms (2020) 13, dmm042747. doi:10.1242/dmm.042747

RESEARCH ARTICLE

Daughterless, the *Drosophila* orthologue of TCF4, is required for associative learning and maintenance of the synaptic proteome

Laura Tamberg[1,*], Mariliis Jaago[1,2], Kristi Säälik[1], Alex Sirp[1], Jürgen Tuvikene[1,2], Anastassia Shubina[1], Carl Sander Kiir[1], Kaja Nurm[1], Mari Sepp[1], Tõnis Timmusk[1,2] and Mari Palgi[1,*]

ABSTRACT

Mammalian transcription factor 4 (TCF4) has been linked to schizophrenia and intellectual disabilities, such as Pitt–Hopkins syndrome (PTHS). Here, we show that similarly to mammalian TCF4, fruit fly orthologue Daughterless (Da) is expressed widely in the *Drosophila* brain. Furthermore, silencing of *da*, using several central nervous system-specific Gal4 driver lines, impairs appetitive associative learning of the larvae and leads to decreased levels of the synaptic proteins Synapsin (Syn) and Discs large 1 (Dlg1), suggesting the involvement of Da in memory formation. Here, we demonstrate that *Syn* and *dlg1* are direct target genes of Da in adult *Drosophila* heads, as Da binds to the regulatory regions of these genes and the modulation of Da levels alter the levels of *Syn* and *dlg1* mRNA. Silencing of *da* also affects negative geotaxis of the adult flies, suggesting the impairment of locomotor function. Overall, our findings suggest that Da regulates *Drosophila* larval memory and adult negative geotaxis, possibly via its synaptic target genes *Syn* and *dlg1*. These behavioural phenotypes can be further used as a PTHS model to screen for therapeutics.

This article has an associated First Person interview with the first author of the paper.

KEY WORDS: TCF4, Daughterless, Pitt–Hopkins syndrome, Intellectual disability, *Drosophila melanogaster*, Appetitive associative learning, Negative geotaxis

INTRODUCTION

Transcription factor 4 (TCF4, also known as ITF2, E2-2, SEF2, etc.) belongs to the family of class I basic helix-loop-helix (bHLH) transcription factors, also called E-proteins (Murre et al., 1994). E-proteins bind to the DNA Ephrussi box (E-box) sequence CANNTG as homodimers or heterodimers with class II bHLH transcription factors (Cabrera and Alonso, 1991). TCF4 should be distinguished from TCF7L2, a downstream effector of the Wnt signalling pathway that is also referred to as TCF4 (T cell factor 4). TCF4 is essential for a range of neurodevelopmental processes including early spontaneous neuronal activity, cell survival, cell cycle regulation, neuronal migration and differentiation, synaptic plasticity and memory formation (Chen et al., 2016; Crux et al., 2018; Forrest et al., 2013; Hill et al., 2017; Jung et al., 2018; Kennedy et al., 2016; Kepa et al., 2017; Li et al., 2019; Page et al., 2018; Thaxton et al., 2018). Genes involved in pathways including nervous system development, synaptic function and axon development are TCF4 targets (Forrest et al., 2018; Xia et al., 2018). Furthermore, TCF4 regulates the expression of ion channels $Na_v1.8$ and $K_v7.1$ (Ekins et al., 2019; Rannals et al., 2016). Recent insights into the mechanisms of activation of TCF4 show that TCF4-dependent transcription in primary neurons is induced by neuronal activity via soluble adenylyl cyclase and protein kinase A (PKA) signalling (Sepp et al., 2017). In addition to the nervous system, TCF4 has been shown to function in the immune system during the development of plasmacytoid dendritic cells (Cisse et al., 2008; Grajkowska et al., 2017).

Deficits in TCF4 function are associated with several human diseases. TCF4 haploinsufficiency causes Pitt–Hopkins syndrome (PTHS; OMIM #610954) (Amiel et al., 2007; Brockschmidt et al., 2007; Zweier et al., 2007). As reviewed in an international consensus statement (Zollino et al., 2019), patients with PTHS have severe intellectual disability, developmental delay, intermittent hyperventilation periods followed by apnea, and display distinct craniofacial features. Currently, there is no treatment for PTHS, but dissecting the functional consequences triggered by mutated *TCF4* alleles could reveal attractive avenues for curative therapies for this disorder (reviewed in Rannals and Maher, 2017). Large-scale genome-wide association studies revealed single nucleotide polymorphisms in *TCF4* among the highest risk loci for schizophrenia (SCZ) (Talkowski et al., 2012). Consistently, *TCF4* is involved in SCZ endophenotypes, such as neurocognition and sensorimotor gating (Lennertz et al., 2011a,b; Quednow et al., 2011). Furthermore, many genes that are mutated in SCZ, autism spectrum disorder and intellectual disability patients are TCF4 target genes (Forrest et al., 2018). How deficits in TCF4 function translate into neurodevelopmental impairments, and whether TCF4 plays an essential role in the mature nervous system, is poorly understood.

We have previously demonstrated that TCF4 function can be modelled in *Drosophila melanogaster* using its orthologue and the sole E-protein in the fruit fly, Daughterless (Da) (Tamberg et al., 2015). PTHS-associated mutations introduced to Da lead to similar consequences in the fruit fly as do the same mutations in TCF4 *in vitro* ranging from hypomorphic to dominant negative effects (Sepp et al., 2012; Tamberg et al., 2015). Furthermore, human TCF4 is capable of rescuing the lack of Da in the development of the *Drosophila* embryonic nervous system (Tamberg et al., 2015).

Da is involved in various developmental processes including sex determination, neurogenesis, myogenesis, oogenesis, intestinal

[1]Department of Chemistry and Biotechnology, Tallinn University of Technology, Akadeemia tee 15, Tallinn 12618, Estonia. [2]Protobios LLC, Mäealuse 4, Tallinn 12618, Estonia.

*Authors for correspondence (laura.tamberg@taltech.ee; mari.palgi@taltech.ee)

A.S., 0000-0002-8132-5258; C.S.K., 0000-0002-0829-0482; M.S., 0000-0003-1733-8385; M.P., 0000-0001-6142-6142

Handling Editor: Ross Cagan
Received 3 October 2019; Accepted 24 June 2020

Disease Models & Mechanisms (2020) 13, dmm042747. doi:10.1242/dmm.042747

stem cell maintenance and the development of the eye, trachea and salivary gland (Bardin et al., 2010; Bhattacharya and Baker, 2011; Brown et al., 1996; Castanon et al., 2001; Caudy et al., 1988; Cline, 1978; Cummings and Cronmiller, 1994; King-Jones et al., 1999; Massari and Murre, 2000; Smith et al., 2002; Wong et al., 2008). In the developing nervous system, the role of Da is well established during neuronal specification as an obligatory heterodimerization partner for proneural class II bHLH transcription factors (Cabrera and Alonso, 1991; Powell et al., 2008). However, the functional role of Da following neurogenesis and nervous system maturation remains unknown.

Here, we set out to characterize the expression of Da in the nervous system. To this end, we created *Drosophila* lines in which Da protein was endogenously tagged with either 3xFLAG or sfGFP epitope tags. We showed that Da is broadly expressed in the larval central nervous system (CNS), including in populations of Kenyon cells contributing to the mushroom body, which is the memory and learning centre of insects. To test whether Da is involved in learning and memory formation in the fruit fly, we used the appetitive associative learning paradigm in larvae (Michels et al., 2017). In this assay, silencing of *da* by several CNS-specific Gal4 drivers resulted in impaired learning and memory formation. Knockdown of *da* using 30Y-Gal4 also impaired negative geotaxis of adult flies. These phenotypes were moderately improved by adding resveratrol or suberoylanilide hydroxamic acid (SAHA) to the food substrate. Therefore, our results show that knockdown of *da* combined with appetitive associative learning paradigm or negative geotaxis assay is further applicable for screening potential therapeutics for the treatment of PTHS, as well as putative genetic interactors of Da and by proxy, TCF4. Furthermore, silencing of *da* resulted in a decreased level of the synaptic proteins Synapsin (Syn) and Discs large 1 (Dlg1) in third instar larval brains. We also demonstrated that Da binds to several areas in the *dlg1* gene and to the *Syn* promoter region in adult *Drosophila* heads, and that overexpression of *da* increases *Syn* and *dlg1* mRNA levels in the adult heads. Collectively, we have shown for the first time that Da is required to sustain elements of the synaptic proteome in a mature nervous system positing a post-developmental function for Da and possibly TCF4.

RESULTS

Da is expressed at all developmental stages of the fruit fly

Although the expression of Da protein has been studied in fruit fly embryos, ovaries, larval optic lobes and imaginal discs using various anti-Da antibodies (Andrade-Zapata and Baonza, 2014; Bhattacharya and Baker, 2011, 2012; Brown et al., 1996; Cronmiller and Cummings, 1993; Li and Baker, 2018; Tanaka-Matakatsu et al., 2014; Yasugi et al., 2014), its expression during adulthood remains largely uncharacterized. Therefore, we first aimed to study Da expression throughout the development of the fruit fly using immunoblot analysis. As there are no commercial antibodies available that recognize Da, we used the CRISPR/Cas9 system to create transgenic flies in which Da was N-terminally tagged with 3xFLAG epitope. The resulting *3xFLAG-da* line was maintained in a homozygous state, indicating that the tagged Da protein is functional as both *da* null mutations and *da* ubiquitous overexpression lead to embryonic lethality (Caudy et al., 1988; Giebel et al., 1997). We then characterized Da expression throughout development, and in adult *Drosophila* heads of the *3xFLAG-da* line, by performing immunoblot analysis with anti-FLAG antibodies. During development, we compared Da expression from embryonic to late pupal stages (Fig. 1A,C). In adults, we analysed the Da levels from the heads of 1-, 4- and 7-day-old males and females (Fig. 1B,D). In addition to the expected FLAG-Da signal at ∼80 kDa, we identified a previously uncharacterized lower molecular weight (∼65 kDa) Da signal (Fig. 1A,B, Fig S1, Fig. S2A). This signal was also present in a western blot of embryos using anti-Da antibody in flies overexpressing Da under the ubiquitous strong driver da^{G32}*-Gal4* but not with nervous system-specific *R123B08-Gal4* (Fig. S1). We found no significant differences in 80 kDa Da protein expression throughout development from the embryonic stage to pupariation (Fig. 1C). During adulthood, Da expression was highest in the heads of 1-day-old females and decreased thereafter in both males and females (Fig. 1D). Expression of ∼65 kDa Da decreased during development, with highest levels at the embryonic stage (Fig. 1C,D). This ∼65 kDa Da protein also seemed to be mostly non-neural as its expression level was very low in larval brains (Fig. S2).

3xFLAG-Da retains the transactivational capability of Da in HEK293 cells

In addition to *3xFLAG-da*, we also created *sfGFP-da* flies, in which Da is tagged with superfolder green fluorescent protein (sfGFP) in the same N-terminal position. This line was also maintained in a homozygous state, indicating that sfGFP tag does not interfere with the function of Da *in vivo*. To determine whether N-terminal tagging of Da proteins influences their transactivation capability, we used an *in vitro* luciferase reporter system in which the expression of the *luciferase* gene was controlled by E-boxes with a minimal promoter. Therefore, we cloned the *3xFLAG*-tagged or *sfGFP*-tagged *da* from the genomes of the tagged lines into mammalian expression vector pcDNA3.1 and overexpressed these constructs in HEK293 cells. The luciferase reporter assay showed that compared to wild-type Da the transactivational capability of 3xFLAG-Da was unchanged (Fig. 2A). In contrast, the transactivational capability of sfGFP-Da was significantly reduced (Fig. 2A). To determine whether the effects seen in the luciferase reporter assay were due to differential expression levels of the Da proteins used, we also performed western blot analysis, which revealed that both *3xFLAG-da* and *sfGFP-da* constructs were expressed at equal levels (Fig. 2B). This suggests that the 3xFLAG tag, unlike the sfGFP tag, does not interfere with the expression and transcriptional activity of Da. Correspondingly, we focused on using 3xFLAG-Da flies in subsequent experiments.

Da is widely expressed in the third instar larval brain

Next, we used the *3xFLAG-da* line to characterize the expression of Da in the third instar larval brain. Da was expressed weakly and ubiquitously throughout the larval CNS, with stronger expression detected in some nuclei and cytoplasm of specific cells (Fig. 3A″,B″,C″). Mutations in, or deletion of, one of the *TCF4* alleles lead to PTHS in humans. One of the hallmarks of PTHS is severe learning disability, and it has been shown that *TCF4* is highly expressed in the adult human and rodent hippocampus, which is the brain structure involved in learning and memory (Sepp et al., 2011; Jung et al., 2018). Additionally, single-cell RNA sequencing data have shown that *da* mRNA is expressed widely in the adult fly brain and also in mushroom body Kenyon cells (Davie et al., 2018). Therefore, we attempted to determine whether TCF4 homologue Da is expressed in the mushroom body, the brain structure of insects responsible for learning and memory. To facilitate this, we deployed the UAS-Gal4 binary expression system (Brand and Perrimon, 1993) by combining the *3xFLAG-da* line with different driver lines with expression in the mushroom body. Resulting lines with *3xFLAG-da* and *Gal4* were then

Disease Models & Mechanisms (2020) 13, dmm042747. doi:10.1242/dmm.042747

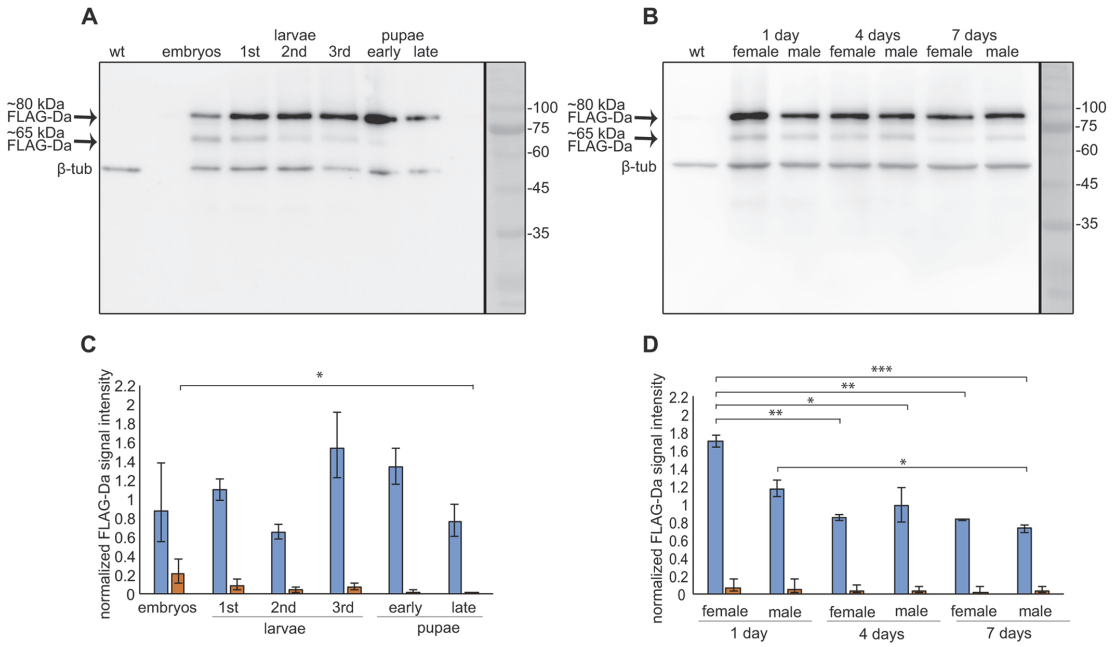


Fig. 1. Da is expressed in all developmental stages of the fruit fly. (A,B) The 3xFLAG-Da fusion protein is expressed throughout fruit fly development. Western blot analysis carried out using anti-FLAG antibody reveals two FLAG-Da-specific signals, one at ∼80 kDa and the other at ∼65 kDa. The w^{1118} wild type (wt) serves as a negative control. The numbers on the right side indicate the molecular weights of proteins in kDa. (C,D) Results of densitometric analysis of western blot. The 3xFLAG-Da signals were normalized using β-tubulin signals. The mean results from three independent western blots are shown. The blue bars represent the mean intensity of the 80 kDa protein signal and the orange bars represent the mean intensity of the 65 kDa protein signal. Data are mean±s.e.m. Statistical significance is shown with asterisks between the groups connected with lines. **P*<0.05, ***P*<0.01, ****P*<0.001, one way ANOVA with post-hoc Bonferroni test.

combined with nuclear-targeted *UAS-nls-GFP*. The *R12B08-Gal4* line directed *Gal4* expression under the control of the single intron of *da* in most regions of the brain, including the mushroom body (Fig. 3A,A‴). Two mushroom body-specific drivers, *201Y-Gal4* and *30Y-Gal4*, were used to express nuclear GFP to visualize Kenyon cells (Fig. 3B,B‴,C,C‴). We observed that the expression of 3xFLAG-Da and *R12B08>nls-GFP* overlapped in many areas of the third instar larval brain (Fig. 3A′). With *201Y-Gal4* and *30Y-Gal4*, the mushroom body-specific driver lines, 3xFLAG-Da showed partial co-expression in cells contributing to the third

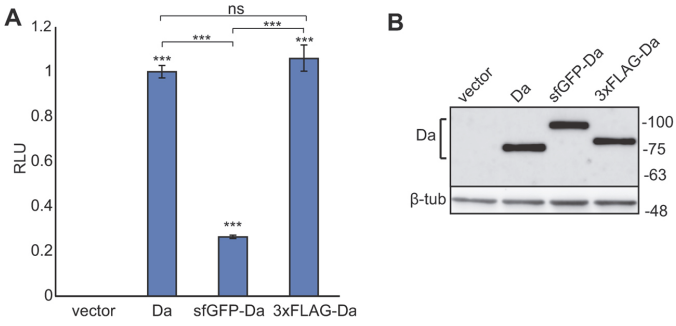


Fig. 2. Transactivational capability of Da is unaffected by an N-terminal 3xFLAG tag, but is reduced by an sfGFP tag. (A) HEK293 cells were co-transfected with constructs encoding wild-type *da*, tagged *da* or empty vector, a firefly luciferase construct carrying 12 μE5 boxes with a minimal promoter, and a *Renilla* luciferase construct without E-boxes for normalization. Luciferase activities were measured and data are presented as fold-induced levels compared to the signals obtained from cells transfected with the wild-type *da* encoding construct. The mean results from six independent transfection experiments performed in duplicates are shown. Data are mean±s.e.m. Statistical significance compared to cells transfected with empty vector is shown with asterisks above the bars. Statistical significance between the groups is indicated with brackets. ****P*<0.001; ns, not significant; one way ANOVA with post-hoc Bonferroni test. RLU, relative luciferase unit. (B) Western blot from transfected HEK293 cells using anti-Da antibody dam109-10. Wild-type Da, sfGFP- and 3xFLAG-tagged Da are all expressed at equal levels. Numbers on the right side indicate the molecular weight of the proteins in kDa.

Disease Models & Mechanisms (2020) 13, dmm042747. doi:10.1242/dmm.042747

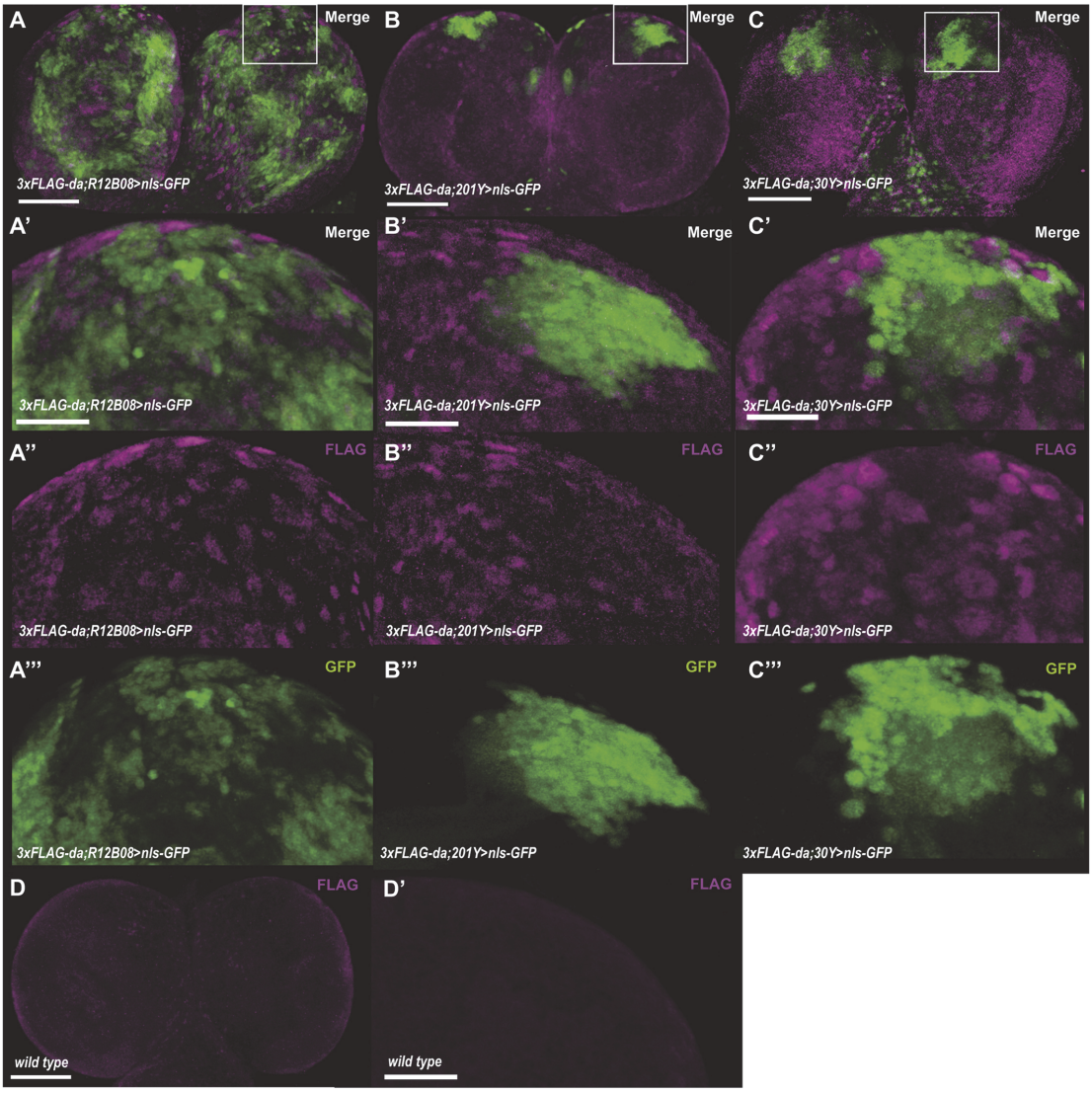


Fig. 3. Da is widely expressed in the third instar larval brain, including several Kenyon cells of the mushroom body. (A-A‴) *R12B08-Gal4* is expressed widely in the larval brain. (B-C‴) *da* is expressed in some of the mushroom body Kenyon cells in third instar larval brains marked by *201Y-Gal4* (B-B‴) and by *30Y-Gal4* (C-C‴). nls-GFP expression shows driver expression pattern (A″,B‴,C‴). The expression of 3xFLAG-Da (A″,B″, C″). Wild-type third instar larval brains showing unspecific binding of anti-FLAG antibodies (D,D′). Scale bars: 70 µm (A,B,C,D; boxes show the area represented in A′-C‴); 20 µm (A′,B′,C′,D′).

instar mushroom body (Fig. 3B′,C′). Thus, Da is expressed broadly in the CNS of third instar larvae, including a portion of the mushroom body. Wild-type larval brains were used to validate the specificity of anti-FLAG antibodies (Fig. 3D,D′).

Silencing of *da* in the CNS leads to impaired memory of the larvae

Heterozygous mutations in *TCF4*, the orthologue of *da*, lead to PTHS syndrome, which is characterized by intellectual disability. This fact, and the observation that Da is expressed in a portion of Kenyon cells contributing to the mushroom body, imply that Da might be involved in learning and memory in flies. To test this, we decided to take advantage of the ease of assaying appetitive associative learning and memory in the *Drosophila* larvae (Michels et al., 2017). In this assay, associative memory between odours and taste reward in the larvae was tested. Larvae were trained three times by being presented with one odour with a fructose reward and the other odour with no reward. During the test, larvae were given a

Disease Models & Mechanisms (2020) 13, dmm042747. doi:10.1242/dmm.042747

choice between the two odours. Then, reciprocal training with a new set of larvae was conducted with subsequent testing. After the second test, the performance index (PI) was calculated. However, this assay showed that learning ability was not impaired in *da* heterozygous mutants da^{10}/*CyO* (Fig. 4A), which could be due to *da* upregulation by autoregulation (Smith and Cronmiller, 2001). As sfGFP-Da showed diminished transactivation capability *in vitro* (see above), we also tested homozygous *sfGFP-da* larvae and found no impairment of learning (Fig. 4A). Thus, we next investigated whether knockdown of *da* with concurrent enhancement by Dicer-2 (Dcr2) expression (Dietzl et al., 2007) in the *Drosophila* CNS could impact memory and learning ability. To silence *da* in the CNS, we used several CNS-specific *Gal4* lines. We found that silencing of *da* using three drivers, *R12B08-Gal4* (Fig. 4B) and mushroom body-specific lines *30Y-Gal4* (Fig. 4C) and *201Y-Gal4* (Fig. 4D) (genotypes: *UAS-Dcr2;UAS-da*^{RNAi}*;R12B08-Gal4*, *UAS-Dcr2/+;UAS-da*^{RNAi}*/+;30Y-Gal4/+* and *UAS-Dcr2;201Y-Gal4;UAS-da*^{RNAi}) caused larvae to have zero PI, meaning their appetitive associative learning was impaired. For controls, we used both the *UAS-da*^{RNAi} line and the *UAS-Dcr2* driven by the CNS-specific *Gal4* line (genotypes: *UAS-Dcr2;+;R12B08-Gal4*, *UAS-Dcr2/+;+;30Y-Gal4/+* or *UAS-Dcr2;201Y-Gal4;+*). All of the control larvae had a non-zero PI with regards to memory. In the case of the *UAS-Dcr2;3xFLAG-da, UAS-da*^{RNAi}*;R12B08-Gal4* line (in which *3xFLAG-da*, *UAS-da*^{RNAi} and *R12B08-Gal4* were all in a homozygous state), Da levels in the larval brains were reduced by ∼25% and ∼35% when compared to *UAS-Dcr2;3xFLAG-da;R12B08-Gal4* and *3xFLAG-da* larval brains, respectively (Fig. S2). To validate that the observed learning phenotype was caused by *da* silencing and not by off-target effects, we conducted rescue experiments using simultaneous *R12B08-Gal4*-driven *da* silencing and overexpression. *UAS-Dcr2;UAS-da*^{RNAi}*;R12B08-Gal4,UAS-da/+* larvae had a non-zero PI, whereas *UAS-Dcr2;UAS-da*^{RNAi}*;R12B08-Gal4/+* had a zero PI (Fig. S3), indicating that overexpressing da partially rescued the memory deficit. Larvae were also tested for their ability to taste and smell. Silencing of *da* using *R12B08-Gal4*, *30Y-Gal4* or *201Y-Gal4* did not impair fructose (Fig. S4A), amyl-acetate (AM) (Fig. S4B) or octanol (OCT) preference (Fig. S4C). Interestingly, *UAS-Dcr2;UAS-da*^{RNAi}*; R12B08-Gal4* larvae had a higher preference for odours; however, this preference was cancelled out due to reciprocal training. In addition, in the memory test situation there were two odours present on the Petri dish, but in the smell sensing test there was only one, which could explain why larvae tended to move towards it. All tested larvae moved around the agar plate, which indicated that they had normal locomotion. Our findings suggest that for normal larval appetitive associative memory, appropriate Da levels are needed in the brain structures specified by *R12B08-Gal4*, *30Y-Gal4* and *201Y-Gal4*.

Reduced levels of Da in the larval CNS lead to decreased expression of synaptic proteins Syn and Dlg1

To investigate the putative mechanisms underlying learning and memory deficits in larvae with lowered levels of Da in the nervous system, we used the driver line *R12B08-Gal4* for silencing *da*, as it had the broadest expression. We compared the expression levels of several known synaptic proteins in the third instar larval brains under both *da* knockdown and overexpression conditions using the *R12B08-Gal4* line (Fig. 5). We quantified the expression levels of the presynaptic protein bruchpilot (Brp) (Fig. 5A), postsynaptic protein Dlg1 (Fig. 5B), presynaptic Syn (Fig. 5C), which is important for learning and memory (Michels et al., 2005), and pan-neuronally expressed neuronal-specific splicing factor embryonic lethal abnormal vision (Elav) (Fig. 5D). We found that the levels of both Dlg1 and Syn were reduced in third instar larval brains with

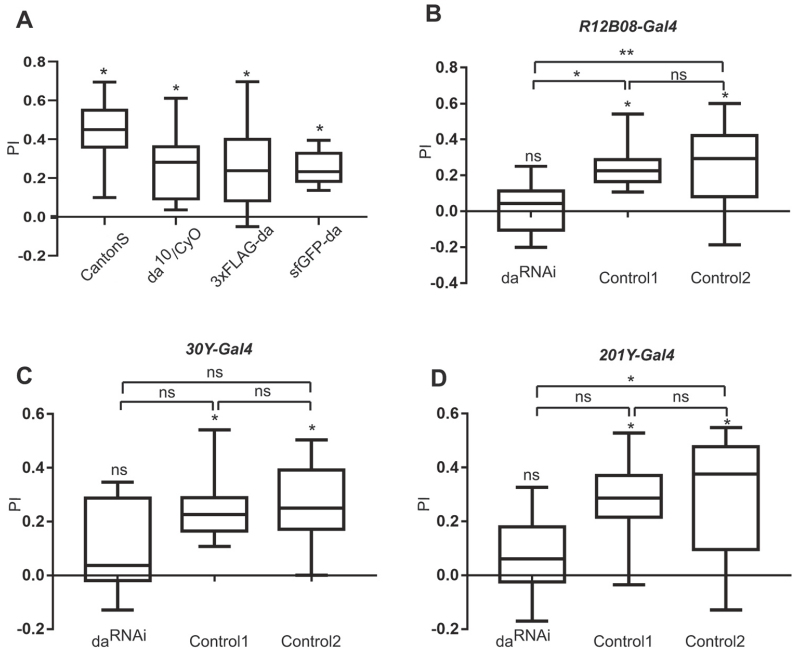


Fig. 4. Knockdown of *da* in the mushroom body leads to impaired olfactory learning of larvae. (A) CantonS, da^{10}/*CyO*, *3xFLAG-da* and *sfGFP-da* larvae have non-zero PI showing appetitive associative memory. (B-D) Larvae have zero PI meaning appetitive associative learning was impaired when *da* was silenced using *R12B08-Gal4* (*UAS-Dcr2; UAS-da*^{RNAi} *KK105258;R12B08-Gal4* homozygotes) (B), *30Y-Gal4* (*UAS-Dcr2/+; UAS-da*^{RNAi} *KK105258/+;30Y-Gal4/+* heterozygotes) (C) and 201Y-Gal4 (*UAS-Dcr2;201Y-Gal4;UAS-da*^{RNAi} *GD51297* homozygotes) (D) indicated by ns over the boxes of da^{RNAi}. For Control1 in B, C and D, *UAS-da*^{RNAi} larvae were used without drivers. For Control2, *UAS-Dcr2;+;R12B08* homozygotes were used in B, *UAS-Dcr2/+;+; 30Y-Gal4/+* heterozygotes in C and *UAS-Dcr2;201Y-Gal4;+* homozygotes in D. PIs are visualized using box-whisker plots that show the median, the 25% and 75% quantiles (boxes), and the minimum to maximum (whiskers). For statistical analysis to determine PI difference compared to zero inside one genotype, a one-sample sign test (asterisks indicated over the boxes) was used. Between the groups, a Kruskal–Wallis ANOVA with Dunn's post-hoc test (indicated on the brackets) was used. *P<0.05, **P<0.01; ns, not significant.

Disease Models & Mechanisms (2020) 13, dmm042747. doi:10.1242/dmm.042747

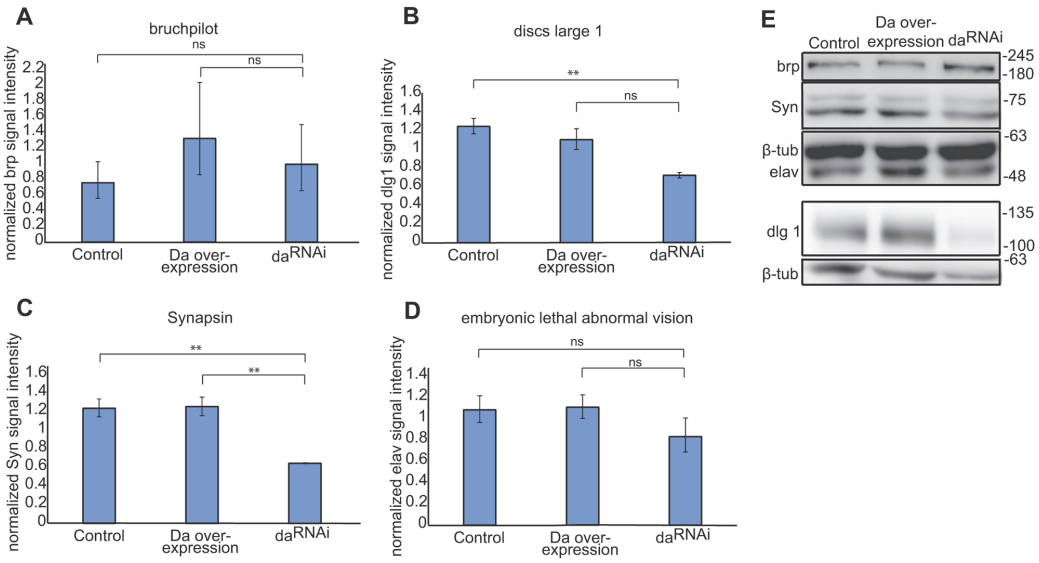


Fig. 5. Silencing of *da* lowers the expression levels of Syn and Dlg1. Western blot was carried out using larval brains in which *da* was silenced with *R12B08-Gal4*. (A-D) Results of densitometric analysis of the western blot. Protein signals were normalized using β-tubulin signals. The mean results from four independent western blots are shown. Data are mean±s.e.m. Statistical significance is shown with asterisks between the groups connected with brackets. ***P*<0.01; ns, not significant. Paired, two-tailed Student's *t*-test. Overexpression of Da does not alter bruchpilot, discs large 1, Syn or Elav levels (A-D). Dlg1 and Syn expression levels were lower when Da was silenced (B,C). Silencing of Da did not change bruchpilot and Elav expression levels (A,D). (E) Representative western blot using third instar larval brains. Numbers indicate the molecular weight of proteins in kDa. Control, *R12B08>Dcr2* larval brains; Da overexpression, *R12B08>Da* larval brains; da^{RNAi}, *R12B08>Dcr2,da^{RNAi}* larval brains.

lower levels of Da (Fig. 5B,C,E). On the other hand, Da overexpression did not result in increased levels of these proteins. To further confirm that lowered Da levels decrease Syn and Dlg1 expression we used immunohistochemistry in third instar larval brains. For the silencing of *da*, we used the *201Y*-Gal4 line as it was the strongest and most specific driver in the mushroom body of the lines used. We detected weaker Syn and Dlg1 levels compared to controls (Fig. S5). The levels of Elav and Brp were not significantly changed by knockdown or overexpression of *da* (Fig. 5A,D). The finding that Elav levels were not affected by Da suggests that reducing Da levels does not affect the number of neurons, and the observed learning impairment might instead stem from lowered expression levels of synaptic proteins or, alternatively, from reduced numbers of synapses.

Syn and *dlg1* are Da target genes

As Syn and Dlg1 protein levels were reduced in third instar larval brains when *da* was silenced, and it has been shown that Da binds to both *Syn* and *dlg1* gene loci at embryonic stages 4 and 5 (MacArthur et al., 2009), we sought to investigate whether Da binds to these areas in adult heads too. To facilitate this, we conducted a chromatin immunoprecipitation (ChIP) assay in *3xFLAG-da* adult heads using anti-FLAG antibodies. As a control we used the *white^{1118}* fly line with no *FLAG* tag. For quantitative PCR (qPCR) from immunoprecipitated chromatin, we designed primers to amplify *Syn* and *dlg1* gene areas containing E-boxes to which Da binds in early embryos (MacArthur et al., 2009). In addition to the previously shown Da-binding site in the *Syn* gene, we also tested Da binding to the *Syn* promoter region (Fig. 6A). For *dlg1*, we designed four primer pairs, as Da has been shown to bind four areas in that gene (Fig. 6B) (MacArthur et al., 2009). As a negative control, we used primers for *achaete* (Andrade-Zapata and Baonza, 2014), as it encodes a proneural protein essential for neuronal development and should not be expressed in adult heads. As a positive control, we used the *peptidylglycine-α-hydroxylating monooxygenase* (*Phm*) gene first intron in which Da binds as a heterodimer with dimmed to activate transcription (Park et al., 2008). qPCR with immunoprecipitated chromatin using *Syn* primers resulted in the enrichment of the *Syn* promoter area (primer pair SynI), whereas the previously reported Da-binding site was not enriched in adult heads (primer pair SynII) (Fig. 6C). All *dlg1* primers resulted in the enrichment of previously reported Da-binding areas (Fig. 6C). This means that Da does not bind to the locus at the 3′ end of the *Syn* gene but binds to the *Syn* promoter and all four *dlg1* gene areas that we selected in the adult heads. To validate *Syn* and *dlg1* as Da target genes, we carried out RT-qPCR analysis in adult *Drosophila* heads under *da* silencing and overexpression conditions. Here, we used pan-neuronal *elav-Gal4* to silence *da* in all neurons. Although upon *da* silencing using *elav-Gal4*, only *Syn* mRNA levels were decreased (Fig. 6D), *da* overexpression using *elav-Gal4* increased mRNA levels of *Syn* and *dlg1* (Fig. 6E). This indicates that both *Syn* and *dlg1* are direct targets of Da in the *Drosophila* nervous system.

Suppressing Da using *30Y-Gal4* leads to impaired negative geotaxis of adult flies

Negative geotaxis has been successfully used to evaluate climbing ability indicative of motor dysfunction in the *Drosophila* model for Angelman syndrome, which has similar symptoms to PTHS (Wu et al., 2008). Thus, we next used this assay to evaluate locomotion in

Disease Models & Mechanisms (2020) 13, dmm042747. doi:10.1242/dmm.042747

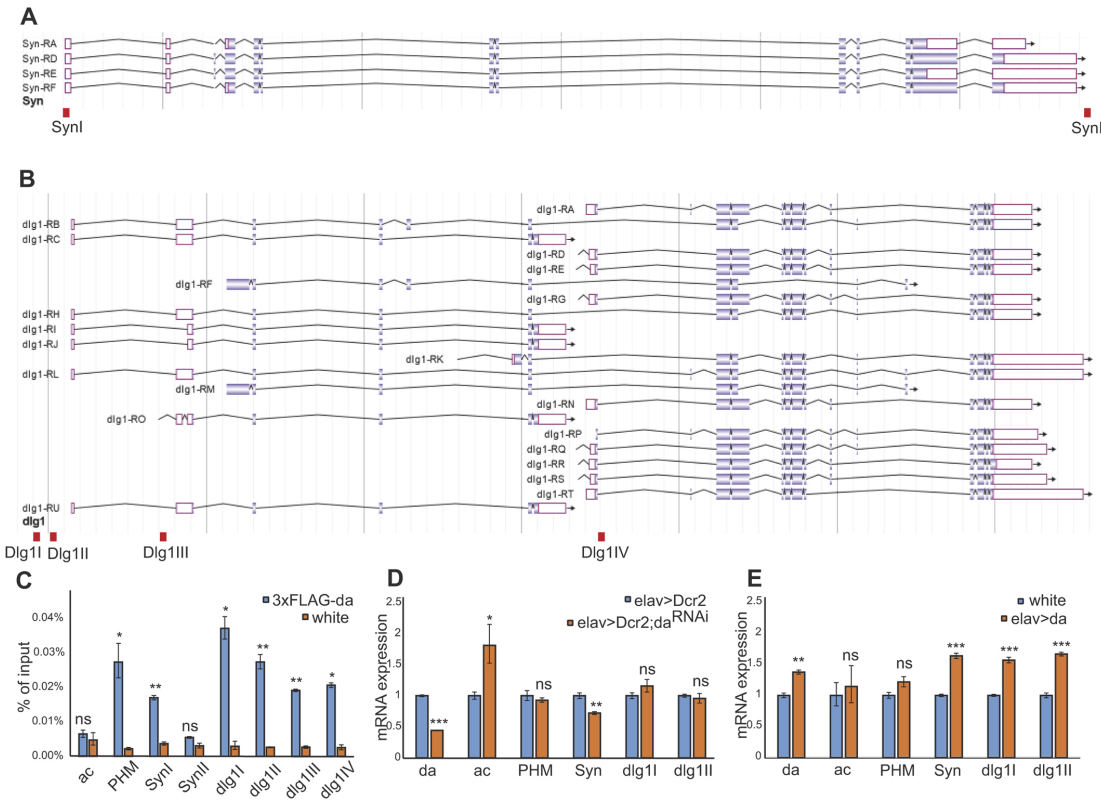


Fig. 6. Da directly regulates *Syn* and *dlg1* in adult fly heads. (A) JBrowse view of the *Syn* gene. Four annotated transcripts are shown. (B) JBrowse view of the *dlg1* gene. A total of 21 annotated transcripts are shown. Red boxes indicate areas in which primer pairs Syn1, SynII, Dlg1I, Dlg1II, Dlg1III and Dlg1IV amplify DNA. (C) qPCR results from the ChIP experiment with 3xFLAG-*da* and *white*^{1118} wild-type adult heads using the anti-FLAG antibody. ac, *achaete* gene locus used as a negative control; PHM, *peptidylglycine alpha-hydroxylating monooxygenase* gene first intron used as a positive control; SynI, promoter region of *Syn*; SynII, 3′ end of *Syn* gene locus; dlg1I, dlg1II, dlg1III and dlg1IV, *discs large 1* gene locus. (D,E) RT-qPCR results showing the effects of *da* silencing (D) or overexpression (E) using *elav-Gal4* on *da*, *ac*, *PHM*, *Syn* and *dlg1* mRNA levels. *da* silencing reduces *da* and *Syn*, and increases *ac* mRNA levels (D). *da* overexpression increases *da*, *Syn* and *dlg1* mRNA levels (E). (C-E) Results from three biological replicates are shown. Data are mean±s.e.m. *P<0.05, **P<0.01, ***P<0.001, Paired, two-tailed Student's *t*-test.

adult flies in which *da* knockdown had been achieved by the same drivers as used for the larval learning test. We found that negative geotaxis was unchanged in homozygotes in which *da* knockdown had been achieved by the larval broad neuronal driver *R12B08-Gal4* or mushroom body-specific driver *201Y-Gal4* (Fig. 7A,C,E,G). Interestingly, both female and male heterozygotes in which *da* was silenced by the *30Y-Gal4* driver had severely impaired negative geotaxis (Fig. 7B,F). Rescue experiments were performed to validate that impaired negative geotaxis was caused by *da* silencing and not by off-target effects. The negative geotaxis phenotype was rescued using simultaneous *30Y-Gal4*-driven *da* silencing and overexpression (Fig. 7J,M). Expression of human *TCF4B* under *da*-silencing conditions had a tendency to improve geotaxis (Fig. 7K,N). To further eliminate the possibility of off-target effects, we used the alternative *da*^{RNAi} fly line, *UAS-da*^{RNAi}*(GD51297)*, which also caused impairment of negative geotaxis (Fig. 7I,L). An alternative mushroom body driver, *OK107-Gal4*, was used to investigate whether the impairment of geotaxis was caused by lowered Da levels in the mushroom body, but the results revealed no change in negative geotaxis compared to controls (Fig. 7D,H). Next, we visualized Da expression in the adult brains using the *3xFLAG-da* line. Da was expressed widely in the adult *Drosophila* brain including the central brain and thoracic ganglion (Fig. S6), and co-expressed with *30Y-Gal4* in many Kenyon cells in the mushroom body (Fig. S6A-A″). Fewer *OK107-Gal4*$^{+}$ cells were also Da^{+} (Fig. S6B-B″). Cells that cause negative geotaxis impairment when *da* is silenced must be marked by *30Y-Gal4* and not by *OK107-Gal4*, as silencing *da* by *30Y-Gal4* but not by *OK107-Gal4* caused negative geotaxis impairment. *30Y-Gal4* has broader expression outside the mushroom body; for example, in the thoracic ganglion. Silencing *da* using *R12B08-Gal4* or *201Y-Gal4* did not result in impaired negative geotaxis, possibly because co-expression of the drivers and Da is limited in the adult *Drosophila* brain (Fig. S6C-D″).

Larval appetitive associative learning and negative geotaxis assays can be used for screening drugs for PTHS treatment

Our finding showing that larval appetitive associative learning and adult negative geotaxis become impaired upon *da* silencing, indicates that these fly lines could be used for modelling certain

Disease Models & Mechanisms (2020) 13, dmm042747. doi:10.1242/dmm.042747

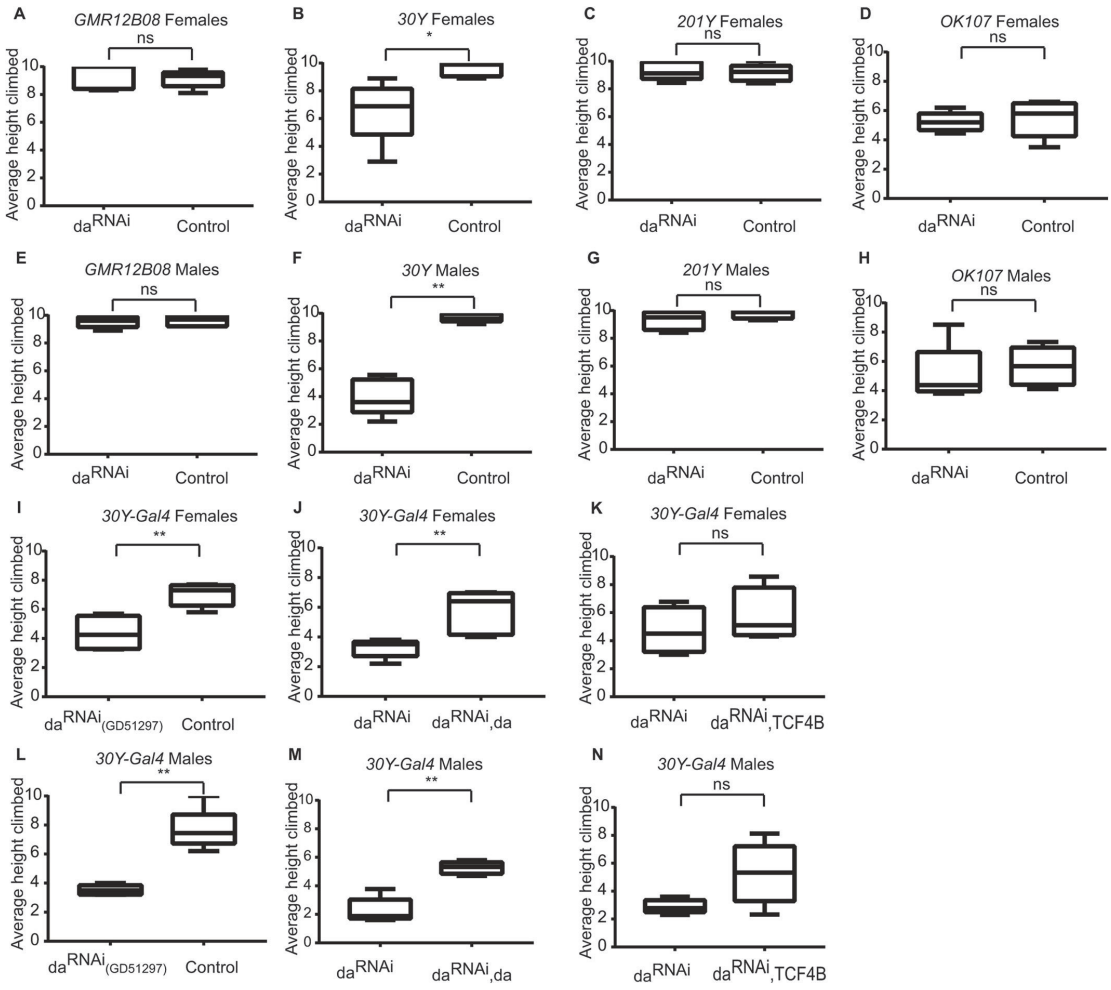


Fig. 7. Silencing of *da* with *30Y-Gal4* impairs negative geotaxis in adult flies. (A-N) Negative geotaxis was not affected when Da was suppressed using *R12B08-Gal4* (A,E), *201Y-Gal4* (C,G) or *OK107-Gal4* (D,H). The climbing height of the flies was significantly lower when Da was silenced using *30Y-Gal4* (B,F,I,L). (A-N) da^{RNAi}, da silencing using *UAS-da^{RNAi} KK105258* (A,B,D-F,H,J,K,M,N) or *UAS-da^{RNAi} GD51297* (C,G,I,L). Control, expressing only Dcr2. (J,M) Overexpression of *da*-rescued geotaxis. (K,N) Expression of human *TCF4B* had a moderate positive effect on geotaxis. Average climbing heights are visualized using box-whisker plots that show the median, the 25% to 75% quantiles (boxes) and the minimum to maximum (whiskers). For statistical significance, Mann–Whitney U tests were used. *P<0.05, **P<0.01; ns, not significant.

aspects of PTHS in *Drosophila* and testing potential drug candidates. For example, various drugs or drug candidates could be tested for their capacity to rescue these behavioural impairments. As TCF4-dependent transcription is activated by cAMP-PKA pathway-mediated phosphorylation of TCF4 in mammals (Sepp et al., 2017), and resveratrol inhibits cAMP-degrading phosphodiesterases, which leads to elevated cAMP levels (Park et al., 2012), we sought to investigate whether resveratrol increases the transactivational capability of Da and TCF4. We also tested the histone deacetylase inhibitor suberoylanilide hydroxamic acid (SAHA), as it has been shown to rescue memory impairment in the mouse model of PTHS (Kennedy et al., 2016). Therefore, we used the luciferase reporter system in cultured rat primary cortical neurons. Treating neurons with resveratrol for 8 h and resveratrol or SAHA for 24 h significantly increased the E-box-dependent transactivational capability of Da, 3xFLAG-da and two human TCF4 isoforms, a shorter isoform TCF4A and a longer isoform TCF4B (Sepp et al., 2011) (Fig. 8A,B). To validate that the increase in luciferase signals seen after treatments with resveratrol and SAHA was caused by an increase in the transcriptional activity of Da, TCF4A and TCF4B, and not by other effects, we performed luciferase reporter assays with Da, TCF4A and TCF4B mutants bearing mutations in the bHLH domain (Fig. S7). Previously, it has been shown that mutations in the bHLH domain abolish transcriptional activity of Da, TCF4A and TCF4B (Forrest et al., 2012; Sepp et al., 2012; Tamberg et al., 2015). Our results

Disease Models & Mechanisms (2020) 13, dmm042747. doi:10.1242/dmm.042747

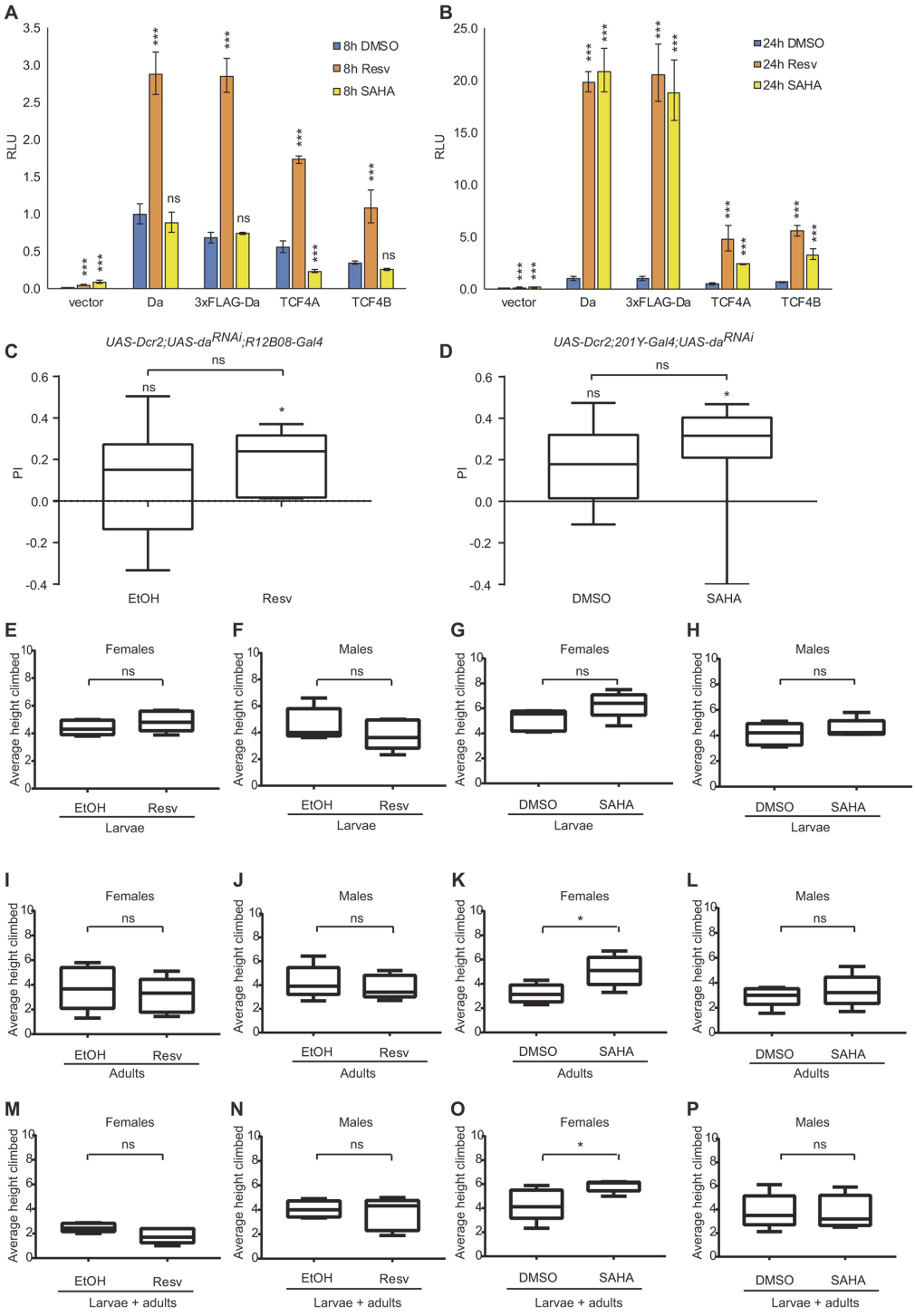


Fig. 8. See next page for legend.

Fig. 8. Resveratrol and SAHA have moderate positive effects on rescuing the impaired learning and negative geotaxis phenotype resulting from decreased levels of Da. (A,B) Cultured rat primary neurons were co-transfected with constructs encoding wild-type *da*, 3x-*FLAG-tagged da*, empty vector, *TCF4A* or *TCF4B*, a firefly luciferase construct carrying 12 µE5 boxes with a *TK* promoter, and a *Renilla* luciferase construct with mouse *Pgk* promoter for normalization, and treated with 50 µM resveratrol or 5 µM SAHA. Luciferase activities were measured and data are presented relative to the luciferase signals obtained from cells transfected with the wild-type *da*-encoding construct and treated with 0.1% DMSO for control. Mean results from three independent transfection experiments performed in duplicates are shown. Data are mean±s.e.m. Statistical significance is shown compared to 0.1% DMSO-treated cells expressing the respective effector protein. ***P<0.001; ns, not significant; one way ANOVA with post-hoc Bonferroni test; RLU, relative luciferase unit. (C,D) In the larval appetitive associative learning paradigm, the addition of 400 µM resveratrol or 2 µM SAHA to the growth media improved memory. (E-P) In the negative geotaxis assay 400 µM resveratrol or 2 µM SAHA was added to the food substrate of *UAS-Dcr2/+;UAS-da*^{RNAi}*/+;30Y-Gal4/+* only during larval development (E-H), during 5 days after eclosion (I-L) or both during larval development and 5 days after eclosion (M-O). Impaired negative geotaxis was improved when SAHA was fed to females only during adulthood (K) or both during larval development and after eclosion (O). In C-P, results are shown as box-whisker plots that show the median, the 25 to 75% quantiles as boxes and the minimum to maximum as whiskers. For statistical analysis, a one-sample sign test was used, as indicated above the boxes, and a Mann–Whitney U test was used indicated above the brackets. ns above the boxes in C and D indicates zero PI, * indicates non-zero PI, ns above the brackets indicates not significant and *P<0.05 (above the brackets).

confirmed that mutations in the bHLH domain cause loss of transcriptional activity of Da, TCF4A and TCF4B. Next, we decided to test these two substances in the appetitive associative learning and negative geotaxis experiments. We observed that although the *da* knockdown larvae fed with 400 µM resveratrol or 2 µM SAHA showed increased associative memory as their median PI was significantly different from zero, i.e. non-zero PI, the rescue of the learning deficit was not significantly different compared to the controls (Fig. 8C,D). We also tested negative geotaxis of *da* knockdown by *30Y-Gal4* flies fed with 400 µM resveratrol or 2 µM SAHA during larval development (Fig. 8E-H) and 5 days after eclosion (Fig. 8I-L), and both treatments during larval development and 5 days from the beginning of adulthood (Fig. 8M-P). SAHA improved the impairment of negative geotaxis of female flies when they were fed after eclosion (Fig. 8K) or during larval development and after eclosion (Fig. 8O). In addition, our results showed that resveratrol and SAHA did not increase TCF4 and Da protein levels (Fig. S8A-D), confirming that the activation of E-box-controlled reporter genes was caused by an increase in transcriptional activity of TCF4 and Da, and not by an increase of their protein levels. Next, we investigated whether feeding resveratrol or SAHA to larvae causes an increase in Da targets Syn and Dlg1, but we were unable to detect any increase in the expression of these proteins (Fig. S8E-G). These results indicate that the improvement of the learning and geotaxis phenotypes by resveratrol and SAHA are possibly caused by other mechanisms rather than rescuing the levels of synaptic proteins Syn and Dlg1. Nevertheless, rescuing negative geotaxis impairment caused by lowered levels of Da provides a powerful tool for finding drugs that can potentially improve PTHS symptoms.

DISCUSSION

Here, we characterized the expression of Da protein in the *Drosophila* larval and adult brain. Da was expressed in many areas of the brain, including populations of Kenyon cells in the mushroom body, which is the centre for learning and memory in the fruit fly and carries out a role that is comparable to the mammalian hippocampus. Single-cell RNA sequencing has shown that *da* is expressed widely in the adult brain and that *da* co-expresses with *eyeless* and *portabella*, which are markers for mushroom body Kenyon cells (Davie et al., 2018). The orthologue of *da*, *TCF4*, is expressed not only in the adult mammalian hippocampus but also in cortical and subcortical structures (Jung et al., 2018).

We created N-terminally tagged *3xFLAG-da* and *sfGFP-da* fly strains. Both strains are homozygous viable and fertile, indicating that the overall functionality of Da *in vivo* is not altered by the molecular tag. However, in a luciferase reporter assay in mammalian HEK293 cells, the sfGFP tag reduced transcription activation capability of Da. E-proteins activate transcription preferably as heterodimers with class II bHLH proteins, but can also act as homodimers (Cabrera and Alonso, 1991). In mammalian HEK293 cells, Da probably activates transcription as a homodimer, as Da levels are high due to overexpression, so homodimer formation is preferred (Sepp et al., 2012). This suggests that the sfGFP tag could interfere with Da function as a homodimer in the luciferase assay but not as a heterodimer *in vivo*. We also compared the appetitive associative learning ability of *3xFLAG-da* and *sfGFP-da* larvae, and both of the lines had no learning impairment in this assay. This provides additional evidence that the 3xFLAG tag does not affect Da function and the sfGFP tag reduces its transactivational capability, probably by interfering with Da homodimer function.

As Da is widely expressed in the third instar larval CNS with some expression in structures associated with learning and memory, and PTHS is caused by heterozygous mutations in *TCF4*, we tested the learning ability of *da* heterozygous mutant larvae. These larvae had no memory impairment, which could be because of *da* upregulation by autoregulation (Smith and Cronmiller, 2001). Learning and memory of the larvae were impaired when *da* was knocked down using a broad neuronal driver and two mushroom body-specific drivers, although in third instar larval brains Da co-expression with the mushroom body drivers was limited. The impaired learning phenotype could be explained by neurodevelopmental issues resulting from lowered levels of Da protein during development or, alternatively, by the contribution of cells outside the mushroom body. Da mammalian orthologue TCF4 is also associated with learning and memory, as when *TCF4* is downregulated in the mouse hippocampus, pathways associated with neuronal plasticity are dysregulated (Kennedy et al., 2016), and silencing of *TCF4* in human pluripotent stem cell-derived neurons results in downregulated signalling pathways that are important for learning and memory (Hennig et al., 2017). In *TCF4* conditional knockout mice, the neurons in the cortex and hippocampus have reduced numbers of dendritic spines, which also suggests that synaptic plasticity is altered (Crux et al., 2018). In multiple PTHS mouse models, spatial learning is defective probably as a result of hippocampal N-methyl-D-aspartate receptor (NMDA) hyperfunction (Thaxton et al., 2018). Furthermore, many genes that code for synaptic proteins and have been linked to autism, intellectual disability, or psychiatric diseases, are direct targets of TCF4 (Forrest et al., 2018; Hennig et al., 2017).

Here, we showed that when *da* was silenced using a driver with broad expression in the *Drosophila* larval brain, expression levels of synaptic proteins Dlg1 and Syn was downregulated. Dlg1 is a member of the membrane-associated guanylate kinase (MAGUK) protein family. Several vertebrate homologues of Dlg1 have been shown to be important for learning and memory. Discs large MAGUK scaffold protein 3 [DLG3; also called synapse-associated protein 102 (SAP102)] knockout mice have spatial learning deficit (Cuthbert et al., 2007), and in human *DLG3*, mutations that cause dysfunctional NMDA receptor signalling have been associated with

X-linked mental retardation (Tarpey et al., 2004; Zanni et al., 2010). We also found that Da is a direct regulator of *dlg1*, as in adult *Drosophila* heads, Da binds to multiple areas in the *dlg1* gene and *dlg1* expression is upregulated when *da* is overexpressed. The gene coding for discs large MAGUK scaffold protein 2 [DLG2; also called postsynaptic density protein 93 (PSD-93)], which is a homologue of *Drosophila dlg1*, is a direct target of TCF4 (Hennig et al., 2017), which indicates that Da and TCF4 share at least some common mechanisms in regulating learning and memory.

Synapsins are presynaptic phosphoproteins that regulate synaptic output (reviewed by Diegelmann et al., 2013). There are three genes that encode vertebrate Synapsins but only one *Syn* gene in *Drosophila* (Klagges et al., 1996). Using knockout experiments in mice, it has been shown that Synapsins are involved in learning and memory (Gitler et al., 2004; Silva et al., 1996), and *SYN1* has been implicated in human neurological diseases, such as learning difficulties and epilepsy (Garcia et al., 2004). Likewise, *Drosophila* syn^{97} mutant larvae have impaired appetitive associative learning (Michels et al., 2005). The fact that the memory of syn^{97} larvae can be rescued by expressing Syn in the mushroom bodies (Michels et al., 2011), is consistent with our findings that lower Da levels affect Syn expression levels and that appropriate Da levels are required for proper memory formation. Syn-dependent memory is likely formed by its phosphorylation by Protein kinase A (PKA) (Michels et al., 2011). When Syn is phosphorylated at its PKA/CamK I/IV (Protein kinase A/Ca^{2+}/calmodulin-dependent protein kinase I/IV) sites, its affinity for actin is reduced and synaptic vesicles from the reserve pool can be exocytosed (reviewed by Benfenati, 2011). We found that Syn is likely a direct target of Da as Da binds to the *Syn* promoter, and both silencing and overexpression of *da* changes *Syn* mRNA levels.

We also sought to rescue the learning phenotype caused by *da* silencing. To facilitate this, we fed the larvae with resveratrol or SAHA, as our luciferase reporter experiments in primary neuronal cultures showed that resveratrol and SAHA significantly improve the transactivational capability of both Da and TCF4. Resveratrol inhibits cAMP-degrading phosphodiesterases, which leads to elevated cAMP levels (Park et al., 2012), and TCF4-dependent transcription upon neuronal activity is activated by cAMP-PKA pathway-mediated phosphorylation of TCF4 (Sepp et al., 2017). It is plausible that Da could also be regulated by phosphorylation by PKA; therefore, resveratrol improves Da transactivational capability. Also, resveratrol had a moderate positive effect on learning and memory in the *da* knockdown larvae. Whether this effect is linked to the cAMP-PKA pathway has yet to be verified. SAHA is a histone deacetylase inhibitor that improves learning and memory in *TCF4*(+/−) mice through the normalization of synaptic plasticity (Kennedy et al., 2016). Here, we showed that feeding SAHA to *Drosophila* larvae also had a moderate effect on learning.

Silencing of *da* by *30Y-Gal4* impaired negative geotaxis. We also sought to rescue the impaired geotaxis of *30Y>Dcr2;*da^{RNAi} flies using resveratrol or SAHA. We administered the drugs in the food substrate either during development or to adult flies or at both developmental stages. Negative geotaxis of female flies was significantly improved when SAHA was administered only after eclosion, or during both larval development and after eclosion. Supplementing the food of larvae only had no effect on negative geotaxis of the adults. The finding that SAHA only improved the phenotype of females could be due to the amount ingested by the males not being enough to rescue the geotaxis phenotype caused by lowered levels of Da. Recently, it has been shown that female flies do indeed consume more food than male flies (Wu et al., 2020), which could be the reason for SAHA improving geotaxis impairment of only female flies. In a recent study in *Drosophila*, in which genes associated with autism spectrum disorders and intellectual disability were suppressed, the knockdown of Da resulted in impaired habituation (Fenckova et al., 2019). The rescue of this habituation phenotype could also be tested to examine whether it can be improved with drugs.

Our study demonstrates that the levels of the TCF4 homologue Da are important for memory and negative geotaxis, possibly via regulation of the synaptic proteome. These novel learning and geotaxis deficiency models can be further used for screening therapeutics for TCF4-related diseases. Recently, using deconvolution analysis, TCF4 was identified as a master regulator in SCZ (Doostparast Torshizi et al., 2019). This opens up new avenues for using *Drosophila* to model TCF4-related diseases.

MATERIALS AND METHODS

Drosophila stocks

All *Drosophila* stocks and crosses were fed with malt and semolina-based food with 12 h light and dark daily rhythms at 25°C with 60% humidity, unless mentioned otherwise. *Drosophila* strains used in this study were *UAS-*da^{RNAi} GD51297 and *UAS-*da^{RNAi} KK105258 from the Vienna *Drosophila* Resource Center, CantonS (a gift from Dr Bertram Gerber, Leibniz Institute for Neurobiology, Magdeburg, Germany), da^{G32}*-Gal4* (a gift from Riitta Lindström, University of Helsinki, Helsinki, Finland), *UAS-TCF4B* (Tamberg et al., 2015), *201Y-Gal4* [Bloomington *Drosophila* Stock Center (BDSC, 4440)], *30Y-Gal4* (BDSC, 30818) (Yao Yang et al., 1995) and *OK107-Gal4* (BDSC, 854) were gifts from Mark Fortini, Thomas Jefferson University, Philadelphia, PA, USA. *R12B08-Gal4* (BDSC, 48489) (Pfeiffer et al., 2008, flweb.janelia.org/cgi-bin/view_flew_imagery.cgi?line=R12B08), *elav-Gal4* (BDSC, 8760) (Luo et al., 1994), *UAS-Dcr2;*Pin^{1}*/CyO* (BDSC, 24644) (Dietzl et al., 2007), *UAS-nls-GFP* (BDSC, 4776), *UAS-*da^{G} (BDSC, 37291), *UAS-mCD8-GFP;* Pin^{1}*/CyO* (BDSC, 5136) and *da* mutant line da^{10} (BDSC, 5531) (Caudy et al., 1988) were obtained from the BDSC. The following transgenic lines were generated in this study: *3xFLAG-*da^{2M4} and *sfGFP-*da^{4MI}.

Endogeneous tagging of Da by CRISPR/Cas9

The coding sequence for *3xFLAG*- or *sfGFP*-tag was inserted into the 5′ coding region of the *da* gene using CRISPR/Cas9 technology. The genomic sequence around the tag was as follows: 5′-ATGGCGA**CCA**GTG|ACGATGAGCC-3′ (PAM sequence shown as bold and the cut site marked with |). For the higher mutagenesis rate, a specific fruit fly line for guide RNA production was created. Partially overlapping oligonucleotides, 5′-CTTCGTGCATCGGCTCATCGTCAC-3′ and 5′-AAACTGGACGATGAGCCGATGCAC-3′, designed to target the N-terminus of the Da protein, were cloned downstream of the polymerase III U6:2 promoter in the pCFD2-dU6:2gRNA plasmid (Addgene #49409). Transgenic flies expressing gRNAs were created by injecting the generated plasmid into *PBac{yellow*$^{+}$*-attP-9A}VK00027* (BDSC, 9744) fly strain embryos. For donor plasmid generation, *pHD-3xFLAG-ScarlessDsRed* or *pHD-sfGFP-ScarlessDsRed* [both were gifts from Kate O'Connor-Giles, *Drosophila* Genomics Resource Center (DGRC), Indiana University, IN, USA] were used with Gibson cloning. The following primer pairs were used for the amplification of upstream and downstream homology arms:

upst F5, 5′-CGGCCGCGAATTCGCCCTTGGTTGTGAATCAGGTGTAGAAACA-3′ and

upst_R, 5′-GCCGGAACCTCCAGATCCACCACTGGTCGCCATTTCAGCA-3′; and

dwns_F, 5′-TTCTGGTGGTTCAGGAGGTTACGATGAGCCGATGCACTTG-3′ and

dwns_R, 5′-GTTTAAACGAATTCGCCCTTAACGCCCTGGAACACCGAGG-3′.

After verification, the obtained donor plasmids *pHD-da-3xFLAG-ScarlessDsRed* and *pHD-da-sfGFP-ScarlessDsRed* were injected into F_1

embryos from a cross between *da-gRNA* (our gRNA-expressing transgenic strain) and $y^1M\{w^{+mC}=nos\text{-}Cas9.P\}ZH\text{-}2A\ w^*$ (BDSC, 54591) fly strains. All embryo injections were ordered from BestGene.

The *dsRed* cassette was removed from selected progeny by crossing to the PiggyBac transposase line *Herm{3xP3-ECFP,atub-piggyBacK10}M10* (BDSC, 32073) (Horn et al., 2003). The obtained *3xFLAG-da* and *sfGFP-da* lines were verified by sequencing.

RNA isolation and cloning

RNA from *3xFLAG-da* or *sfGFP-da Drosophila* embryos was isolated using an RNeasy Mini Kit (Qiagen) according to the manufacturer's protocol. cDNA was synthesized using 2 µg of RNA. Primer sequences for cloning were 5′-ACTAGTTGAAGTCGACTGGAC-3′ and 5′-CCAGGTCCTCCAATTCCACC-3′. PCR products containing either *3xFLAG-da* or *sfGFP-da* cDNA sequences were sequenced and cloned into the *pCDNA3.1* expression vector (Tamberg et al., 2015) using BcuI (SpeI, 10 U; Thermo Scientific) and BstII (Eco 91I, 10 U; Thermo Scientific) restriction enzymes. The pcDNA3.1 constructs encoding Da, and reporter vectors pGL4.29[luc2P/12µE5/Hygro], pGL4[hRlucP/min/Hygro], pGL4[hRlucP/PGK/Hygro] and pGL4.29[luc2P/12µE5-TK/Hygro] have been described previously (Sepp et al., 2011, 2012, 2017; Tamberg et al., 2015).

Cell culture, transfections and luciferase reporter assay

Human embryonic kidney cells HEK-293 were obtained from ATCC (LGC Standards GmbH, Wesel, Germany), routinely tested for contamination and were grown in minimal essential media (Capricorn Scientific) supplemented with 10% fetal bovine serum (PAA Laboratories), 100 U/ml penicillin and 0.1 mg/ml streptomycin (Gibco). For transfection, 0.375 µg of DNA and 0.75 µg of polyethylenimine (Sigma-Aldrich) were used for each well of a 48-well plate, or scaled up accordingly. For co-transfections, equal amounts of pGL4.29[luc2P/12µE5/Hygro], pGL4[hRlucP/min/Hygro] and effector constructs were used. Cells were lysed 24 h after transfection.

Rat cortical neuronal cultures from Sprague Dawley embryonic day (E)22.5 rat embryos were obtained and maintained as described previously (Sepp et al., 2017). All animal procedures were approved by the local ethics committee. Neuronal cultures were transfected at 6 days *in vitro* (DIV) in conditioned medium. For transfection, 120 ng of expression plasmid, 60 ng of pGL4.29[luc2P/12µE5-TK/Hygro], 20 ng of pGL4[hRlucP/PGK/Hygro] and 0.6 µl of Lipofectamine 2000 (Invitrogen) were used. Neurons were treated with resveratrol, SAHA or DMSO as a vehicle and lysed at 8 DIV.

Luciferase assays were performed as described previously (Sepp et al., 2011) using passive lysis buffer (Promega) and the Dual-Glo luciferase assay system (Promega). For data analysis, background signals from untransfected cells were subtracted and firefly luciferase signals were normalised to *Renilla* luciferase signals. The data were then log transformed and auto scaled, means and standard deviations were calculated and paired, two-tailed Student's *t*-tests were performed. The data were back-transformed for graphical representation.

Protein electrophoresis and western blotting

For SDS-PAGE, embryos, larvae, pupae, adult heads or larval brains were lysed in 2× SDS sample buffer. Equal amounts of protein were loaded to the gel. The following mouse monoclonal antibodies were obtained from the Developmental Studies Hybridoma Bank (DSHB; University of Iowa, Iowa City, IA, USA): β-tubulin E7 (dilution 1:3000; DSHB AB_2315513, developed by M. Klymkowsky); Synapsin SYNORF1 3C11 (dilution 1:1000; DSHB AB_528479, developed by E. Buchner); Discs large 1 4F3 (dilution 1:2000; DSHB AB_528203, developed by C. Goodman); Elav 9F8A9 (dilution 1:1000; DSHB AB_528217, developed by G. M. Rubin); and Bruchpilot nc82 (dilution 1:100; DSHB AB_2314866, developed by E. Buchner). Other antibodies used were: mouse anti-Da dam109-10 (dilution 1:10; a gift from C. Cronmiller, University of Virginia, Charlottesville, VA, USA); mouse anti-FLAG M2 horseradish peroxidase (HRP)-conjugated (dilution 1:6000; Sigma-Aldrich A8592); and goat anti-mouse IgM HRP-conjugated secondary antibody (dilution 1:5000; Invitrogen 32430).

Immunohistochemical staining

The anterior parts of third instar larvae were dissected in PBS and fixed using 4% paraformaldehyde in PBS. Adult flies were first fixed in 4% paraformaldehyde in PBS and then dissected. Primary antibody labelling was performed overnight, or for 72 h with the anti-FLAG antibody, on an overhead rotator at 4°C in PBS with 0.1% Triton X-100, or 0.5% Triton X-100 for the anti-FLAG antibody. The antibodies used were as follows: Synapsin SYNORF1 3C11 (dilution 1:10; DSHB AB_528479); Discs large 1 4F3 (dilution 1:400; DSHB AB_528203; mouse anti-FLAG M2 (dilution 1:1000; Sigma-Aldrich F1804); and goat anti-mouse Alexa594 (dilution 1:1000; ImmunoResearch Laboratories 115-585-003). Secondary antibodies were pre-adsorbed to wild-type tissues before use. Incubation with secondary antibodies was performed for 3 h on an overhead rotator at room temperature in PBS with 0.1% Triton X-100. The labelled larval brains were dissected and mounted in Vectashield mounting medium (Vector Laboratories). For image collection, a Zeiss LSM 510 Meta confocal microscope with a Pln Apo 20×/0.8 DICII objective or a Pln Apo 63×/1.4 Oil DICII objective was used. Suitable layers were selected using Imaris software (Bitplane).

Appetitive associative learning assay

The appetitive associative memory assay in the *Drosophila* larvae was performed as described previously (Michels et al., 2017). Briefly, the larvae were trained three times for 5 min on Petri dishes; the odour-amyl acetate (AM) was presented on plain agar and odour–OCT on agar containing fructose as a reward. Then, the larvae were placed in the midline of a plain agar plate and given a choice between the two odours placed on separate halves of the Petri dish. After 3 min, larvae were counted on each half of the Petri dish. Then reciprocal training was performed with AM and fructose and OCT with plain agar. Using data from two reciprocally trained tests, the PI was calculated $PI=(PREF\ AM_{AM+/OCT-}-PREF\ AM_{AM/OCT+})/2$. The odours and the reward were presented in four different orders to eliminate any non-specific preferences. Altogether, 12 training and test cycles were conducted per genotype, each time with new larvae, and PIs were calculated and used for statistical analysis. The PIs were visualized as box-whisker plots that showed the median, the 25% and 75% quantiles and the minimum and maximum. For statistical analysis inside one genotype, a one-sample sign test was applied with an error threshold smaller than 5% and between the groups Kruskal–Wallis ANOVA with Dunn's post-hoc test was used. SAHA was dissolved in dimethyl sulfoxide (DMSO) and the same concentration of DMSO (0.1%) was used in the food substrate for a control. Resveratrol was dissolved in 96% ethanol and 1% ethanol in the food that was used for the control.

ChIP

Chromatin preparations were carried out as described previously (Chanas et al., 2004). Adult heads (~150 mg) were collected on dry ice and homogenized in buffer A1 [60 mM KCl, 15 mM NaCl, 4 mM $MgCl_2$, 15 mM HEPES (pH 7.6), 0.5% Triton X-100, 0.5 mM DTT, 10 mM sodium butyrate and 1× EDTA-free protease inhibitor cocktail (Roche)] with 1.8% formaldehyde at room temperature using a Kontes pellet pestle followed by three strokes using a Dounce homogenizer with a loose pestle. Homogenate was incubated for 15 min and glycin was added to 225 mM of the homogenate followed by 5 min incubation. The homogenate was then centrifuged for 5 min at 4000 ***g*** at 4°C and the supernatant was discarded. The pellet was washed three times with 3 ml of buffer A1 followed by a wash with 3 ml of lysis buffer [14 mM NaCl, 15 mM HEPES (pH 7.6), 1 mM EDTA, 0.5 mM EGTA, 1% Triton X-100, 0.5 mM DTT, 0.1% sodium deoxycholate, 0.05% SDS, 10 mM sodium butyrate and 1× EDTA-free protease inhibitor cocktail (Roche)]. Crosslinked material was resuspended in 0.5 ml of lysis buffer with 0.1% SDS and 0.5% N-lauroylsarcosine, and incubated for 10 min at 4°C on a rotator followed by sonication using a Sonics Vibra-Cell processor at 70% amplitude for 30 times at 15 s intervals. Crosslinked material was then rotated for 10 min at 4°C and centrifuged for 5 min at room temperature at maximum speed. The supernatant was then transferred to a new tube and 0.5 ml of lysis buffer was added to the pellet followed by rotation and centrifugation. Supernatants were combined and centrifuged two times for 10 min each time at maximum

speed. Chromatin extract was transferred to Microcon DNA Fast Flow Centrifugal Filter Units (Merck Millipore), blocked with 1 mg/ml bovine serum albumin in PBS, and purified using lysis buffer. The volume of chromatin extract was brought to 1 ml using lysis buffer. Protein concentrations were determined using a bicinchoninic acid assay (Pierce).

After removing equal amounts of inputs, chromatin extracts were diluted 10× using dilution buffer [1% Triton X-100, 150 mM NaCl, 2 mM EDTA (pH 8.0), 20 mM Tris-HCl (pH 8.0) and 1× EDTA-free protease inhibitor cocktail (Roche)] and added to 50 µl of Dynabeads Protein G (Invitrogen) beads that were previously incubated with 5 µg of monoclonal anti-FLAG M2 antibody (Sigma-Aldrich F1804) in 400 µl of 0.05% PBS Tween 20 overnight (antibody dilution 1:80). ChIP was carried out overnight at 4°C. Beads with chromatin were then washed in wash buffer [1% Triton X-100, 0.1% SDS, 150 mM NaCl, 2 mM EDTA (pH 8.0), 20 mM Tris-HCl (pH 8.0) and 1× EDTA-free protease inhibitor cocktail (Roche)] using a magnetic rack for 10 min for three times at 4°C on a rotator, followed by final wash with wash buffer [1% Triton X-100, 0.1% SDS, 500 mM NaCl, 2 mM EDTA (pH 8.0), 20 mM Tris-HCl (pH 8.0) and 1× EDTA-free protease inhibitor cocktail (Roche)]. Chromatin was eluted using three aliquots of 50 µl elution buffer (1% SDS, 100 mM $NaHCO_3$ and 1 mM EDTA) for 10 min each time at 37°C. The volume of inputs was brought to 150 µl with elution buffer. For decrosslinking, 8 µl of 5 M NaCl was added and the samples were incubated at 65°C overnight. Then, 2 µl of RNase A (10 mg/ml) was added and the samples were incubated at 37°C for 30 min, followed by incubation with 2 µl of EDTA (0.5 M) and 4 µl Proteinase K (10 mg/ml) at 45°C for 30 min. DNA was extracted using a QIAquick PCR Purification Kit (Qiagen).

qPCR

For RT-qPCR, 15 heads were collected from 2- to 3-day-old adult flies on dry ice. RNA was extracted using an RNeasy Mini Kit (Qiagen). cDNA was synthesized with Superscript IV Reverse Transcriptase (Invitrogen) and oligo$(dT)_{20}$ primers. qPCR was performed using a LightCycler 480 II (Roche) with Hot FIREPol EvaGreen qPCR Mix Plus (Solis Biodyne). Primer sequences are shown in Table S1.

Negative geotaxis assay

Ten females and males were separated in fresh vials 48 h before the assay to allow recovery from anaesthesia. Before the test, males and females from control and *da* silencing groups were transferred to empty vials without anaesthesia, which were closed with another upside down vial using sticky tape. The flies were knocked down three times on the table and a photo was taken after 10 s. The height of the vial was divided into ten equal parts and the number of flies in each compartment was counted, and average height was calculated. The experiment was repeated five times, each time with new flies. Average climbing heights were visualized using box-whisker plots that showed the median, the 25% to 75% quantiles and the minimum and maximum. For statistical significance, pairwise Mann–Whitney U tests were used.

Acknowledgements

We thank Mark Fortini, Bertram Gerber, Riitta Lindström, and Claire Cronmiller for sharing fly stocks and reagents. We thank Epp Väli for technical assistance; and Kaisa Roots, Hanna Vihma, Nadežda Pupina and Jan Erik Alliksaar for their contribution in the early stages of this study. We are grateful to Allan-Hermann Pool and Richard Tamme for critical comments on the manuscript. We thank the 'TUT Institutional Development Program for 2016-2022' Graduate School in Clinical Medicine, which received funding from the European Regional Development Fund under program ASTRA 2014-2020.4.01.16-0032 in Estonia. Stocks obtained from the Bloomington *Drosophila* Stock Center (NIH P40OD018537) were used in this study. We also acknowledge the *Drosophila* Genomics Resource Center, which is supported by NIH grant 2P40OD010949.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: L.T., M.S., T.T., M.P.; Methodology: L.T., M.J., J.T., M.P.; Formal analysis: L.T., A. Sirp, J.T., A. Shubina, C.S.K., K.N.; Investigation: L.T., M.J., K.S., A. Sirp, J.T., A. Shubina, C.S.K., K.N., M.P.; Resources: T.T.; Writing - original draft: L.T.; Writing - review & editing: L.T., A. Sirp, J.T., K.N., M.S., T.T., M.P.; Visualization: L.T., M.P.; Supervision: L.T., A. Sirp, M.S., M.P.; Project administration: T.T., M.P.; Funding acquisition: M.S., T.T.

Funding

This study was supported by the Estonian Research Council (IUT19-18 and PRG805 to T.T.); the European Union through the European Regional Development Fund [Project No. 2014-2020.4.01.15-0012 to T.T.) and H2020-MSCA-RISE-2016 (EU734791 to T.T.)]; the Pitt Hopkins Research Foundation (grants 8 and 21 to T.T and M.S.); and the Million Dollar Bike Ride Pilot Grant Program for Rare Disease Research at the University of Pennsylvania Orphan Disease Center (MDBR-16-122-PHP and MDBR-17-127-Pitt Hopkins to T.T and M.S).

Supplementary information

Supplementary information available online at
https://dmm.biologists.org/lookup/doi/10.1242/dmm.042747.supplemental

References

Amiel, J., Rio, M., de Pontual, L., Redon, R., Malan, V., Boddaert, N., Plouin, P., Carter, N. P., Lyonnet, S., Munnich, A. et al. (2007). Mutations in TCF4, encoding a class I basic helix-loop-helix transcription factor, are responsible for Pitt-Hopkins syndrome, a severe epileptic encephalopathy associated with autonomic dysfunction. *Am. J. Hum. Genet.* **80**, 988-993. doi:10.1086/515582

Andrade-Zapata, I. and Baonza, A. (2014). The bHLH factors extramacrochaetae and daughterless control cell cycle in Drosophila imaginal discs through the transcriptional regulation of the Cdc25 phosphatase string. *PLoS Genet.* **10**, e1004233. doi:10.1371/journal.pgen.1004233

Bardin, A. J., Perdigoto, C. N., Southall, T. D., Brand, A. H. and Schweisguth, F. (2010). Transcriptional control of stem cell maintenance in the *Drosophila* intestine. *Development* **137**, 705-714. doi:10.1242/dev.039404

Benfenati, F. (2011). Synapsins—Molecular function, development and disease. *Semin. Cell Dev. Biol.* **22**, 377. doi:10.1016/j.semcdb.2011.07.016

Bhattacharya, A. and Baker, N. E. (2011). A network of broadly expressed HLH genes regulates tissue-specific cell fates. *Cell* **147**, 881-892. doi:10.1016/j.cell.2011.08.055

Bhattacharya, A. and Baker, N. E. (2012). The role of the bHLH protein hairy in morphogenetic furrow progression in the developing Drosophila eye. *PLoS ONE* **7**, e47503. doi:10.1371/journal.pone.0047503

Brand, A. H. and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401-415.

Brockschmidt, A., Todt, U., Ryu, S., Hoischen, A., Landwehr, C., Birnbaum, S., Frenck, W., Radlwimmer, B., Lichter, P., Engels, H. et al. (2007). Severe mental retardation with breathing abnormalities (Pitt-Hopkins syndrome) is caused by haploinsufficiency of the neuronal bHLH transcription factor TCF4. *Hum. Mol. Genet.* **16**, 1488-1494. doi:10.1093/hmg/ddm099

Brown, N. L., Paddock, S. W., Sattler, C. A., Cronmiller, C., Thomas, B. J. and Carroll, S. B. (1996). daughterless is required for Drosophila photoreceptor cell determination, eye morphogenesis, and cell cycle progression. *Dev. Biol.* **179**, 65-78. doi:10.1006/dbio.1996.0241

Cabrera, C. V. and Alonso, M. C. (1991). Transcriptional activation by heterodimers of the achaete-scute and daughterless gene products of Drosophila. *EMBO J.* **10**, 2965-2973. doi:10.1002/j.1460-2075.1991.tb07847.x

Castanon, I., Von Stetina, S., Kass, J. and Baylies, M. K. (2001). Dimerization partners determine the activity of the Twist bHLH protein during Drosophila mesoderm development. *Development* **128**, 3145-3159.

Caudy, M., Vässin, H., Brand, M., Tuma, R., Jan, L. Y. and Jan, Y. N. (1988). daughterless, a Drosophila gene essential for both neurogenesis and sex determination, has sequence similarities to myc and the achaete-scute complex. *Cell* **55**, 1061-1067. doi:10.1016/0092-8674(88)90250-4

Chanas, G., Lavrov, S., Iral, F., Cavalli, G. and Maschat, F. (2004). Engrailed and polyhomeotic maintain posterior cell identity through cubitus-interruptus regulation. *Dev. Biol.* **272**, 522-535. doi:10.1016/j.ydbio.2004.05.020

Chen, T., Wu, Q., Zhang, Y., Lu, T., Yue, W. and Zhang, D. (2016). Tcf4 controls neuronal migration of the cerebral cortex through regulation of Bmp7. *Front. Mol. Neurosci.* **9**, 94. doi:10.3389/fnmol.2016.00094

Cisse, B., Caton, M. L., Lehner, M., Maeda, T., Scheu, S., Locksley, R., Holmberg, D., Zweier, C., den Hollander, N. S., Kant, S. G. et al. (2008). Transcription factor E2-2 is an essential and specific regulator of plasmacytoid dendritic cell development. *Cell* **135**, 37-48. doi:10.1016/j.cell.2008.09.016

Cline, T. W. (1978). Two closely linked mutations in Drosophila melanogaster that are lethal to opposite sexes and interact with daughterless. *Genetics* **90**, 683-698.

Cronmiller, C. and Cummings, C. A. (1993). The daughterless gene product in Drosophila is a nuclear protein that is broadly expressed throughout the organism during development. *Mech. Dev.* **42**, 159-169. doi:10.1016/0925-4773(93)90005-I

Crux, S., Herms, J. and Dorostkar, M. M. (2018). Tcf4 regulates dendritic spine density and morphology in the adult brain. *PLoS ONE* **13**, e0199359. doi:10.1371/journal.pone.0199359

Cummings, C. A. and Cronmiller, C. (1994). The daughterless gene functions together with Notch and Delta in the control of ovarian follicle development in Drosophila. *Development* **120**, 381-394.
Cuthbert, P. C., Stanford, L. E., Coba, M. P., Ainge, J. A., Fink, A. E., Opazo, P., Delgado, J. Y., Komiyama, N. H., O'Dell, T. J. and Grant, S. G. N. (2007). Synapse-associated protein 102/dlgh3 Couples the NMDA receptor to specific plasticity pathways and learning strategies. *J. Neurosci. Off. J. Soc. Neurosci.* **27**, 2673-2682. doi:10.1523/JNEUROSCI.4457-06.2007
Davie, K., Janssens, J., Koldere, D., De Waegeneer, M., Pech, U., Kreft, Ł., Aibar, S., Makhzami, S., Christiaens, V., Bravo González-Blas, C. et al. (2018). A single-cell transcriptome atlas of the aging Drosophila brain. *Cell* **174**, 982-998.e20. doi:10.1016/j.cell.2018.05.057
Diegelmann, S., Klagges, B., Michels, B., Schleyer, M. and Gerber, B. (2013). Maggot learning and Synapsin function. *J. Exp. Biol.* **216**, 939-951. doi:10.1242/jeb.076208
Dietzl, G., Chen, D., Schnorrer, F., Su, K.-C., Barinova, Y., Fellner, M., Gasser, B., Kinsey, K., Oppel, S., Scheiblauer, S. et al. (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature* **448**, 151-156. doi:10.1038/nature05954
Doostparast Torshizi, A., Armoskus, C., Zhang, H., Forrest, M. P., Zhang, S., Souaiaia, T., Evgrafov, O. V., Knowles, J. A., Duan, J. and Wang, K. (2019). Deconvolution of transcriptional networks identifies TCF4 as a master regulator in schizophrenia. *Sci. Adv.* **5**, eaau4139. doi:10.1126/sciadv.aau4139
Ekins, S., Gerlach, J., Zorn, K. M., Antonio, B. M., Lin, Z. and Gerlach, A. (2019). Repurposing Approved Drugs as Inhibitors of Kv7.1 and Nav1.8 to Treat Pitt Hopkins Syndrome. *Pharm. Res.* **36**, 137. doi:10.1007/s11095-019-2671-y
Fenckova, M., Blok, L. E. R., Asztalos, L., Goodman, D. P., Cizek, P., Singgih, E. L., Glennon, J. C., IntHout, J., Zweier, C., Eichler, E. E., et al. (2019). Habituation learning Is a widely affected mechanism in Drosophila models of intellectual disability and autism spectrum disorders. *Biol. Psychiatry* **86**, 294-305. doi:10.1016/j.biopsych.2019.04.029
Forrest, M., Chapman, R. M., Doyle, A. M., Tinsley, C. L., Waite, A. and Blake, D. J. (2012). Functional analysis of TCF4 missense mutations that cause Pitt-Hopkins syndrome. *Hum. Mutat.* **33**, 1676-1686. doi:10.1002/humu.22160
Forrest, M. P., Waite, A. J., Martin-Rendon, E. and Blake, D. J. (2013). Knockdown of human TCF4 affects multiple signaling pathways involved in cell survival, epithelial to mesenchymal transition and neuronal differentiation. *PLoS ONE* **8**, e73169. doi:10.1371/journal.pone.0073169
Forrest, M. P., Hill, M. J., Kavanagh, D. H., Tansey, K. E., Waite, A. J. and Blake, D. J. (2018). The psychiatric risk gene Transcription Factor 4 (TCF4) regulates neurodevelopmental pathways associated with schizophrenia, autism, and intellectual disability. *Schizophr. Bull.* **44**, 1100-1110. doi:10.1093/schbul/sbx164
Garcia, C. C., Blair, H. J., Seager, M., Coulthard, A., Tennant, S., Buddles, M., Curtis, A. and Goodship, J. A. (2004). Identification of a mutation in synapsin I, a synaptic vesicle protein, in a family with epilepsy. *J. Med. Genet.* **41**, 183-186. doi:10.1136/jmg.2003.013680
Giebel, B., Stüttem, I., Hinz, U. and Campos-Ortega, J. A. (1997). Lethal of Scute requires overexpression of Daughterless to elicit ectopic neuronal development during embryogenesis in Drosophila. *Mech. Dev.* **63**, 75-87. doi:10.1016/S0925-4773(97)00029-4
Gitler, D., Takagishi, Y., Feng, J., Ren, Y., Rodriguiz, R. M., Wetsel, W. C., Greengard, P. and Augustine, G. J. (2004). Different presynaptic roles of synapsins at excitatory and inhibitory synapses. *J. Neurosci.* **24**, 11368-11380. doi:10.1523/JNEUROSCI.3795-04.2004
Grajkowska, L. T., Ceribelli, M., Lau, C. M., Warren, M. E., Tiniakou, I., Higa, S. N., Bunin, A., Haecker, H., Mirny, L. A., Staudt, L. M. et al. (2017). Isoform-specific expression and feedback regulation of E protein TCF4 control dendritic cell lineage specification. *Immunity* **46**, 65-77. doi:10.1016/j.immuni.2016.11.006
Hennig, K. M., Fass, D. M., Zhao, W.-N., Sheridan, S. D., Fu, T., Erdin, S., Stortchevoi, A., Lucente, D., Cody, J. D., Sweetser, D. et al. (2017). WNT/β-catenin pathway and epigenetic mechanisms regulate the Pitt-Hopkins syndrome and schizophrenia risk gene TCF4. *Mol. Neuropsychiatry* **3**, 53-71. doi:10.1159/000475666
Hill, M. J., Killick, R., Navarrete, K., Maruszak, A., McLaughlin, G. M., Williams, B. P. and Bray, N. J. (2017). Knockdown of the schizophrenia susceptibility gene TCF4 alters gene expression and proliferation of progenitor cells from the developing human neocortex. *J. Psychiatry Neurosci.* **42**, 181-188. doi:10.1503/jpn.160073
Horn, C., Offen, N., Nystedt, S., Häcker, U. and Wimmer, E. A. (2003). piggyBac-based insertional mutagenesis and enhancer detection as a tool for functional insect genomics. *Genetics* **163**, 647-661.
Jung, M., Häberle, B. M., Tschaikowsky, T., Wittmann, M.-T., Balta, E.-A., Stadler, V.-C., Zweier, C., Dörfler, A., Gloeckner, C. J. and Lie, D. C. (2018). Analysis of the expression pattern of the schizophrenia-risk and intellectual disability gene TCF4 in the developing and adult brain suggests a role in development and plasticity of cortical and hippocampal neurons. *Mol. Autism* **9**, 20. doi:10.1186/s13229-018-0200-1
Kennedy, A. J., Rahn, E. J., Paulukaitis, B. S., Savell, K. E., Kordasiewicz, H. B., Wang, J., Lewis, J. W., Posey, J., Strange, S. K., Guzman-Karlsson, M. C. et al. (2016). Tcf4 regulates synaptic plasticity, DNA methylation, and memory function. *Cell Rep.* **16**, 2666-2685. doi:10.1016/j.celrep.2016.08.004
Kepa, A., Medina, L. M., Erk, S., Srivastava, D. P., Fernandes, A., Toro, R., Lévi, S., Ruggeri, B., Fernandes, C., Degenhardt, F. et al. (2017). Associations of the intellectual disability gene MYT1L with helix-loop-helix gene expression, hippocampus volume and hippocampus activation during memory retrieval. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.* **42**, 2516-2526. doi:10.1038/npp.2017.91
King-Jones, K., Korge, G. and Lehmann, M. (1999). The helix-loop-helix proteins dAP-4 and daughterless bind both in vitro and in vivo to SEBP3 sites required for transcriptional activation of the Drosophila gene Sgs-4. *J. Mol. Biol.* **291**, 71-82. doi:10.1006/jmbi.1999.2963
Klagges, B. R. E., Heimbeck, G., Godenschwege, T. A., Hofbauer, A., Pflugfelder, G. O., Reifegerste, R., Reisch, D., Schaupp, M., Buchner, S. and Buchner, E. (1996). Invertebrate synapsins: a single gene codes for several isoforms in *Drosophila*. *J. Neurosci.* **16**, 3154-3165. doi:10.1523/JNEUROSCI.16-10-03154.1996
Lennertz, L., Rujescu, D., Wagner, M., Frommann, I., Schulze-Rauschenbach, S., Schuhmacher, A., Landsberg, M. W., Franke, P., Möller, H.-J., Wölwer, W. et al. (2011a). Novel schizophrenia risk gene TCF4 influences verbal learning and memory functioning in schizophrenia patients. *Neuropsychobiology* **63**, 131-136. doi:10.1159/000317844
Lennertz, L., Quednow, B. B., Benninghoff, J., Wagner, M., Maier, W. and Mössner, R. (2011b). Impact of TCF4 on the genetics of schizophrenia. *Eur. Arch. Psychiatry Clin. Neurosci.* **261** Suppl. 2, S161-S165. doi:10.1007/s00406-011-0256-9
Li, K. and Baker, N. E. (2018). Regulation of the Drosophila ID protein Extra macrochaetae by proneural dimerization partners. *eLife* **7**, e33967. doi:10.7554/eLife.33967.020
Li, H., Zhu, Y., Morozov, Y. M., Chen, X., Page, S. C., Rannals, M. D., Maher, B. J. and Rakic, P. (2019). Disruption of TCF4 regulatory networks leads to abnormal cortical development and mental disabilities. *Mol. Psychiatry* **24**, 1235-1246. doi:10.1038/s41380-019-0353-0
Luo, L., Liao, Y. J., Jan, L. Y. and Jan, Y. N. (1994). Distinct morphogenetic functions of similar small GTPases: Drosophila Drac1 is involved in axonal outgrowth and myoblast fusion. *Genes Dev.* **8**, 1787-1802. doi:10.1101/gad.8.15.1787
MacArthur, S., Li, X.-Y., Li, J., Brown, J. B., Chu, H. C., Zeng, L., Grondona, B. P., Hechmer, A., Simirenko, L., Keränen, S. V. E. et al. (2009). Developmental roles of 21 Drosophila transcription factors are determined by quantitative differences in binding to an overlapping set of thousands of genomic regions. *Genome Biol.* **10**, R80. doi:10.1186/gb-2009-10-7-r80
Massari, M. E. and Murre, C. (2000). Helix-Loop-Helix proteins: regulators of transcription in eucaryotic organisms. *Mol. Cell. Biol.* **20**, 429-440. doi:10.1128/MCB.20.2.429-440.2000
Michels, B., Diegelmann, S., Tanimoto, H., Schwenkert, I., Buchner, E. and Gerber, B. (2005). A role for Synapsin in associative learning: the Drosophila larva as a study case. *Learn. Mem.* **12**, 224-231. doi:10.1101/lm.92805
Michels, B., Chen, Y., Saumweber, T., Mishra, D., Tanimoto, H., Schmid, B., Engmann, O. and Gerber, B. (2011). Cellular site and molecular mode of synapsin action in associative learning. *Learn. Mem.* **18**, 332-344. doi:10.1101/lm.2101411
Michels, B., Saumweber, T., Biernacki, R., Thum, J., Glasgow, R. D. V., Schleyer, M., Chen, Y., Eschbach, C., Stocker, R. F., Toshima, N. et al. (2017). Pavlovian conditioning of larval Drosophila: an illustrated, multilingual, hands-on manual for odor-taste associative learning in maggots. *Front. Behav. Neurosci.* **11**, 45. doi:10.3389/fnbeh.2017.00045
Murre, C., Bain, G., van Dijk, M. A., Engel, I., Furnari, B. A., Massari, M. E., Matthews, J. R., Quong, M. W., Rivera, R. R. and Stuiver, M. H. (1994). Structure and function of helix-loop-helix proteins. *Biochim. Biophys. Acta* **1218**, 129-135. doi:10.1016/0167-4781(94)90001-9
Page, S. C., Hamersky, G. R., Gallo, R. A., Rannals, M. D., Calcaterra, N. E., Campbell, M. N., Mayfield, B., Briley, A., Phan, B. N., Jaffe, A. E. et al. (2018). The schizophrenia- and autism-associated gene, transcription factor 4 regulates the columnar distribution of layer 2/3 prefrontal pyramidal neurons in an activity-dependent manner. *Mol. Psychiatry* **23**, 304-315. doi:10.1038/mp.2017.37
Park, D., Shafer, O. T., Shepherd, S. P., Suh, H., Trigg, J. S. and Taghert, P. H. (2008). The Drosophila basic helix-loop-helix protein DIMMED directly activates PHM, a gene encoding a neuropeptide-amidating enzyme. *Mol. Cell. Biol.* **28**, 410-421. doi:10.1128/MCB.01104-07
Park, S.-J., Ahmad, F., Philp, A., Baar, K., Williams, T., Luo, H., Ke, H., Rehmann, H., Taussig, R., Brown, A. L. et al. (2012). Resveratrol ameliorates aging-related metabolic phenotypes by inhibiting cAMP phosphodiesterases. *Cell* **148**, 421-433. doi:10.1016/j.cell.2012.01.017
Pfeiffer, B. D., Jenett, A., Hammonds, A. S., Ngo, T.-T. B., Misra, S., Murphy, C., Scully, A., Carlson, J. W., Wan, K. H., Laverty, T. R. et al. (2008). Tools for neuroanatomy and neurogenetics in Drosophila. *Proc. Natl. Acad. Sci. USA* **105**, 9715-9720. doi:10.1073/pnas.0803697105
Powell, L. M., Deaton, A. M., Wear, M. A. and Jarman, A. P. (2008). Specificity of Atonal and Scute bHLH factors: analysis of cognate E box binding sites and the

influence of Senseless. *Genes Cells Devoted Mol. Cell. Mech.* **13**, 915-929. doi:10.1111/j.1365-2443.2008.01217.x

Quednow, B. B., Ettinger, U., Mössner, R., Rujescu, D., Giegling, I., Collier, D. A., Schmechtig, A., Kühn, K.-U., Möller, H.-J., Maier, W. et al. (2011). The schizophrenia risk allele C of the TCF4 rs9960767 polymorphism disrupts sensorimotor gating in schizophrenia spectrum and healthy volunteers. *J. Neurosci. Off. J. Soc. Neurosci.* **31**, 6684-6691. doi:10.1523/JNEUROSCI.0526-11.2011

Rannals, M. D. and Maher, B. J. (2017). Molecular mechanisms of transcription factor 4 in Pitt-Hopkins syndrome. *Curr. Genet. Med. Rep.* **5**, 1-7. doi:10.1007/s40142-017-0110-0

Rannals, M. D., Page, S. C., Campbell, M. N., Gallo, R. A., Mayfield, B. and Maher, B. J. (2016). Neurodevelopmental models of transcription factor 4 deficiency converge on a common ion channel as a potential therapeutic target for Pitt Hopkins syndrome. *Rare Dis. Austin Tex* **4**, e1220468. doi:10.1080/21675511.2016.1220468

Sepp, M., Kannike, K., Eesmaa, A., Urb, M. and Timmusk, T. (2011). Functional diversity of human basic helix-loop-helix transcription factor TCF4 isoforms generated by alternative 5′ exon usage and splicing. *PLoS ONE* **6**, e22138. doi:10.1371/journal.pone.0022138

Sepp, M., Pruunsild, P. and Timmusk, T. (2012). Pitt-Hopkins syndrome-associated mutations in TCF4 lead to variable impairment of the transcription factor function ranging from hypomorphic to dominant-negative effects. *Hum. Mol. Genet.* **21**, 2873-2888. doi:10.1093/hmg/dds112

Sepp, M., Vihma, H., Nurm, K., Urb, M., Page, S. C., Roots, K., Hark, A., Maher, B. J., Pruunsild, P. and Timmusk, T. (2017). The intellectual disability and schizophrenia associated transcription factor TCF4 is regulated by neuronal activity and protein kinase A. *J. Neurosci.* **37**, 10516-10527. doi:10.1523/JNEUROSCI.1151-17.2017

Silva, A. J., Rosahl, T. W., Chapman, P. F., Marowitz, Z., Friedman, E., Frankland, P. W., Cestari, V., Cioffi, D., Südhof, T. C. and Bourtchuladze, R. (1996). Impaired learning in mice with abnormal short-lived plasticity. *Curr. Biol.* **6**, 1509-1518. doi:10.1016/S0960-9822(96)00756-7

Smith, J. E. and Cronmiller, C. (2001). The Drosophila daughterless gene autoregulates and is controlled by both positive and negative cis regulation. *Development* **128**, 4705-4714. doi:10.3410/f.1002603.28865

Smith, J. E., Cummings, C. A. and Cronmiller, C. (2002). Daughterless coordinates somatic cell proliferation, differentiation and germline cyst survival during follicle formation in Drosophila. *Development* **129**, 3255-3267.

Talkowski, M. E., Rosenfeld, J. A., Blumenthal, I., Pillalamarri, V., Chiang, C., Heilbut, A., Ernst, C., Hanscom, C., Rossin, E., Lindgren, A. M. et al. (2012). Sequencing chromosomal abnormalities reveals neurodevelopmental loci that confer risk across diagnostic boundaries. *Cell* **149**, 525-537. doi:10.1016/j.cell.2012.03.028

Tamberg, L., Sepp, M., Timmusk, T. and Palgi, M. (2015). Introducing Pitt-Hopkins syndrome-associated mutations of TCF4 to Drosophila daughterless. *Biol. Open* **4**, 1762-1771. doi:10.1242/bio.014696

Tanaka-Matakatsu, M., Miller, J., Borger, D., Tang, W.-J. and Du, W. (2014). Daughterless homodimer synergizes with Eyeless to induce Atonal expression and retinal neuron differentiation. *Dev. Biol.* **392**, 256-265. doi:10.1016/j.ydbio.2014.05.019

Tarpey, P., Parnau, J., Blow, M., Woffendin, H., Bignell, G., Cox, C., Cox, J., Davies, H., Edkins, S., Holden, S. et al. (2004). Mutations in the DLG3 gene cause nonsyndromic X-linked mental retardation. *Am. J. Hum. Genet.* **75**, 318-324. doi:10.1086/422703

Thaxton, C., Kloth, A. D., Clark, E. P., Moy, S. S., Chitwood, R. A. and Philpot, B. D. (2018). Common pathophysiology in multiple mouse models of Pitt–Hopkins syndrome. *J. Neurosci.* **38**, 918-936. doi:10.1523/JNEUROSCI.1305-17.2017

Wong, M.-C., Castanon, I. and Baylies, M. K. (2008). Daughterless dictates Twist activity in a context-dependent manner during somatic myogenesis. *Dev. Biol.* **317**, 417-429. doi:10.1016/j.ydbio.2008.02.020

Wu, Y., Bolduc, F. V., Bell, K., Tully, T., Fang, Y., Sehgal, A. and Fischer, J. A. (2008). A Drosophila model for Angelman syndrome. *Proc. Natl. Acad. Sci. USA* **105**, 12399-12404. doi:10.1073/pnas.0805291105

Wu, Q., Yu, G., Park, S. J., Gao, Y., Ja, W. W. and Yang, M. (2020). Excreta Quantification (EX-Q) for longitudinal measurements of food intake in Drosophila. *iScience* **23**, 100776. doi:10.1016/j.isci.2019.100776

Xia, H., Jahr, F. M., Kim, N.-K., Xie, L., Shabalin, A. A., Bryois, J., Sweet, D. H., Kronfol, M. M., Palasuberniam, P., McRae, M. P. et al. (2018). Building a schizophrenia genetic network: transcription factor 4 regulates genes involved in neuronal development and schizophrenia risk. *Hum. Mol. Genet.* **27**, 3246-3256. doi:10.1093/hmg/ddy222

Yao Yang, M., Armstrong, J. D., Vilinsky, I., Strausfeld, N. J. and Kaiser, K. (1995). Subdivision of the drosophila mushroom bodies by enhancer-trap expression patterns. *Neuron* **15**, 45-54. doi:10.1016/0896-6273(95)90063-2

Yasugi, T., Fischer, A., Jiang, Y., Reichert, H. and Knoblich, J. A. (2014). A regulatory transcriptional loop controls proliferation and differentiation in Drosophila neural stem cells. *PLoS ONE* **9**, e97034. doi:10.1371/journal.pone.0097034

Zanni, G., van Esch, H., Bensalem, A., Saillour, Y., Poirier, K., Castelnau, L., Ropers, H. H., de Brouwer, A. P. M., Laumonnier, F., Fryns, J.-P. et al. (2010). A novel mutation in the DLG3 gene encoding the synapse-associated protein 102 (SAP102) causes non-syndromic mental retardation. *Neurogenetics* **11**, 251-255. doi:10.1007/s10048-009-0224-y

Zollino, M., Zweier, C., van Balkom, I. D., Sweetser, D. A., Alaimo, J., Bijlsma, E. K., Cody, J., Elsea, S. H., Giurgea, I., Macchiaiolo, M. et al. (2019). Diagnosis and management in Pitt-Hopkins syndrome: first international consensus statement. *Clin. Genet.* **95**, 462-478. doi:10.1111/cge.13506

Zweier, C., Peippo, M. M., Hoyer, J., Sousa, S., Bottani, A., Clayton-Smith, J., Reardon, W., Saraiva, J., Cabral, A., Göhring, I. et al. (2007). Haploinsufficiency of TCF4 causes syndromal mental retardation with intermittent hyperventilation (Pitt-Hopkins syndrome). *Am. J. Hum. Genet.* **80**, 994-1001. doi:10.1086/515583

Publication III

Sirp, A.*, Shubina, A.*, Tuvikene, J., **Tamberg, L.,** Kiir, C. S., Kranich, L. and Timmusk, T. Expression of alternative transcription factor 4 mRNAs and protein isoforms in the developing and adult rodent and human tissues.
Frontiers in Molecular Neuroscience 15, 2022. DOI 10.3389/fnmol.2022.1033224.

frontiers | Frontiers in Molecular Neuroscience

TYPE Original Research
PUBLISHED 02 November 2022
DOI 10.3389/fnmol.2022.1033224

OPEN ACCESS

EDITED BY
Enrico Tongiorgi,
University of Trieste,
Italy

REVIEWED BY
Marc Forrest,
Northwestern University,
United States
Brady J. Maher,
Lieber Institute for Brain Development,
United States
Germana Meroni,
University of Trieste,
Italy

*CORRESPONDENCE
Tõnis Timmusk
tonis.timmusk@taltech.ee

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Neuroplasticity and Development, a section of the Frontiers in Molecular Neuroscience

RECEIVED 31 August 2022
ACCEPTED 05 October 2022
PUBLISHED 02 November 2022

CITATION
Sirp A, Shubina A, Tuvikene J, Tamberg L, Kiir CS, Kranich L and Timmusk T (2022) Expression of alternative transcription factor 4 mRNAs and protein isoforms in the developing and adult rodent and human tissues.
Front. Mol. Neurosci. 15:1033224.
doi: 10.3389/fnmol.2022.1033224

Expression of alternative transcription factor 4 mRNAs and protein isoforms in the developing and adult rodent and human tissues

Alex Sirp[1†], Anastassia Shubina[1†], Jürgen Tuvikene[1,2], Laura Tamberg[1], Carl Sander Kiir[1], Laura Kranich[1] and Tõnis Timmusk[1,2*]

[1]Department of Chemistry and Biotechnology, Tallinn University of Technology, Tallinn, Estonia, [2]Protobios LLC, Tallinn, Estonia

Transcription factor 4 (TCF4) belongs to the class I basic helix–loop–helix family of transcription factors (also known as E-proteins) and is vital for the development of the nervous system. Aberrations in the *TCF4* gene are associated with several neurocognitive disorders such as schizophrenia, intellectual disability, post-traumatic stress disorder, depression, and Pitt-Hopkins Syndrome, a rare but severe autism spectrum disorder. Expression of the human *TCF4* gene can produce at least 18 N-terminally distinct protein isoforms, which activate transcription with different activities and thus may vary in their function during development. We used long-read RNA-sequencing and western blot analysis combined with the analysis of publicly available short-read RNA-sequencing data to describe both the mRNA and protein expression of the many distinct TCF4 isoforms in rodent and human neural and nonneural tissues. We show that TCF4 mRNA and protein expression is much higher in the rodent brain compared to nonneural tissues. TCF4 protein expression is highest in the rodent cerebral cortex and hippocampus, where expression peaks around birth, and in the rodent cerebellum, where expression peaks about a week after birth. In human, highest *TCF4* expression levels were seen in the developing brain, although some nonneural tissues displayed comparable expression levels to adult brain. In addition, we show for the first time that out of the many possible TCF4 isoforms, the main TCF4 isoforms expressed in the rodent and human brain and other tissues are TCF4-B, -C, -D, -A, and-I. Taken together, our isoform specific analysis of TCF4 expression in different tissues could be used for the generation of gene therapy applications for patients with TCF4-associated diseases.

KEYWORDS

transcription factor TCF4, basic helix–loop–helix transcription factor, western blot analysis, neurodevelopment, long-read RNA sequencing, brain tissue, peripheral tissue

Introduction

Transcription factor 4 (TCF4) is a member of the class I basic helix–loop–helix transcription factor family (also known as E-proteins) and is the main E-protein expressed in the adult mouse brain (Massari and Murre, 2000; Fischer et al., 2014). TCF4 regulates numerous genes involved in neurodevelopment (Forrest et al., 2018) and has been shown to mediate its function by forming either homo-or heterodimers with proneural interaction partners such as achaete-scute homolog 1 (ASCL1; Persson et al., 2000) and neurogenic differentiation factor 2 (NEUROD2; Brzózka et al., 2010) as well as negative regulators known as inhibitor of DNA binding (ID) proteins (Chiaramello et al., 1995; Einarson and Chao, 1995). The expression of TCF4 interaction partners is strictly regulated, allowing TCF4 to possibly exert different functions during the development of the nervous system (Quednow et al., 2011).

Changes in the *TCF4* gene are linked to the development of many severe neurocognitive disorders such as schizophrenia (Stefansson et al., 2009; Ripke et al., 2014; Doostparast Torshizi et al., 2019), intellectual disability (Kharbanda et al., 2016), post-traumatic stress disorder (Gelernter et al., 2019), and depression (Wray et al., 2018). In addition, *de novo* mutations in one of the *TCF4* alleles cause Pitt-Hopkins syndrome (Brockschmidt et al., 2007; Zweier et al., 2007)—an autism spectrum disorder described by severe cognitive impairment, breathing abnormalities, motor delay, and distinctive facial features (Zollino et al., 2019). Interestingly, in addition to deletions and translocations, just a single nucleotide mutation in the basic helix–loop–helix encoding domain can completely impair the normal functionality of the TCF4 protein (Amiel et al., 2007; Zweier et al., 2007; Sepp et al., 2012; Sirp et al., 2021). *Tcf4* heterozygous mutant mice exhibit memory deficits, impaired motor control, and social isolation (Kennedy et al., 2016; Thaxton et al., 2018). Similar results have been noted in *Drosophila melanogaster*, where downregulation of Daughterless, the orthologue of TCF4, impairs memory and learning (Tamberg et al., 2020). Overexpression of *Tcf4* in mouse brain causes impairments in cognition and sensorimotor gating (Brzózka et al., 2010), and increased long term depression at synapses (Badowska et al., 2020). Homozygous *Tcf4* knockout mice have low viability and usually die around birth (Zhuang et al., 1996).

Expression of *Tcf4* was first described during late embryonic and early postnatal development in different mouse brain regions using northern blot analysis and *in situ* hybridization (Soosaar et al., 1994; Pscherer et al., 1996; Ravanpay and Olson, 2008). More recent studies have used quantitative droplet digital PCR and reverse-transcription quantitative PCR to show that in the cerebral cortex *TCF4* mRNA expression peaks around birth and declines rapidly in the following 2 weeks (Li et al., 2019; Phan et al., 2020). Expression of TCF4 protein in the developing and adult mouse brain has been described in detail by Jung et al. (2018) using immunostaining with antibodies specific for longer TCF4 protein isoforms. During embryonic development of the brain, expression levels of long TCF4 isoforms are high in the areas which will develop into the cortex and the hippocampus. More specifically, long TCF4 isoforms are largely expressed in the germinal regions that will give rise to GABAergic and glutamatergic neurons of the cortex (Jung et al., 2018). In the brain of adult mice, TCF4 expression of long TCF4 isoforms is high in the cortex, hippocampus, and cerebellum (Jung et al., 2018). Similar results have been obtained by Kim and colleagues who used TCF4-GFP mice to characterize total TCF4 expression (Kim et al., 2020).

In human, TCF4 is expressed broadly, with the expression of different TCF4 isoforms varying between tissues (de Pontual et al., 2009; Sepp et al., 2011). Further analysis of human RNA sequencing (RNA-seq) data has revealed that the mRNA expression dynamics of *TCF4* in human and mouse appear to be conserved in the cerebral cortex—*TCF4* mRNA is highly expressed during fetal stages of development and reaches the maximum before birth, rapidly declines around birth until entering a relatively stable expression level from the early postnatal period to adulthood (Ma et al., 2018). As *TCF4* remains stably expressed in adult humans and rodents alike (de Pontual et al., 2009; Jung et al., 2018; Ma et al., 2018; Li et al., 2019) its expression is probably important for the normal functioning of the organism (Sarkar et al., 2021).

To date, none of the previous studies of *TCF4* mRNA (Ma et al., 2018; Li et al., 2019; Phan et al., 2020) and protein (Jung et al., 2018; Kim et al., 2020) expression have described expression of the variety of TCF4 isoforms. Expression of the mouse and human *TCF4* gene results in many different transcripts encoding N-terminally distinct protein isoforms which vary in their intracellular localization, transactivation capability (Sepp et al., 2011, 2017; Nurm et al., 2021) and possibly mediate their function depending on dosage (Ravanpay and Olson, 2008). Here, we investigated the complex expression dynamics of different TCF4 mRNAs and protein isoforms in the developing and adult rodent and human tissues. Our results can be used to estimate which TCF4 isoforms and at which proportions should be introduced into different tissues during development to generate gene therapy applications of the TCF4-associated diseases.

Abbreviations: TCF4, Transcription factor 4; RNA-seq, RNA sequencing; RT, Room temperature; RT-qPCR, Reverse transcription quantitative PCR; gRNA, Guide RNA; P, Postnatal; E, Embryonic; CTX, Cerebral cortex; HC, Hippocampi; CB, Cerebellum; OB, Olfactory bulb; HTH, Hypothalamus; MB, Midbrain; TH, Thalamus; STRT, Striatum.

Materials and methods

Direct TCF4 RNA sequencing

Total RNA was extracted from a mixture of cerebral cortices from three P3 BALB/c mice using the RNeasy lipid tissue mini kit

(Qiagen). Genomic DNA was digested on-column using RNase-Free DNase Set (Qiagen). Concentration of the purified RNA was determined with BioSpec-nano spectrophotometer (Shimadzu).

Before RNA-seq library preparation, 20 µg of P3 BALB/c cortical RNA was enriched for *Tcf4* transcripts in wash/binding buffer (0.5 M NaCl, 20 mM Tris–HCl pH 7.5, and 1 mM EDTA) with 8 µM each of three 5′ biotinylated oligonucleotides (Microsynth AG)—two oligonucleotides were complementary to the 3` untranslated region of *Tcf4*, and one was complementary to the bHLH region of *Tcf4* (Supplementary Table S1). As there are no known *Tcf4* transcripts that lack the bHLH region or exon 21, we expect that our *Tcf4* mRNA enrichment strategy is unbiased and enriches all possible TCF4 transcripts (Sepp et al., 2011; Nurm et al., 2021). First the mixture was incubated at 70°C for 2 min and then cooled to room temperature in about 30 min in a heating block. When the heating block reached 60°C, 100 units of RiboLock RNase Inhibitor (Thermo Fisher Scientific) was added to the mixture.

Next, Pierce streptavidin magnetic beads (Thermo Fisher Scientific) were prepared for the binding reaction. For that, 40 µl of magnetic beads was washed in 800 µl wash/binding buffer and then suspended in 30 µl of wash/binding buffer. The magnetic beads were then added to the previously annealed oligonucleotide-RNA mixture for the binding reaction. The bead-oligonucleotide-RNA mixture was incubated at RT for 90 min with occasional agitation by hand. After 90 min, the flow-through sample was collected and after that the beads were washed twice with 100 µl of wash/bind buffer followed by three washes with 100 µl of ice-cold low salt buffer (0.15 M NaCl, 20 mM Tris–HCl pH 7.5, and 1 mM EDTA). For elution, the magnetic beads were incubated at 70°C for 5 min in 25 µl of nuclease free water (Qiagen) twice, with the total volume of eluted RNA being 50 µl. In total, two *Tcf4* RNA enrichments were done—the first *Tcf4* enriched RNA sample was sequenced twice and the second once.

RNA sequencing library was prepared separately for each of the three sequencing experiments according to the Sequence-specific direct RNA sequencing protocol SQK-RNA002 (Oxford Nanopore Technologies). The RNAClean XP beads (Agencourt) were substituted with the Mag-Bind total pure NGS magnetic beads (Omega Bio-tek). Sequencing was done three times with the MinION sequencer, FLO-MIN106 flow-cell (new flow cells were used for each experiment), and SQK-RNA002 kit using MinKNOW software (version 3.6.5; Oxford Nanopore Technologies).

Base-calling of the direct RNA sequencing data was performed using Guppy Basecalling Software (version 4.0.11 + f1071ceb, Oxford Nanopore Technologies) with high-accuracy basecalling algorithm. Failed reads were discarded and the passed reads were mapped to mouse GRCm38.p6 genome (obtained from Gencode) using Minimap2 (version 2.17-r941) with the following settings: -ax splice-uf -k14. The generated sam files were converted to bam format using Samtools (version 1.9), the alignments of all three replicates were combined and only reads mapping to the *Tcf4* gene locus were kept. The resulting merged and filtered bam file was then converted to bed12 file format using bedtools (version 2.28.0) for easier visualization. All reads mapping to the *Tcf4* locus in bed12 format can be found in Data Sheet 1. Raw sequencing reads mapping to the *Tcf4* locus can be found in Data Sheet 2. The final data was visualized in Integrated Genomics Viewer and the transcripts encoding *Tcf4* isoforms were manually quantified. Aberrant transcripts were excluded from the analysis.

Guide RNA design and cloning

The University of California Santa Cruz Genome Browser Gateway[1] was used to define the genomic region of mouse exon 3 and exon 10a protein coding regions for guide RNA (gRNA) design. The genomic region for mouse exon 3 and exon 10a was chr18:69,347,299–69,347,369 and chr18:69,593,516–69,593,584, respectively, according to mouse GRCm38/mm10 (Dec. 2011) assembly. In total, three gRNAs were designed for exon 3 and two for exon 10a using Benchling Inc.[2] CRISPR guide design software.

To insert the gRNA targeting region-containing oligonucleotides effectively into the PX459 (Addgene #62988) expression vector, nucleotides were added to the 5` ends of gRNAs that were complementary to the sticky ends produced after restriction of the PX459 plasmid with the BbsI restriction enzyme (Thermo Scientific). In addition, a guanine nucleotide was added to the 5′ end of each forward oligonucleotide sequence of gRNA as it has been found to increase targeting efficiency.[3] The designed sequences are included in Supplementary Table S2. The oligonucleotides were ordered from Microsynth AG.

Cell culture and transfection

Mouse Neuro2a and human SH-SY5Y cells were grown in DMEM (Dulbecco's modified Eagle's medium, Thermo Scientific) medium, supplemented with 10% fetal bovine serum (Pan Biotech), 100 U/ml penicillin, and 0.1 mg/ml streptomycin (Thermo Scientific).

For transfection, cells were plated on a 12-well plate (Greiner) in 800 µl medium per well 24–48 h before transfection. At the time of the transfection, the cells were at 50–70% confluency. Neuro2a cells were transfected with 500 ng of the pEGFP plasmid and 500 ng of the PX459 plasmid expressing the respective gRNA, Cas9, and Puromycin resistance gene using Lipofectamine 2000 (Invitrogen). In each experiment, DNA to transfection reagent ratio was 1:2. For preparation of protein lysates, the cells were lysed in in 1x Laemmli buffer [0.062 M Tris–HCl pH 6.8, 2% SDS, 5% 2-mercaptoethanol (Roth), 10% glycerol, and 0.01% bromophenol blue].

1 https://genome.ucsc.edu
2 https://benchling.com
3 http://www.addgene.org/crispr/zhang/

Reverse transcription PCR

Total RNA was extracted from Neuro2a cells using the RNeasy mini kit (Qiagen). Genomic DNA was digested on-column using RNase-Free DNase Set (Qiagen). Concentrations of the purified RNAs were determined with BioSpec-nano spectrophotometer (Shimadzu). cDNA was synthesized from Neuro2a total RNA using Superscript IV Reverse Transcriptase (Invitrogen) according to the manufacturer's instructions. Primers used for reverse transcription PCR are listed in Supplementary Table S3.

Animal husbandry

The protocols involving animals were approved by the ethics committee of animal experiments at Ministry of Agriculture of Estonia (Permit Number: 45). All experiments were performed in accordance with the relevant guidelines and regulations. WISTAR rats (RccHan:WI, Envigo) and C57BL/6 and BALB/c mouse strains (Envigo) were used in this study. Animals were maintained in conventional polycarbonate or H-TEMP polysulfone cages (2–4 animals per cage) with *ad libitum* access to clean water and food pellets (ssniff Spezialdiäten) under a 12-h light/dark cycle in humidity and temperature-controlled room (temperature 22 ± 1°C and humidity 50 ± 10%).

To establish timed pregnancy for studying embryonic (E) development, the female mouse estrous cycle was monitored by visual observations of the vaginal opening of each female mouse based on the criteria described by Champlin et al. (1973). Mice in the proestrus or estrous phase of the cycle were selected for mating. Animals were bred in the evening and vaginal post-coitum protein plug was checked in the next morning no more than 12 h later. The morning that a plug was found was designated as E0.5 gestational stage. The day of the animal birth was designated as postnatal (P) 0 stage.

Tissue isolation and protein extraction

Mice and rats were euthanized by carbon dioxide inhalation and decapitated with a guillotine. Dissection of tissue samples was done in ice-cold 1x phosphate-buffered saline solution. Each sample contained tissues pooled together from three different animals for biological diversity and sufficient protein extraction at early developmental stages. The mouse and rat cerebral cortex, hippocampus, cerebellum, olfactory bulb, hypothalamus, and pons including medulla, midbrain, and thalamus were collected. Striatum was collected only for the BALB/c mouse strain. Tissue collection for mouse and rat brain regions occurred at P0, 3, 5, 7, 10, 14, 21, 60, and P0, 3, 5, 10, 14, 30, 60, respectively. In addition to postnatal days, collection of total mouse brain samples occurred at E13, 15, and 18. Mouse and rat peripheral tissues, skin, lung, kidney, heart, diaphragm, muscle, bladder, stomach, pancreas, thymus, spleen, liver, and blood, were collected at developmental stages P0, 14, and 60. After collection, tissue samples were stored at −80°C until further processing.

Tissues were homogenized on ice with tissue grinder PELLET PESTLE® Cordless Motor (Kimble-Chase, DWK Life Sciences) in ice-cold Radioimmunoprecipitation assay buffer [RIPA, 50 mM Tris pH 8.0, 150 mM NaCl, 1% NP-40, 0.5% Na-deoxycholate, 0.5% sodium dodecyl sulfate (SDS), and 1x Roche Protease Inhibitor Cocktail Complete]. Lysates were sonicated for 15 s with Torbeo Ultrasonic probe sonicator (36810-series, Cole Parmer), and insoluble material was removed by centrifugation at 4°C for 20 min at 16,000 *g*. Protein concentration was measured with Pierce BCA Protein Assay Kit (Thermo Scientific).

Protein lysates from human post-mortem cerebral cortex and hippocampus were prepared like rodent lysates. All protocols using human tissue samples were approved by Tallinn Committee for Medical Studies, National Institute for Health Development (Permit Number 402). All experiments were performed in accordance with relevant guidelines and regulations.

In vitro protein translation

TCF4-A$^+$, -A$^-$, -B$^+$, B$^-$, -C$^-$, -D$^-$ and TCF4-I$^-$isoforms were translated *in vitro* using pcDNA3 plasmids encoding the respective TCF4 isoforms (Sepp et al., 2011) with TnT Quick Coupled Transcription/Translation System (Promega). Equal volumes of *in vitro* translated TCF4 protein mixtures were used for western blot analysis.

Western blot analysis

For western blot analysis, protein lysates in RIPA were diluted to the same concentration in 1x Laemmli buffer. 55 µg of each sample was electrophoretically separated by SDS-polyacrylamide gel electrophoresis in 8% gel and transferred to polyvinylidene difluoride membrane (Millipore) in Towbin buffer (25 mM Tris, 192 mM glycine, 20% methanol, 0.1% SDS, and pH 8.3) using wet transfer. Membranes were blocked with 5% skimmed milk (Sigma-Aldrich) in 1x Tris Buffered Saline with 0.1% Tween-20 (TBST, Sigma Aldrich) before incubating with primary [mouse monoclonal anti-ITF-2 (TCF4); C-8, 1:1,000 dilution, Santa Cruz] and secondary (goat polyclonal anti-mouse IgG HRP-conjugated antibody 1:5,000 dilution, Thermo Scientific) antibody in 2.5% skimmed milk-TBST solution overnight at 4°C and 1 h at room temperature, respectively. Specificity of the anti-ITF-2 (TCF4; C-8, Santa Cruz) antibody has been previously validated using tissue lysates from *Tcf4* knockout mice (Nurm et al., 2021). For peripheral tissues, mouse IgG kappa binding protein conjugated to HRP was used as a secondary antibody (1:5,000 dilution, Santa Cruz). Chemiluminescence signal detection with SuperSignal West Femto or Atto Chemiluminescence Substrate (Thermo Scientific). The chemiluminescence signal was visualized with ImageQuant LAS 4000 bioimager (GE Healthcare) and densitometric quantification was performed with ImageQuant TL

8.2 image analysis software (GE Healthcare). Membrane staining with Coomassie solution (0.1% Coomassie Brilliant Blue R-250, 25% ethanol, and 7% acetic acid) was used as a loading control and for total protein normalization.

Analysis of publicly available RNA-seq datasets

Raw RNA-seq datasets of human, mouse, and rat were obtained from EMBL-EBI European Nucleotide Archive database using www.sra-explorer.info (Keane et al., 2011; ENCODE Project Consortium, 2012; Schmitt et al., 2014; Yu et al., 2014; Vied et al., 2016; Li et al., 2017; Söllner et al., 2017; Cardoso-Moreira et al., 2019; Luo et al., 2020; Shafik et al., 2021; see Supplementary Table S4 for accession numbers and sample information). Adapter and quality trimming were done using BBDuk (part of BBMap version 38.90) with the following parameters: ktrim = r k = 23 mink = 11 hdist = 1 tbo qtrim = lr trimq = 10 maq = 10 minlen = 25. Mouse reads were mapped to mm10 (primary assembly and annotation obtained from GENCODE, release M25, GRCm38), rat reads were mapped to rn6 (primary assembly and annotation obtained from Ensembl, release 104, RGSC 6.0/Rnor_6.0), and human reads were mapped to hg19 (primary assembly and annotation obtained from GENCODE, release 37, GRCh37) using STAR aligner (version 2.7.4a) with default parameters. To increase sensitivity for unannotated splice junctions, splice junctions obtained from the first pass were combined per dataset and filtered as follows: junctions on mitochondrial DNA and non-canonical intron motifs were removed; only junctions detected in at least 10% of samples (rounded up to the nearest integer) in the whole dataset were kept. Filtered junctions were added to the second pass mapping using STAR. Intron spanning reads were quantified using FeatureCounts (version 2.0.1). The following parameters were used for paired-end data: -p –B –C –J; and single-end data: -J. To count reads from *TCF4* extended exons (exons 4c and 7bII), reads crossing a 1 bp region 2 bp 5′ from the internal exon 4 and 7, respectively, were quantified using FeatureCounts and a custom-made saf file. Splice junctions in the *TCF4* locus were manually curated and annotated to TCF4 isoforms according to Sepp et al. (2011).

A custom R script[4] was used to quantify the expression of different *TCF4* transcripts from analyzed RNA-seq datasets. Briefly, RNA-seq reads crossing the indicated *TCF4* splice-junctions were normalized using all splice-junction crossing reads in the respective sample. Then, the data were summarized by the Exon column (Supplementary Tables S5–S7). To acquire total *TCF4* mRNA levels, the mean value of exon-junctions from 10–11 to 19–20 was taken for the analysis. Aggregated mouse and rat data was meta-analyzed in tandem with human data to study *TCF4* expression during development. The expression of *TCF4* transcripts encoding specific isoforms was assessed by quantifying the number of reads crossing the exon-junctions specific for the TCF4 isoforms. Data were summarized in each tissue and age group by the Isoform column. Values for each isoform were then divided with the sum of all annotated isoforms to show isoform composition in percentages. The results were visualized using ggplot2 (version 3.3.5) in R (version 4.1.2). The *TCF4* exon-junction data used for the analysis of *TCF4* transcripts in mouse, rat, and human datasets can be found in Supplementary Tables S5–S7, respectively.

Data mining and visualization was also performed on human GTEx portal exon-exon junction dataset (dbGaP Accession phs000424.v8.p2) and human developmental transcriptome data from BrainSpan (RNA-Seq Gencode v10 summarized to genes). The human GTEx data used for the analyses described in this manuscript were obtained from the GTEx Portal[5] on 12/01/2021 and the human BrainSpan data was obtained from the BrainSpan Atlas of the Developing Human Brain[6] on 12/01/2021. For information about rodent and human developmental stages and the number of individual data points per developmental stage, see Supplementary Tables S8–S12.

Results

Five N-terminally distinct TCF4 protein isoforms are expressed in the developing mouse cerebral cortex

The use of numerous alternative 5′ exons results in the expression of many transcripts from the *Tcf4* gene, resulting in a variety of TCF4 protein isoforms with different expression patterns between tissue types (Sepp et al., 2011; Nurm et al., 2021). We have previously described transcripts from the mouse *Tcf4* gene based on available mRNA and expressed sequence tag data from various tissues (Nurm et al., 2021) but characterizing all the isoform encoding transcripts using short-read sequencing is complicated. Here, we did long-read direct RNA-sequencing (RNA-seq) on the Oxford Nanopore Technologies platform from postnatal day 3 (P3) mouse cerebral cortex. Direct RNA-seq eliminates the bias which may result from complementary DNA synthesis used in conventional RNA-seq methods. Briefly, we extracted total RNA from the P3 mouse cerebral cortex which was then enriched for *Tcf4* transcripts using a combination of oligonucleotides complementary to the 3′ untranslated region and the basic helix–loop–helix region of *Tcf4* (Figure 1A; Supplementary Table S1). The Oxford Nanopore Technologies platform begins sequencing from the 3′ end of RNA meaning that early sequencing termination can result in reads which do not reach the 5′ exons of *Tcf4* transcripts. Analysis of our results

4 https://github.com/CSKiir/Sirp_et_al_2022

5 https://gtexportal.org

6 http://brainspan.org

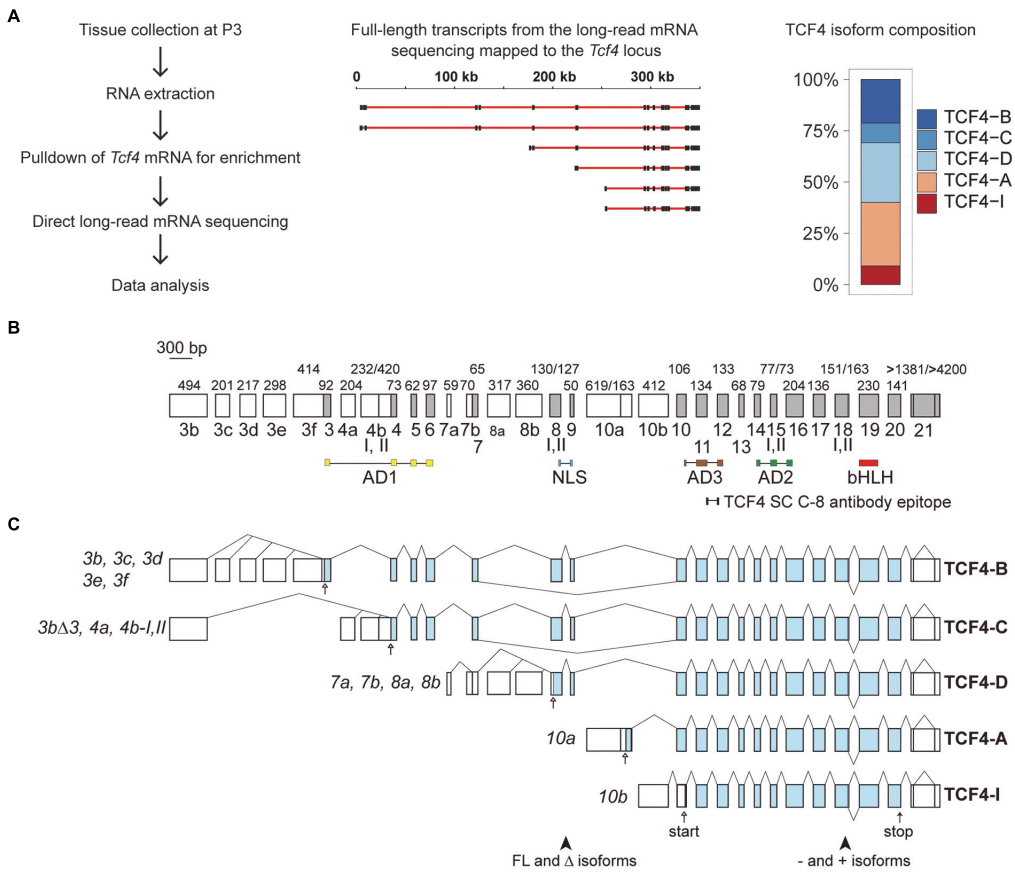


FIGURE 1
Tcf4 mRNA isoforms expressed in the developing mouse cerebral cortex encode five N-terminally distinct protein isoforms. **(A)** Schema of experimental design for direct RNA-sequencing. The cerebral cortices of postnatal day 3 (P3) mice were collected, followed by RNA extraction and *Tcf4* mRNA enrichment before direct long-read mRNA sequencing. A selection of *Tcf4* transcripts mapped to the mouse *Tcf4* locus is shown for reference. Black boxes represent exons and red lines show introns. Scale bar in kilobases is shown on top. *Tcf4* transcripts encoding the different TCF4 isoforms (TCF4-B, -C, -D, -A, and-I) were quantified and the distribution is shown on the right. Each isoform is represented with different color as shown in the legend on the right. **(B)** Mouse *Tcf4* genomic organization with exons drawn in scale. Exons are named according to the human *TCF4* gene (Sepp et al., 2011). 5′ exons are shown as white boxes while internal and 3′ exons are shown as gray boxes. Exon names are shown below boxes. Numbers above the exons designate the size of the exon in base pairs. Roman numerals below exons show alternative splice sites. Regions encoding different domains are marked below the gene structure (AD1, NLS, AD3, AD2, and bHLH) as well as the epitope for the TCF4 antibody C-8 (Santa Cruz, SC) used in the present study. **(C)** Schematic structure of *Tcf4* transcripts expressed in the developing mouse cerebral cortex. Untranslated regions are shown as white boxes and translated regions as blue boxes. Each transcript is named (shown on the left) according to the 5′ exon and with the number of the splice site where indicated. The names of the protein isoforms encoded by the transcripts are shown on the right. Positions of alternative splice region that generates full-length (FL), Δ, – and+isoforms are shown at the bottom. The position of the first in-frame start codon is shown with an arrow for each transcript and the common stop codon with an arrow at the bottom. AD, activation domain; NLS, nuclear localization signal; bHLH, basic helix–loop–helix; FL, full-length; and Δ, lack of exons 8–9.

showed that most of the 1,336 RNAs which mapped to the mouse *Tcf4* gene were short and mapped only to the last exon of *Tcf4* gene. However, we obtained 163 *Tcf4* transcripts that reached from the 3′untranslated region to the 5′ terminal exons and were thus considered full-length based on previous knowledge about the rodent *Tcf4* gene structure (Nurm et al., 2021). We then quantified the potential TCF4 protein isoforms encoded by these transcripts (Figure 1A). The results showed that in the mouse P3 cerebral cortex ~20% of mRNAs transcribed from *Tcf4* gene encode isoform TCF4-B; 10% encode isoform TCF4-C; 30% encode isoform TCF4-D; 30% encode isoform TCF4-A and 10% encode isoform TCF4-I (Figure 1A). The presence of plus (containing the RSRS amino acid sequence) and minus (without the RSRS amino acid sequence; Corneliussen et al., 1991; Nurm et al., 2021) TCF4

isoforms encoded by *Tcf4* mRNAs was roughly equal. Overall, our data show that the expression of mouse *Tcf4* gene leads to numerous transcripts due to the presence of 32 exons out of which 13 are alternative 5′ exons and 18 are internal exons and one is a terminal 3′ exon (Figures 1B,C). Using long-read sequencing, we discovered two novel 5′ exons (exon 3e and 3f) that are included in *Tcf4* transcripts encoding TCF4-B isoform (Figures 1B,C).

We then described the expression pattern of TCF4 protein isoforms in SDS-PAGE/western blot analysis. For that, we used *in vitro* translated TCF4-B, -C, -D, -A, and-I plus or minus TCF4 isoforms. This allowed us to compare different combinations of *in vitro* translated TCF4 isoforms in western blot to the TCF4 protein pattern in the mouse cerebral cortex. In good agreement with our RNA-seq experiment, the combination of *in vitro* translated proteins TCF4-B, -C, -D, -A, and -I resembled the protein pattern of TCF4 in the cerebral cortex (Figure 2A). The variability in *in vitro* translated TCF4 isoform levels in western blot (Figure 2A) could arise from differential translation rate of the respective TCF4 isoform encoding plasmids. Next, we determined the apparent molecular weight and location of endogenously expressed TCF4 isoforms—TCF4-B and TCF4-A in SDS-PAGE/western blot analysis. To this end, we constructed a CRISPR-Cas9 system to inhibit the expression of these TCF4 isoforms in Neuro2a cell line by generating frameshift mutations in the unique exons encoding these isoforms (exons 3 and 10a, respectively). We could not specifically silence the expression of TCF4-D since its translation start site is in internal exon 8—a frameshift mutation in exon 8 would cause the silencing of not only TCF4-D, but also TCF4-B and-C isoforms. We used Neuro2a cells, which show high endogenous expression of TCF4 as confirmed by RT-PCR and western blot analysis (Figure 2B; Supplementary Figure S1). By expressing the generated CRISPR-Cas9 system in Neuro2a cells, we were able to inhibit the expression of TCF4-B and-A and thus confirm the location of these protein isoforms in western blot analysis (Figure 2B). Furthermore, western blot analysis showed that expression pattern of TCF4 isoforms was similar in Neuro2a cells and P3 mouse cerebral cortex (Figure 2B).

Taken together, the main N-terminally distinct TCF4 isoforms expressed in the early postnatal mouse cerebral cortex are TCF4-B, -C, -D, -A, and -I. We classified the detected TCF4 signals into three groups based on their molecular weight: long isoforms (TCF4-B and-C), medium isoforms (TCF4-D), and short isoforms (TCF4-A and-I; Figure 2C).

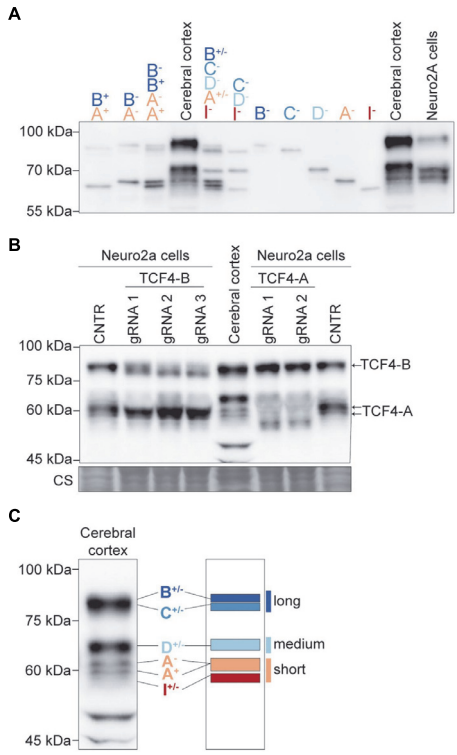


FIGURE 2
TCF4 protein isoforms expressed in the mouse cerebral cortex can be grouped into long, medium, and short isoforms based on apparent molecular weight in SDS-PAGE. **(A)** Western blot analysis of different combinations of *in vitro* translated TCF4 isoforms (shown on the top) to identify the mobility of TCF4 isoforms in the lysates from the mouse cerebral cortex and Neuro2a cells in SDS-PAGE. **(B)** Western blot analysis of Neuro2a cells transfected with CRISPR-Cas9 silencing system. Two exon 10a-specific gRNAs were used to silence TCF4-A and three exon 3-specific gRNAs were used to silence TCF4-B. Cells overexpressing CRISPR-Cas9 vector without the exon-specific gRNA targeting sequence was used as control (CNTR). Mouse cerebral cortex tissue lysate was used to compare TCF4 isoform expression pattern to the pattern in Neuro2a cells. Locations of TCF4-A and TCF4-B isoforms are depicted on the right. Coomassie staining (CS) was used a loading control and is shown at the bottom. **(C)** Schematic layout of the locations of TCF4 isoforms in the protein lysate of the mouse cerebral cortex in western blot. TCF4 isoforms were grouped into three—long, medium and short isoforms. The locations of TCF4 isoform groups are color coded and shown on the right. In each panel, molecular weight is shown on the left in kilodaltons.

Expression of TCF4 protein in the mouse brain is highest around birth

Next, we studied the changes in TCF4 protein expression in the mouse brain throughout pre-and postnatal development. For that, we made mouse whole brain lysates from two strains (BALB/C and C57BL/6) at 11 different developmental stages ranging from E13.5 to P60, and performed western blot analysis (Figures 3A,B). Both mouse strains exhibited expression of long, medium, and short TCF4 isoforms, with the highest expression of total TCF4 detected at late prenatal and early postnatal development. After peaking, TCF4 expression gradually declined

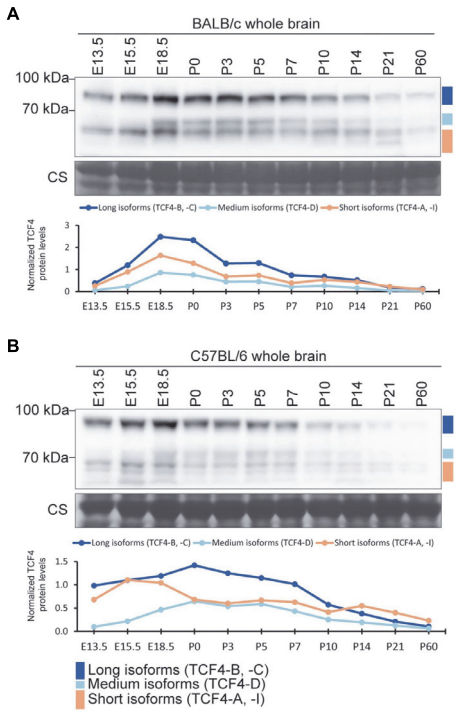


FIGURE 3
Protein expression of TCF4 in the mouse brain peaks around birth. **(A,B)** Western blot analysis of TCF4 protein expression through the pre-and postnatal development in the BALB/c **(A)** and C57BL/6 **(B)** mouse whole brain. Tissue lysates from the whole brain were made at different embryonic (E) and postnatal (P) days. The locations of TCF4 isoform groups are color coded and shown on the right. In each panel, molecular weight is shown on the left in kilodaltons. Under each western plot panel, a line graph depicting the quantification of long, medium, and short TCF4 protein isoforms during development is shown. TCF4 protein levels were normalized to the total TCF4 signal of the P10 cerebral cortex of the respective mouse strain.

during postnatal development of the brain (Figures 3A,B). While long and short TCF4 isoforms were detected at all stages, the medium-sized TCF4 isoforms became more apparent at later fetal stages and were almost undetectable before stage E18.5 (Figures 3A,B).

To better compare TCF4 total levels and isoform expression patterns between the two mouse strains, brain samples of the two strains from early postnatal development (P0–10) were analyzed in the same western blot experiment (Supplementary Figure S2). Our results revealed that the two mouse strains showed no major differences in TCF4 expression levels or expression patterns (Supplementary Figure S2). Altogether, these results indicated that TCF4 is expressed at both pre-and postnatal stages of the mouse brain development, with long and short TCF4 isoforms presented at all stages.

Expression of TCF4 is highest in the cerebral cortex, hippocampus, cerebellum, and olfactory bulb in the rodent brain

Next, we analyzed mRNA and protein expression of TCF4 in various rodent brain regions. First, we conducted a meta-analysis of available short-read RNA-seq data (Keane et al., 2011; ENCODE Project Consortium, 2012; Schmitt et al., 2014; Yu et al., 2014; Vied et al., 2016; Li et al., 2017; Söllner et al., 2017; Cardoso-Moreira et al., 2019; Luo et al., 2020; Shafik et al., 2021) to quantify the expression of total *Tcf4* mRNA and different *Tcf4* transcripts encoding distinct protein isoforms in rodents. Where possible, *Tcf4* expression dynamics was studied during different stages of pre-and postnatal development.

In the mouse brain, *Tcf4* mRNA expression was highest in the cerebral cortex, followed by the cerebellum, midbrain and hypothalamus (Figure 4A; Supplementary Figure S3A). A decrease in *Tcf4* mRNA expression after birth was seen for all the studied brain regions except for the cerebellum, which displayed relatively stable *Tcf4* mRNA levels during postnatal development. The majority of expressed *Tcf4* transcripts encoded isoforms TCF4-B, -C, -D, -A, and -I, with transcripts encoding TCF4-A showing the highest overall expression (Figure 4A; Supplementary Figure S4A) in the mouse brain. The cerebral cortex was the only brain region that displayed a notable change in the expression pattern of transcripts encoding different TCF4 isoforms—during development the expression of TCF4-A decreased and the expression of TCF4-D increased (Figure 4A).

We then sought to describe TCF4 expression at the protein level in mouse brain regions. For this, we dissected the cerebral cortex, hippocampus, cerebellum, striatum, pons, olfactory bulb, hypothalamus, thalamus, and midbrain at eight postnatal stages (P0, 3, 5, 7, 10, 14, 21, and 60) from BALB/C (Figure 4) and C57BL/6 (Supplementary Figure S5) mice, prepared protein lysates and analyzed TCF4 levels by western blot. First, we compared TCF4 protein expression across distinct brain regions at two postnatal stages, P0 and P10 by western blot analysis (Figures 4B,C; Supplementary Figures S5A,B). We observed high TCF4 expression levels in the cerebral cortex and hippocampus at P0, and in the cerebellum at P10 (Figures 4B,C; Supplementary Figures S5A,B). The long and short TCF4 protein isoforms were present in all studied brain regions (Figures 4B,C; Supplementary Figures S5A,B). The medium isoforms had more restricted patterns being detected at high levels in the cerebral cortex and hippocampus, at low levels in the cerebellum, olfactory bulb, and pons, and were below the detection limits in other brain regions (Figures 4B,C).

Next, we focused on the developmental dynamics of TCF4 protein expression in all the dissected mouse brain regions during postnatal development (P0-60; Figure 4D; Supplementary Figure S5C). To better compare TCF4 signals between individual brain regions across development, we used tissue lysate from the P10 cerebral cortex in each experiment as

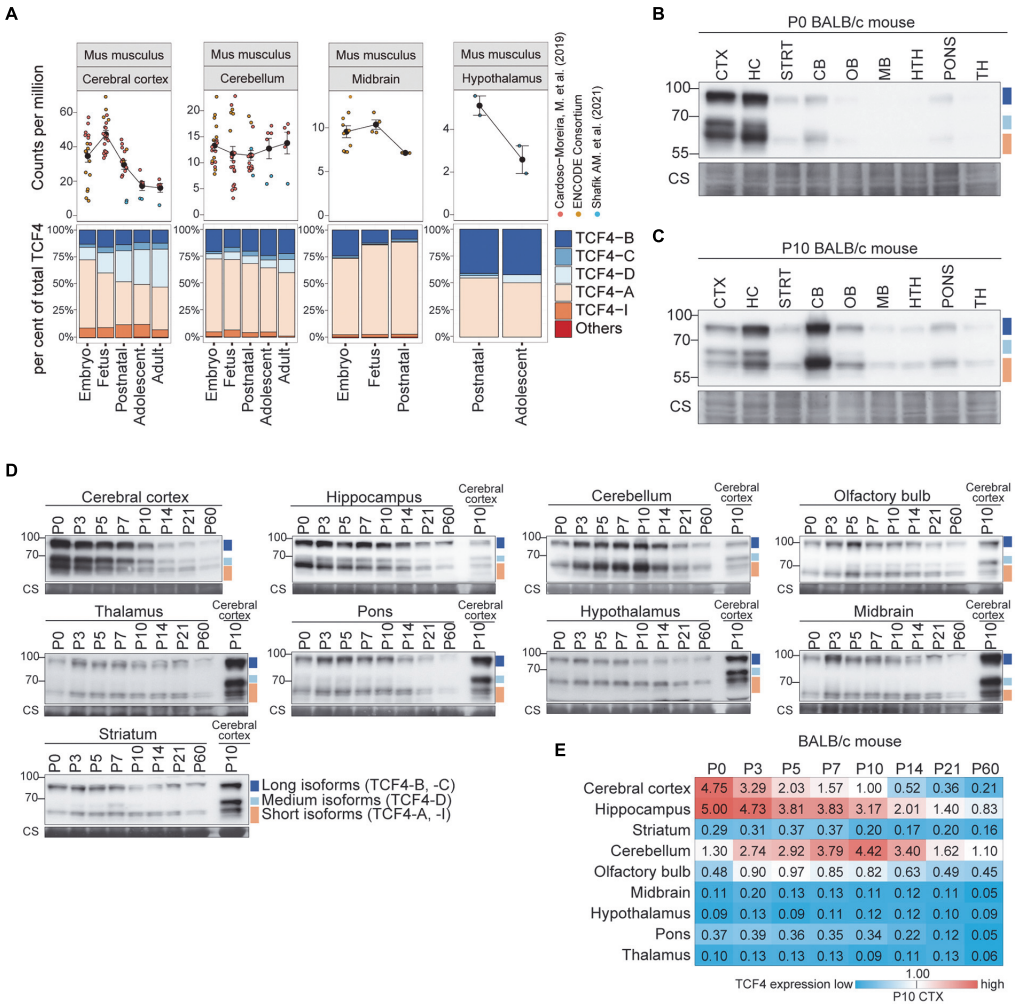


FIGURE 4
Expression of TCF4 is high in the mouse cerebral cortex, hippocampus and cerebellum. **(A)** Three independent datasets from Cardoso-Moreira et al. (2019), ENCODE Consortium (ENCODE Project Consortium, 2012; Luo et al., 2020), and Shafik et al. (2021) were combined for meta-analysis of *Tcf4* mRNA expression in mouse cerebral cortex, cerebellum, midbrain, and hypothalamus throughout development. mRNA expression of total *Tcf4* is visualized as a line chart (upper panel) where the solid line connects the mean of *Tcf4* expression for each developmental stage and error bars represent standard error of the mean (SEM). The distribution of isoform-specific transcripts is shown as bars (lower panel). Each isoform is represented with different color, as shown in the legend on the right. (**B–D**) Western blot analysis of TCF4 protein expression in different brain areas of BALB/c mouse at P0 **(B)**, P10 **(C)**, and throughout postnatal development **(D)**. The examined brain areas are shown on the top of each panel together with the day of postnatal development. P10 cerebral cortex was used for normalization **(D)**. Coomassie membrane staining (CS) shown at the bottom of each western blot was used as a loading control. The locations of TCF4 isoform groups are color coded and shown on the right. In each panel, molecular weight is shown on the left in kilodaltons. **(E)** TCF4 signals from western blot analysis of different brain areas of BALB/c mouse were quantified and normalized using Coomassie staining. The normalized signal from P10 cerebral cortex was set as 1, and the quantification results are visualized as a heatmap. Color scale gradient represents the relative TCF4 expression level, where blue and red color represents the lowest and the highest total TCF4 protein level, respectively. The studied brain regions are shown on the left and developmental stages on top. CTX, cerebral cortex; HC, hippocampus; CB, cerebellum; STRT, striatum; OB, olfactory bulb; MB, midbrain; HTH, hypothalamus; TH, thalamus; P, postnatal day; and CS, Coomassie staining.

a calibrator and quantified the results (Figure 4E; Supplementary Figure S5F). In agreement with our direct comparisons (Figures 4B,C; Supplementary Figures S5A,B), the highest levels of TCF4 expression were observed in the cerebral cortex, hippocampus, and cerebellum, with all the other studied brain regions showing either moderate (olfactory bulb) or low (striatum, midbrain, hypothalamus, pons, and thalamus) TCF4 expression (Figure 4E; Supplementary Figure S5F). While in the

cerebral cortex and hippocampus TCF4 expression peaked just after birth (Figure 4E; Supplementary Figure S5F), cerebellum displayed a slightly delayed increase in TCF4 expression, peaking around a week after birth (Figure 4E; Supplementary Figure S5F). Notably, cerebellum and hippocampus retained a higher expression of TCF4 for a longer period compared to the cerebral cortex (Figure 4E; Supplementary Figure S5F).

To extend our observations for the mouse, we next characterized TCF4 expression in the brain of another important model organism, the rat (Figure 5). *Tcf4* mRNA expression was highest in the rat cortex, hippocampus, and cerebellum (Figure 5A; Supplementary Figure S3B). At P0 and P10, the highest TCF4 protein expression levels were seen in the cerebral cortex, hippocampus, and cerebellum (Figures 5B,C). The

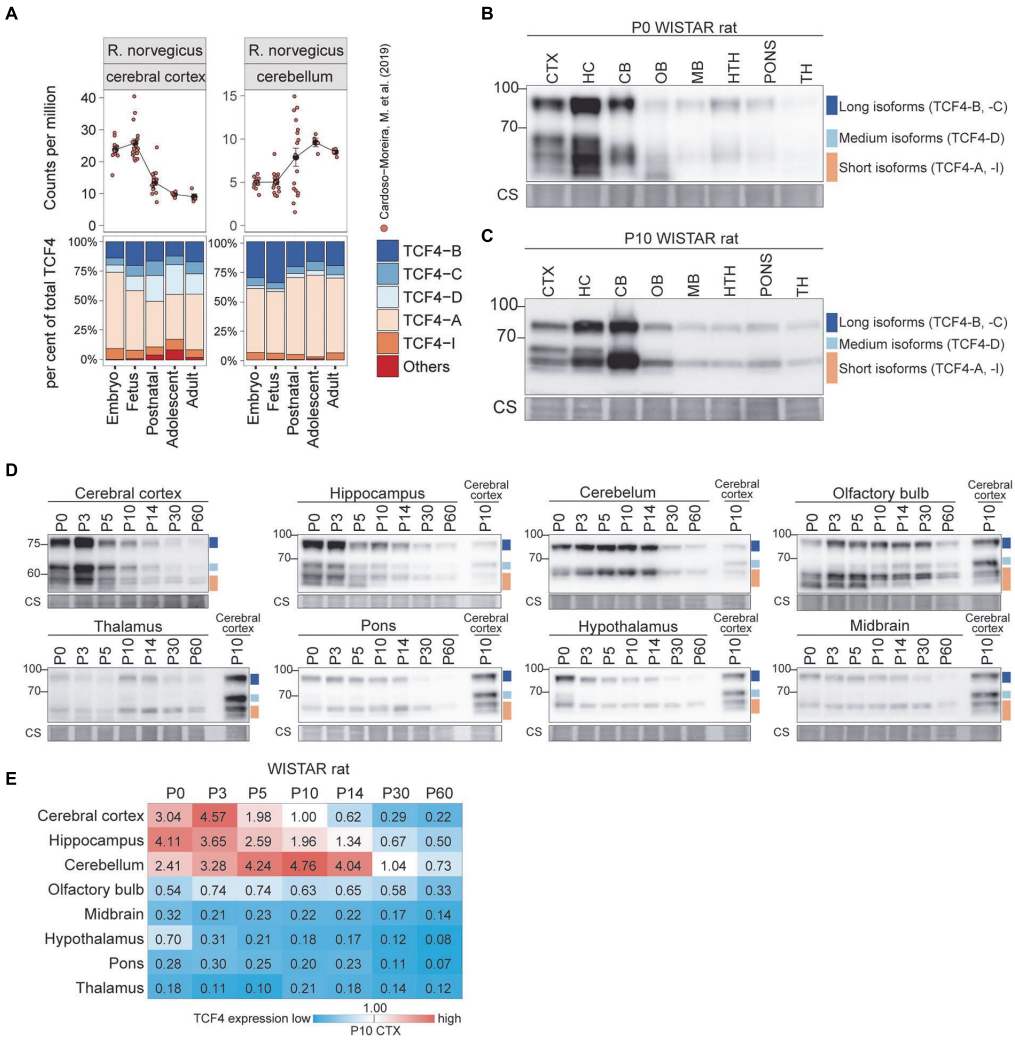


FIGURE 5
Expression of TCF4 is high in the rat cerebral cortex, hippocampus, and cerebellum. **(A)** Dataset from Cardoso-Moreira et al. (2019) was analyzed for TCF4 expression in rat cerebral cortex and cerebellum throughout development. Total *Tcf4* mRNA expression and the TCF4 isoform distribution was analyzed and visualized similarly to mouse. For more information see legend of Figure 4A. **(B–D)** Western blot analyses of TCF4 protein expression in different brain areas of WISTAR rat at P0 **(B)**, P10 **(C)**, and throughout postnatal development **(D)**. For more details see legend of Figures 4B–D. **(E)** TCF4 signals from western blot analysis of different brain areas of WISTAR rat were quantified, normalized using Coomassie staining, and visualized as a heatmap. For more details see legend of Figure 4E. CTX, cerebral cortex; HC, hippocampus; CB, cerebellum; OB, olfactory bulb; MB, midbrain; HTH, hypothalamus; TH, thalamus; P, postnatal day; and CS, Coomassie staining.

expression dynamics of TCF4 protein isoforms across postnatal development were similar to mouse (Figures 5D,E). Notably, in P0-5 rat cerebral cortex, hippocampus, and olfactory bulb an isoform migrating faster than the TCF4-A isoforms was seen, potentially corresponding to TCF4-I (Figures 5B–D).

Taken together, our data showed that TCF4 expression pattern and dynamics were similar in the mouse and rat brain—the highest TCF4 protein expression was seen in the cerebral cortex and hippocampus around birth, and in the cerebellum 1–2 weeks after birth. TCF4 expression in other brain regions was relatively low. In addition to the differences in overall expression of TCF4, our data revealed that the composition of TCF4 isoforms expressed varies across brain regions in mouse and rat.

Expression of TCF4 is lower in rodent nonneural organs compared to the brain

We then focused on mouse nonneural organs and performed a meta-analysis of available short-read RNA-seq data (Keane et al., 2011; ENCODE Project Consortium, 2012; Schmitt et al., 2014; Yu et al., 2014; Vied et al., 2016; Li et al., 2017; Söllner et al., 2017; Cardoso-Moreira et al., 2019; Luo et al., 2020; Shafik et al., 2021). We analyzed *Tcf4* mRNA expression in the lung, kidney, thymus, spleen, liver, heart, and stomach (Figure 6A; Supplementary Figures S3A, S4A). These tissues displayed comparable *Tcf4* mRNA expression levels except for the liver, where almost no *Tcf4* mRNA expression was seen after birth (Figure 6A; Supplementary Figure S3A). Of the transcripts encoding different TCF4 protein isoforms in nonneural organs, the ones encoding TCF4-A were most prominently expressed, followed by TCF4-B-encoding transcripts (Figure 6A; Supplementary Figure S4A).

Next, we prepared protein lysates from BALB/c (Figure 6) and C57BL/6 (Supplementary Figure S5) mouse heart, diaphragm, muscle, skin, lung, kidney, bladder, stomach, pancreas, thymus, spleen, liver, and blood cells at P0, 14, and 60 for western blot analysis (Figures 6B,C, Supplementary Figure S5). In nonneural tissues, the composition of TCF4 protein isoforms was similar to the one in the brain—both long and short TCF4 protein isoforms were always present, whereas medium-sized TCF4 isoforms were not observed in any of the nonneural tissues (Figures 6B,C; Supplementary Figure S5D,E). Among the studied nonneural tissues, the highest levels of TCF4 protein were seen in the skin at P0 (Figure 6D; Supplementary Figure S5G). Very low TCF4 protein levels were detected in the pancreas, spleen, kidney and liver, and TCF4 protein expression was not seen in the blood cells (Figure 6D; Supplementary Figure S5G).

We also investigated TCF4 expression in rat nonneural tissues (Figure 7; Supplementary Figures S3B, S4B). Rat *Tcf4* mRNA expression was comparable in all the nonneural tissues except for the liver, where *Tcf4* expression was very low (Figure 7A; Supplementary Figure S3B). Different from mouse, transcripts encoding TCF4-A did not account for the majority of rat *Tcf4* transcripts expressed in nonneural tissues, as also high expression of transcripts encoding TCF4-B and TCF4-C were present (Figure 7A; Supplementary Figure S4B).

Western blot analysis of rat nonneural tissues showed that different from mouse, TCF4 protein expression levels were more uniform between tissues (Figures 7B–D). In rat, TCF4 protein expression was highest in the thymus and was not observed in the pancreas (Figures 7B–D). The expression pattern of TCF4 isoforms in rat nonneural tissues was similar to mouse, i.e., mainly long and short TCF4 isoforms being present (Figures 7B,C).

Overall, the expression of TCF4 in the rodent nonneural tissues was much lower compared to the expression levels observed in the early postnatal development of the central nervous system. In addition, medium-sized TCF4 protein isoforms were almost non-existent in rodent nonneural tissues.

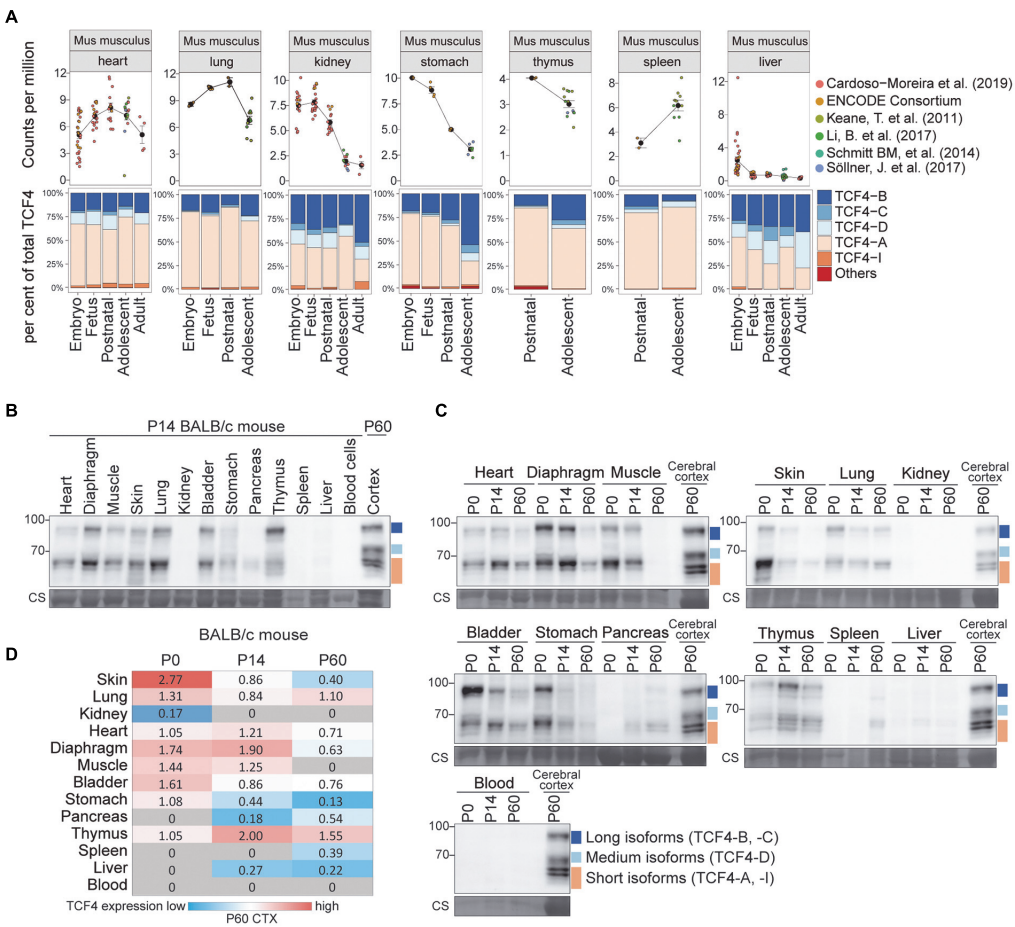
Expression of TCF4 in human tissues is highest around birth

Next, we analyzed available short-read RNA-seq data to describe *TCF4* total and isoform-specific mRNA expression in humans. We first analyzed the dataset published by Cardoso-Moreira and colleagues, which contained RNA-seq data from the human brain, heart, kidney, liver, and testis (Cardoso-Moreira et al., 2019). Of the noted tissues, the highest TCF4 mRNA expression was detected in the brain (Figure 8A). Human nonneural tissues showed detectable but lower *TCF4* mRNA levels compared to the brain, especially in the earlier stages of development (Figure 8A). Very low *TCF4* mRNA expression was noted for the liver (Figure 8A). In contrast to other tissues where *TCF4* mRNA levels were relatively stable throughout development, *TCF4* mRNA expression in the forebrain and kidney peaked during prenatal development (Figure 8A). We then used developmental transcriptome data from the BrainSpan project[7] to describe the changes in total *TCF4* mRNA expression in different brain regions during human development (Supplementary Figure S6). Results were similar in all brain regions—*TCF4* mRNA expression peaked during embryonic development and decreased after birth (Supplementary Figure S6).

Next, we conducted a similar analysis for adult human RNA-seq data from the Genotype-Tissue Expression (GTEx) project.[8] A selection of adult human tissues is shown in Figure 8B and all the studied tissues can be found in Supplementary Figure S7. When comparing different adult human tissues, the highest *TCF4* mRNA expression levels were seen in the adult human brain and adipose tissues (Figure 8B; Supplementary Figure S7). Almost no *TCF4* mRNA was detected in the human pancreas, liver, and whole blood (Figure 8B; Supplementary Figure S7).

7 http://brainspan.org

8 https://gtexportal.org



	P0	P14	P60
Skin	2.77	0.86	0.40
Lung	1.31	0.84	1.10
Kidney	0.17	0	0
Heart	1.05	1.21	0.71
Diaphragm	1.74	1.90	0.63
Muscle	1.44	1.25	0
Bladder	1.61	0.86	0.76
Stomach	1.08	0.44	0.13
Pancreas	0	0.18	0.54
Thymus	1.05	2.00	1.55
Spleen	0	0	0.39
Liver	0	0.27	0.22
Blood	0	0	0

FIGURE 6
Expression of TCF4 in mouse nonneural tissues. **(A)** Six independent datasets shown on the right were combined for meta-analysis *of Tcf4* expression in mouse heart, lung, kidney, stomach, thymus, spleen, and liver throughout development. Total *Tcf4* levels and the distribution of isoform-specific transcripts is visualized. For more information see legend of Figure 4A. **(B,C)** Western blot analysis of TCF4 protein expression in different nonneural tissues of BALB/c mouse at P14 **(B)**, and throughout postnatal development **(C)**. TCF4 signal from the P60 cerebral cortex was used in each experiment for normalization. Coomassie membrane staining (CS) shown at the bottom of each western blot was used as a loading control. The locations of TCF4 isoform groups are color coded and shown on the right. In each panel, molecular weight is shown on the left in kilodaltons. **(D)** TCF4 signals from western blot analysis of different peripheral tissues of BALB/c mouse were quantified and normalized using Coomassie staining. The signal was then normalized to the signal of the P60 cerebral cortex, and the quantification result is visualized as a heatmap. Color scale gradient represents the relative TCF4 expression level, where blue and red color represents the lowest and the highest total TCF4 protein level, respectively. Gray boxes indicate no detectable TCF4 expression. The studied nonneural tissues are shown on the left and developmental stages on top. P, postnatal day; CS, Coomassie staining.

For both Cardoso-Moreira et al. and GTEx datasets, transcripts encoding TCF4-A made up around 50% of the total TCF4 mRNA levels in all the studied tissues, with the only exception being the testis (Figures 8A,B; Supplementary Figure S7). In the human testis, mRNA transcripts encoding TCF4-J accounted for the majority of total *TCF4* levels beginning from adolescence, which coincides with the start of spermatogenesis (Figure 8A; Supplementary Figure S6). The other major isoform-specific transcripts expressed in human tissues were TCF4-B, -C, and-D (Figures 8A,B; Supplementary Figure S6).

Next, we aimed to investigate TCF4 isoform composition in the adult human cerebral cortex and hippocampus. For this, we prepared protein lysates from these brain regions and human neuroblastoma cell line SH-SY5Y used for isoforms' mobility comparison. Western blot analysis revealed that TCF4 protein

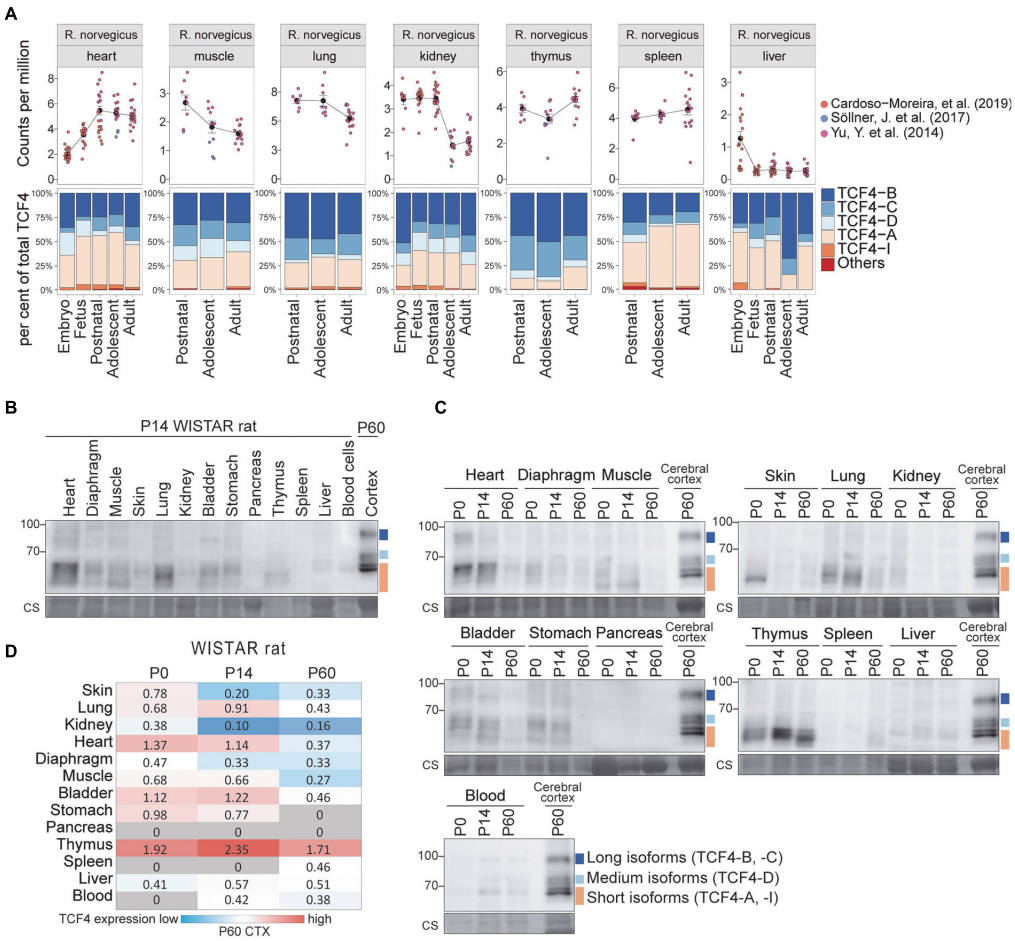


FIGURE 7
Expression of TCF4 in rat nonneural tissues. **(A)** Three independent datasets shown on the right were combined for meta-analysis of *Tcf4* expression in rat heart, muscle, lung, kidney, thymus, spleen, and liver throughout development. Total *Tcf4* levels and the distribution of isoform-specific transcripts is visualized. For more information see Figure 4A. (**B,C**) Western blot analyses of TCF4 protein expression in different peripheral tissues of WISTAR rat at P14 **(B)**, and throughout postnatal development **(C)**. For more details see legend of Figures 6B,C. **(D)** TCF4 signals from western blot analysis of different nonneural tissues of WISTAR rat were quantified, normalized using Coomassie staining, and visualized as a heatmap. For more details see legend of Figure 6D. P, postnatal day; CTX, cerebral cortex.

signal was detectable in both adult human brain and SH-SY5Y cell line (Figure 9). We also detected a possible non-specific signal located between the long and medium TCF4 isoforms in both SH-SY5Y and human brain lysates (Figure 9) since this signal was not detected using other TCF4 antibodies (data not shown) validated by us before (Nurm et al., 2021). Different to SH-SY5Y cell line, we detected all three TCF4 isoform groups in the adult human brain, however expression level of longer TCF4 isoforms was higher compared to the medium and short isoforms (Figure 9; Supplementary Figure S8). This result was on the contrary with protein isoform patterns seen in the rodent brain and the results from our human RNA-seq data analysis, which could result from protein stability, post-mortem artifacts or signal masking by other similar-sized proteins. Nevertheless, the presence of long, medium and short TCF4 isoforms in adult human brain matched with TCF4 isoform pattern in rodents, however the species-and tissue-specific temporal expression dynamics of different TCF4 isoforms during the development cannot be emphasized more.

Altogether, our results show that *TCF4* mRNA is expressed at high levels in the human brain during development and the expression is retained in the adulthood. In most tissues transcripts

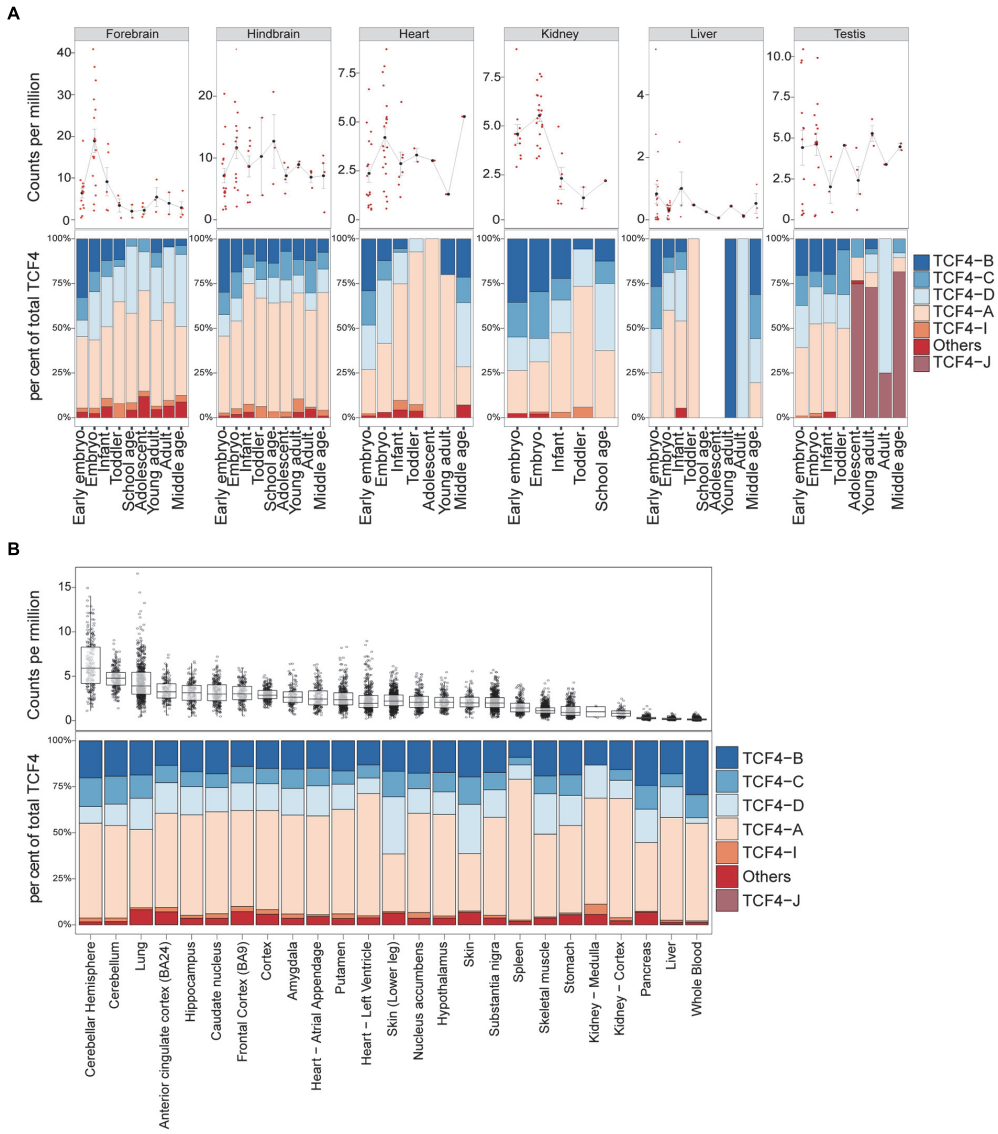


FIGURE 8
TCF4 transcripts encoding TCF4-A account for the majority of total *TCF4* expression in the human brain and nonneural tissues. Data from Cardoso-Moreira et al. **(A)** (Cardoso-Moreira et al., 2019) and the Genotype-Tissue Expression (GTEx) project **(B)** was analyzed for *TCF4* expression in the brain and nonneural tissues in humans through development **(A)** or in adults **(B)**. mRNA expression of total *TCF4* is visualized either as a line chart **(A)** or box plot **(B)**, and the distribution of isoform-specific transcripts is shown as bars **(A,B)**. **(A)** Average values are presented as dots and error bars represent SEM. **(B)** The hinges show 25 and 75% quartiles, the horizontal line shows the median value, the upper whisker extends from the hinge to the largest value no further than 1.5 of the inter-quartile range from the hinge, the lower whisker extends from the hinge to the smallest value at most 1.5 * inter-quartile range of the hinge. Each isoform is represented with different color, as shown in the legend on the right. Individual data points are presented as small dots.

encoding for TCF4-A were the most prominent ones, while in the testis TCF4-J encoding transcripts were mostly expressed. In adult human brain, long, medium and short TCF4 protein isoforms are expressed.

Discussion

Transcription factor TCF4 has been extensively studied due to its linkage with neurocognitive disorders such as intellectual

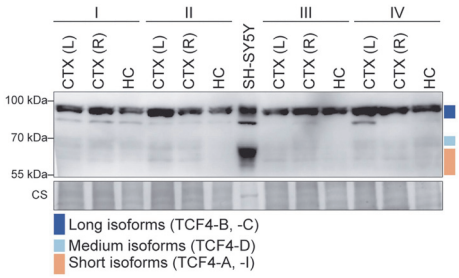


FIGURE 9
Expression of TCF4 protein isoforms in adult human cerebral cortex and hippocampus. Western blot analysis of TCF4 protein expression in the adult human cerebral cortex left (L) and right (R) hemisphere, hippocampus, and SH-SY5Y cell line. Samples from four individuals aged 62 (I), 65 (II), 67 (III), and 70 (IV) yearswere used. Coomassie staining (CS) shown at the bottom was used as loading control. The locations of TCF4 isoform groups are color coded and shown on the right. In each panel, molecular weight is shown on the left in kilodaltons. CS, coomassie staining.

disability, schizophrenia and Pitt-Hopkins syndrome (Stefansson et al., 2009; Kharbanda et al., 2016; Zollino et al., 2019). Knowledge of TCF4 expression across tissues and development would lay the foundation to understanding how these diseases develop and may help with the generation of gene therapy applications for the many TCF4 associated diseases. Transcripts from the mouse and human *TCF4* gene have been previously described in our lab using mRNA and expressed sequence tag data from public databases (Sepp et al., 2011; Nurm et al., 2021). Short read RNA-seq data can also be used to describe *Tcf4* transcripts. However, due to the structure of the *Tcf4* gene, it can be complicated to describe expression of transcripts encoding different isoforms based only on short read RNA-seq data as splicing features or 5′ exons can be difficult to detect. Our direct long-read RNA-seq analysis of *Tcf4* transcripts in the rodent brain revealed that transcription from the *Tcf4* gene results in transcripts encoding 5 N-terminally distinct TCF4 protein isoforms in the rodent brain – TCF4-B, -C, -D, -A, and-I. This result falls in line with previous observations by Nurm et al. (2021).

Expression of total *Tcf4* mRNA during development has been extensively studied mainly in the mouse cerebral cortex at the total mRNA level, with the highest expression reported around birth (E16-P6; Li et al., 2019; Phan et al., 2020). This is in accordance with our RNA-seq meta-analysis and applies for both mouse and rat. In addition, we show that the expression dynamics of *Tcf4* in the rodent brain and nonneural tissues are similar—highest *Tcf4* expression can be detected around birth, followed by a decline during postnatal development. However, studying only total *Tcf4* mRNA levels provides only partial information about *Tcf4* expression as transcription from the *Tcf4* gene results in numerous transcripts, and encoded protein isoforms have different functional protein domains and transactivation capability (Sepp et al., 2011, 2017; Nurm et al., 2021). We have previously developed a method which quantifies the expression of different *Tcf4* protein isoform-encoding transcripts using short read RNA-seq data (Sirp et al., 2020). Here, we applied the same approach to describe the expression of different TCF4 isoforms throughout development using previously published RNA-seq data. When leaving aside the great increase in the expression of TCF4-J in the adolescent human testis (transcripts encoding TCF4-J are not present in rodents), no drastic change concerning switching from the expression of one TCF4 isoform to the other was detected during the rodent and human development. This suggests that the same TCF4 isoforms that are necessary during development may also be vital for the TCF4-mediated normal functioning of the adult organism.

The necessity of so many different TCF4 isoforms remains unknown. In humans, mutations in the 5′ region of *TCF4* gene, which affect only the longer isoforms, lead to mild–moderate intellectual disability (Kharbanda et al., 2016). As the resulting disease is not as severe as the Pitt-Hopkins syndrome, it may mean that a slight decrease in overall TCF4 expression causes the phenotype. However, it is also possible that the longer TCF4 isoforms have specific functions which cannot be compensated by other TCF4 isoforms, and mutations affecting only a subset of TCF4 isoforms result in less severe effects than seen for mutations affecting all the isoforms. Recently, it has been shown that postnatal normalization of TCF4 expression to wild type levels can rescue the phenotype of TCF4 heterozygous knockout mice (Kim et al., 2022). In addition, studies of Daughterless, the orthologue of TCF4 in the fruit fly, have shown that it is possible to partially rescue the severe embryonic neuronal phenotype of Daughterless null mutation by overexpressing either human TCF4-A or TCF4-B (Tamberg et al., 2015). The generation of TCF4 isoform-specific mutant mice would help to identify whether TCF4 isoforms have distinct or similar functions. Such a model could be used to determine whether it is possible to rescue the negative phenotype resulting from a knock-out of a single TCF4 isoform by increasing the level of an another TCF4 isoform. However, generating such a model comes with many challenges. To begin with, it can be complicated to silence all the TCF4 isoforms individually by causing just frameshift mutations as only some isoforms (e.g., TCF4-B and-A) have their translation start sites located in independent 5′ exons. In addition, mutating one *TCF4* transcript can result in the upregulation of another *TCF4* transcript – an effect that we saw when silencing TCF4-A in Neuro2a cells that resulted in an increase in TCF4-I levels. We have also previously shown that Fuchs' Endothelial Dystrophy-related endogenous downregulation of transcripts encoding longer TCF4 isoforms results in the upregulation of shorter isoforms (Sirp et al., 2020).

To fully characterize TCF4 expression, it is important to consider all TCF4 protein isoforms. Previously, a large study on the expression of TCF4 protein during neurodevelopment has been performed by Matthias Jung and colleagues using immunohistochemical analysis with an antibody specific only for the long TCF4 protein isoforms (Jung et al., 2018). Another study

by Kim and colleagues used TCF4-GFP reporter mice to characterize total TCF4 expression in the mouse brain (Kim et al., 2020). A major limitation of these methods is that they cannot be used to describe the expression of different TCF4 protein isoforms. Overall, our results of total TCF4 protein expression levels during postnatal development of different mouse brain areas agree with the previously reported data. However, by using a TCF4 antibody specific for all the TCF4 isoforms in western blot analysis, we were able to distinguish TCF4 protein expression in three different groups – long (TCF4-B, TCF4-C), medium (TCF4-D) and short TCF4 isoforms (TCF4-A, TCF4-I). Isoform-specific silencing of TCF4 and *in vitro* translated TCF4 protein isoforms confirmed the locations of TCF4 isoforms in western blot. However, it should be noted that a similar pattern of TCF4 isoforms in western blot analysis between different tissues may not necessarily indicate the presence of exactly the same TCF4 isoforms. We acknowledge that *in vitro* and *in vivo* translated proteins can migrate differently in western blot analysis due to the differences in post-translational modifications of the proteins in various cell types.

The expression dynamics of TCF4 during the development varied in different brain regions. In contrast to the cerebral cortex where TCF4 expression levels decline after birth, in the cerebellum, hippocampus and olfactory bulb we saw a more prolonged high TCF4 protein expression. In the cerebellum TCF4 protein expression peaks about a week later than in any of the other brain regions. While the majority of the neurogenesis in the central nervous system happens during prenatal development, the granule cell precursors of the cerebellum and olfactory bulb, and the dentate gyrus of the hippocampus proliferate and differentiate after birth (Chen et al., 2017), where TCF4 was shown to be highly expressed (Jung et al., 2018; Kim et al., 2020), and regulate the maturation of the cerebellar granule cells (Kim et al., 2020). We propose that high TCF4 expression is necessary for the maturation of distinct brain regions, whereas fully developed brain areas display low and stable TCF4 expression necessary for normal function of the adult nervous system.

Expression of long and short TCF4 protein isoforms was seen in all brain regions and nonneural tissues where TCF4 was detectable. However, in rodents the medium TCF4 isoforms (TCF4-D) were only observed in the brain, specifically in the cerebral cortex, hippocampus, and olfactory bulb. Interestingly, in the whole rodent brain the expression of medium isoforms became apparent only in later stages of embryonic development. The only well-known functional protein domain located in the N-terminal region of longer TCF4 isoforms is activation domain 1. While the long TCF4 isoforms (TCF4-B and-C) contain this domain, TCF4-D lacks it. In addition, different from short TCF4 isoforms (TCF4-A and-I), TCF4-D contains a nuclear localisation signal. It remains to be studied what the function of TCF4-D in the development of the nervous system is and why this TCF4 isoform is missing in the cerebellum where TCF4 is otherwise highly expressed.

Based on the results of the present study we propose that a mixture consisting of TCF4-B, -C, -D, -A, and-I encoding constructs could be used in gene therapy approaches for Pitt-Hopkins syndrome. However, it should be noted that TCF4 expression levels vary between brain regions and cell types during development (Jung et al., 2018; Kim et al., 2020), suggesting that the dosage of TCF4 isoforms needs to be highly regulated. The direct administration of a cocktail of TCF4 isoforms may allow easier control of each isoform compared to other gene therapy approaches such as activation of endogenous promoters and enhancers. As a next step of this study, a similar TCF4 protein expression analysis should be done for human brain regions with a focus on the hippocampus and cerebral cortex, as studies of structural brain anomalies in PTHS-patients and *Tcf4*-heterozygous mice have shown hypoplasia of these brain regions (Marangi and Zollino, 2015). In addition, the expression of TCF4 different transcripts and the protein isoforms they encode should be studied at the single cell level to better understand how the many TCF4 isoforms are regulated between cell types.

Data availability statement

The datasets analyzed and presented in this study can be found in online repositories and in the Supplementary material. The names of the repository/repositories and accession number(s) can be found in the article.

Ethics statement

The studies involving human participants were reviewed and approved by Tallinn Committee for Medical Studies, National Institute for Health Development (Permit Number 402). The patients/participants provided their written informed consent to participate in this study. The animal study was reviewed and approved by Ministry of Agriculture of Estonia (Permit Number: 45).

Author contributions

ASi and JT designed research, performed research, analyzed data, and wrote the paper. ASh, LT, and CK performed research, analyzed data, and wrote the paper. LK performed research. TT designed research, wrote the paper, and acquired funding. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by Estonian Research Council (grant PRG805), European Union through the European Regional Development Fund (project no. 2014-2020.4.01.15-0012) and H2020-MSCA-RISE-2016 (EU734791), and Pitt Hopkins Research Foundation (grant no. 21). The funding sources were not

involved in study design, analysis and interpretation of data, writing of the report and in the decision to submit the article for publication.

Acknowledgments

We thank Epp Väli for technical assistance, Enn Jõeste for the dissection of human tissue samples and Mari Sepp for critical reading of the manuscript. We would like to acknowledge the ENCODE Consortium and the ENCODE production laboratories for generating the datasets used in this study. We thank the "TUT Institutional Development Program for 2016–2022" Graduate School in Clinical Medicine, which received funding from the European Regional Development Fund under program ASTRA 2014-2020.4.01.16-0032 in Estonia.

Conflict of interest

JT and TT were employed by Protobios LLC.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnmol.2022.1033224/full#supplementary-material

References

Amiel, J., Rio, M., de Pontual, L., Redon, R., Malan, V., Boddaert, N., et al. (2007). Mutations in TCF4, encoding a class I basic helix-loop-helix transcription factor, are responsible for Pitt-Hopkins syndrome, a severe epileptic encephalopathy associated with autonomic dysfunction. *Am. J. Hum. Genet.* 80, 988–993. doi: 10.1086/515582

Badowska, D. M., Brzózka, M. M., Kannaiyan, N., Thomas, C., Dibaj, P., Chowdhury, A., et al. (2020). Modulation of cognition and neuronal plasticity in gain-and loss-of-function mouse models of the schizophrenia risk gene Tcf4. *Transl. Psychiatry* 10:343. doi: 10.1038/s41398-020-01026-7

Brockschmidt, A., Todt, U., Ryu, S., Hoischen, A., Landwehr, C., Birnbaum, S., et al. (2007). Severe mental retardation with breathing abnormalities (Pitt–Hopkins syndrome) is caused by haploinsufficiency of the neuronal bHLH transcription factor TCF4. *Hum. Mol. Genet.* 16, 1488–1494. doi: 10.1093/hmg/ddm099

Brzózka, M. M., Radyushkin, K., Wichert, S. P., Ehrenreich, H., and Rossner, M. J. (2010). Cognitive and sensorimotor gating impairments in transgenic mice overexpressing the schizophrenia susceptibility gene Tcf4 in the brain. *Biol. Psychiatry* 68, 33–40. doi: 10.1016/j.biopsych.2010.03.015

Cardoso-Moreira, M., Halbert, J., Valloton, D., Velten, B., Chen, C., Shao, Y., et al. (2019). Gene expression across mammalian organ development. *Nature* 571, 505–509. doi: 10.1038/s41586-019-1338-5

Champlin, A. K., Dorr, D. L., and Gates, A. H. (1973). Determining the stage of the estrous cycle in the mouse by the appearance of the vagina. *Biol. Reprod.* 8, 491–494. doi: 10.1093/biolreprod/8.4.491

Chen, V. S., Morrison, J. P., Southwell, M. F., Foley, J. F., Bolon, B., and Elmore, S. A. (2017). Histology atlas of the developing prenatal and postnatal mouse central nervous system, with emphasis on prenatal days E7.5 to E18.5. *Toxicol. Pathol.* 45, 705–744. doi: 10.1177/0192623317728134

Chiaramello, A., Neuman, K., Palm, K., Metsis, M., and Neuman, T. (1995). Helix-loop-helix transcription factors mediate activation and repression of the p75LNGFR gene. *Mol. Cell. Biol.* 15, 6036–6044. doi: 10.1128/MCB.15.11.6036

Corneliussen, B., Thornell, A., Hallberg, B., and Grundström, T. (1991). Helix-loop-helix transcriptional activators bind to a sequence in glucocorticoid response elements of retrovirus enhancers. *J. Virol.* 65, 6084–6093. doi: 10.1128/jvi.65.11.6084-6093.1991

de Pontual, L., Mathieu, Y., Golzio, C., Rio, M., Malan, V., Boddaert, N., et al. (2009). Mutational, functional, and expression studies of the TCF4 gene in Pitt-Hopkins syndrome. *Hum. Mutat.* 30, 669–676. doi: 10.1002/humu.20935

Doostparast Torshizi, A., Armoskus, C., Zhang, H., Forrest, M. P., Zhang, S., Souaiaia, T., et al. (2019). Deconvolution of transcriptional networks identifies TCF4 as a master regulator in schizophrenia. *Sci. Adv.* 5:eaau4139. doi: 10.1126/sciadv.aau4139

Einarson, M. B., and Chao, M. V. (1995). Regulation of Id1 and its association with basic helix-loop-helix proteins during nerve growth factor-induced differentiation of PC12 cells. *Mol. Cell. Biol.* 15, 4175–4183. doi: 10.1128/MCB.15.8.4175

ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74. doi: 10.1038/nature11247

Fischer, B., Azim, K., Hurtado-Chong, A., Ramelli, S., Fernández, M., and Raineteau, O. (2014). E-proteins orchestrate the progression of neural stem cell differentiation in the postnatal forebrain. *Neural Dev.* 9:23. doi: 10.1186/1749-8104-9-23

Forrest, M. P., Hill, M. J., Kavanagh, D. H., Tansey, K. E., Waite, A. J., and Blake, D. J. (2018). The psychiatric risk gene transcription factor 4 (TCF4) regulates neurodevelopmental pathways associated with schizophrenia, autism, and intellectual disability. *Schizophr. Bull.* 44, 1100–1110. doi: 10.1093/schbul/sbx164

Gelernter, J., Sun, N., Polimanti, R., Pietrzak, R., Levey, D. F., Bryois, J., et al. (2019). Genome-wide association study of posttraumatic stress disorder (PTSD) re-experiencing symptoms in >165,000 US veterans. *Nat. Neurosci.* 22, 1394–1401. doi: 10.1038/s41593-019-0447-7

Jung, M., Häberle, B. M., Tschaikowsky, T., Wittmann, M.-T., Balta, E.-A., Stadler, V.-C., et al. (2018). Analysis of the expression pattern of the schizophrenia-risk and intellectual disability gene TCF4 in the developing and adult brain suggests a role in development and plasticity of cortical and hippocampal neurons. *Mol. Autism.* 9:20. doi: 10.1186/s13229-018-0200-1

Keane, T. M., Goodstadt, L., Danecek, P., White, M. A., Wong, K., Yalcin, B., et al. (2011). Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* 477, 289–294. doi: 10.1038/nature10413

Kennedy, A. J., Rahn, E. J., Paulukaitis, B. S., Savell, K. E., Kordasiewicz, H. B., Wang, J., et al. (2016). Tcf4 regulates synaptic plasticity, DNA methylation, and memory function. *Cell Rep.* 16, 2666–2685. doi: 10.1016/j.celrep.2016.08.004

Kharbanda, M., Kannike, K., Lampe, A., Berg, J., Timmusk, T., and Sepp, M. (2016). Partial deletion of TCF4 in three generation family with non-syndromic intellectual disability, without features of Pitt-Hopkins syndrome. *Eur. J. Med. Genet.* 59, 310–314. doi: 10.1016/j.ejmg.2016.04.003

Kim, H., Berens, N. C., Ochandarena, N. E., and Philpot, B. D. (2020). Region and cell type distribution of TCF4 in the postnatal mouse brain. *Front. Neuroanat.* 14:42. doi: 10.3389/fnana.2020.00042

Kim, H., Gao, E. B., Draper, A., Berens, N. C., Vihma, H., Zhang, X., et al. (2022). Rescue of behavioral and electrophysiological phenotypes in a Pitt-Hopkins syndrome mouse model by genetic restoration of Tcf4 expression. *elife* 11:e72290. doi: 10.7554/eLife.72290

Li, B., Qing, T., Zhu, J., Wen, Z., Yu, Y., Fukumura, R., et al. (2017). A comprehensive mouse Transcriptomic BodyMap across 17 tissues by RNA-seq. *Sci. Rep.* 7:4200. doi: 10.1038/s41598-017-04520-z

Li, H., Zhu, Y., Morozov, Y. M., Chen, X., Page, S. C., Rannals, M. D., et al. (2019). Disruption of TCF4 regulatory networks leads to abnormal cortical development and mental disabilities. *Mol. Psychiatry* 24, 1235–1246. doi: 10.1038/s41380-019-0353-0

Luo, Y., Hitz, B. C., Gabdank, I., Hilton, J. A., Kagda, M. S., Lam, B., et al. (2020). New developments on the encyclopedia of DNA elements (ENCODE) data portal. *Nucleic Acids Res.* 48, D882–D889. doi: 10.1093/nar/gkz1062

Ma, C., Gu, C., Huo, Y., Li, X., and Luo, X.-J. (2018). The integrated landscape of causal genes and pathways in schizophrenia. *Transl. Psychiatry* 8:67. doi: 10.1038/s41398-018-0114-x

Marangi, G., and Zollino, M. (2015). Pitt–Hopkins syndrome and differential diagnosis: a molecular and clinical challenge. *J. Pediatr. Genet.* 4, 168–176. doi: 10.1055/s-0035-1564570

Massari, M. E., and Murre, C. (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol. Cell. Biol.* 20, 429–440. doi: 10.1128/MCB.20.2.429-440.2000

Nurm, K., Sepp, M., Castany-Pladevall, C., Creus-Muncunill, J., Tuvikene, J., Sirp, A., et al. (2021). Isoform-specific reduction of the basic helix-loop-helix transcription factor TCF4 levels in Huntington's disease. *eNeuro* 8:ENEURO.0197-21.2021. doi: 10.1523/ENEURO.0197-21.2021

Persson, P., Jögi, A., Grynfeld, A., Påhlman, S., and Axelson, H. (2000). HASH-1 and E2-2 are expressed in human neuroblastoma cells and form a functional complex. *Biochem. Biophys. Res. Commun.* 274, 22–31. doi: 10.1006/bbrc.2000.3090

Phan, B. N., Bohlen, J. F., Davis, B. A., Ye, Z., Chen, H.-Y., Mayfield, B., et al. (2020). A myelin-related transcriptomic profile is shared by Pitt-Hopkins syndrome models and human autism spectrum disorder. *Nat. Neurosci.* 23, 375–385. doi: 10.1038/s41593-019-0578-x

Pscherer, A., Dörflinger, U., Kirfel, J., Gawlas, K., Rüschoff, J., Buettner, R., et al. (1996). The helix-loop-helix transcription factor SEF-2 regulates the activity of a novel initiator element in the promoter of the human somatostatin receptor II gene. *EMBO J.* 15, 6680–6690. doi: 10.1002/j.1460-2075.1996.tb01058.x

Quednow, B. B., Ettinger, U., Mössner, R., Rujescu, D., Giegling, I., Collier, D. A., et al. (2011). The schizophrenia risk allele C of the TCF4 rs9960767 polymorphism disrupts sensorimotor gating in schizophrenia Spectrum and healthy volunteers. *J. Neurosci.* 31, 6684–6691. doi: 10.1523/JNEUROSCI.0526-11.2011

Ravanpay, A. C., and Olson, J. M. (2008). E protein dosage influences brain development more than family member identity. *J. Neurosci. Res.* 86, 1472–1481. doi: 10.1002/jnr.21615

Ripke, S., Neale, B. M., Corvin, A., Walters, J. T., Farh, K.-H., Holmans, P. A., et al. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511, 421–427. doi: 10.1038/nature13595

Sarkar, D., Shariq, M., Dwivedi, D., Krishnan, N., Naumann, R., Bhalla, U. S., et al. (2021). Adult brain neurons require continual expression of the schizophrenia-risk gene Tcf4 for structural and functional integrity. *Transl. Psychiatry* 11, 494–411. doi: 10.1038/s41398-021-01618-x

Schmitt, B. M., Rudolph, K. L. M., Karagianni, P., Fonseca, N. A., White, R. J., Talianidis, I., et al. (2014). High-resolution mapping of transcriptional dynamics across tissue development reveals a stable mRNA-tRNA interface. *Genome Res.* 24, 1797–1807. doi: 10.1101/gr.176784.114

Sepp, M., Kannike, K., Eesmaa, A., Urb, M., and Timmusk, T. (2011). Functional diversity of human basic helix-loop-helix transcription factor TCF4 isoforms generated by alternative 5′ exon usage and splicing. *PLoS One* 6:e22138. doi: 10.1371/journal.pone.0022138

Sepp, M., Pruunsild, P., and Timmusk, T. (2012). Pitt–Hopkins syndrome-associated mutations in TCF4 lead to variable impairment of the transcription factor function ranging from hypomorphic to dominant-negative effects. *Hum. Mol. Genet.* 21, 2873–2888. doi: 10.1093/hmg/dds112

Sepp, M., Vihma, H., Nurm, K., Urb, M., Page, S. C., Roots, K., et al. (2017). The intellectual disability and schizophrenia associated transcription factor TCF4 is regulated by neuronal activity and protein kinase a. *J. Neurosci.* 37, 10516–10527. doi: 10.1523/JNEUROSCI.1151-17.2017

Shafik, A. M., Zhang, F., Guo, Z., Dai, Q., Pajdzik, K., Li, Y., et al. (2021). N6-methyladenosine dynamics in neurodevelopment and aging, and its potential role in Alzheimer's disease. *Genome Biol.* 22:17. doi: 10.1186/s13059-020-02249-z

Sirp, A., Leite, K., Tuvikene, J., Nurm, K., Sepp, M., and Timmusk, T. (2020). The Fuchs corneal dystrophy-associated CTG repeat expansion in the TCF4 gene affects transcription from its alternative promoters. *Sci. Rep.* 10:18424. doi: 10.1038/s41598-020-75437-3

Sirp, A., Roots, K., Nurm, K., Tuvikene, J., Sepp, M., and Timmusk, T. (2021). Functional consequences of TCF4 missense substitutions associated with Pitt-Hopkins syndrome, mild intellectual disability, and schizophrenia. *J. Biol. Chem.* 297:101381. doi: 10.1016/j.jbc.2021.101381

Söllner, J. F., Leparc, G., Hildebrandt, T., Klein, H., Thomas, L., Stupka, E., et al. (2017). An RNA-Seq atlas of gene expression in mouse and rat normal tissues. *Sci. Data* 4:170185. doi: 10.1038/sdata.2017.185

Soosaar, A., Chiaramello, A., Zuber, M. X., and Neuman, T. (1994). Expression of basic-helix-loop-helix transcription factor ME2 during brain development and in the regions of neuronal plasticity in the adult brain. *Mol. Brain Res.* 25, 176–180. doi: 10.1016/0169-328X(94)90297-6

Stefansson, H., Ophoff, R. A., Steinberg, S., Andreassen, O. A., Cichon, S., Rujescu, D., et al. (2009). Common variants conferring risk of schizophrenia. *Nature* 460, 744–747. doi: 10.1038/nature08186

Tamberg, L., Jaago, M., Säälik, K., Sirp, A., Tuvikene, J., Shubina, A., et al. (2020). Daughterless, the drosophila orthologue of TCF4, is required for associative learning and maintenance of the synaptic proteome. *Dis. Model. Mech.* 13:dmm042747. doi: 10.1242/dmm.042747

Tamberg, L., Sepp, M., Timmusk, T., and Palgi, M. (2015). Introducing Pitt-Hopkins syndrome-associated mutations of TCF4 to drosophila daughterless. *Biol. Open* 4, 1762–1771. doi: 10.1242/bio.014696

Thaxton, C., Kloth, A. D., Clark, E. P., Moy, S. S., Chitwood, R. A., and Philpot, B. D. (2018). Common pathophysiology in multiple mouse models of Pitt-Hopkins syndrome. *J. Neurosci.* 38, 918–936. doi: 10.1523/JNEUROSCI.1305-17.2017

Vied, C., Ray, S., Badger, C.-D., Bundy, J. L., Arbeitman, M. N., and Nowakowski, R. S. (2016). Transcriptomic analysis of the hippocampus from six inbred strains of mice suggests a basis for sex-specific susceptibility and severity of neurological disorders. *J. Comp. Neurol.* 524, 2696–2710. doi: 10.1002/cne.23989

Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., et al. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* 50, 668–681. doi: 10.1038/s41588-018-0090-3

Yu, Y., Fuscoe, J. C., Zhao, C., Guo, C., Jia, M., Qing, T., et al. (2014). A rat RNA-Seq transcriptomic BodyMap across 11 organs and 4 developmental stages. *Nat. Commun.* 5:3230. doi: 10.1038/ncomms4230

Zhuang, Y., Cheng, P., and Weintraub, H. (1996). B-lymphocyte development is regulated by the combined dosage of three basic helix-loop-helix genes, E2A, E2-2, and HEB. *Mol. Cell. Biol.* 16, 2898–2905. doi: 10.1128/MCB.16.6.2898

Zollino, M., Zweier, C., Van Balkom, I. D., Sweetser, D. A., Alaimo, J., Bijlsma, E. K., et al. (2019). Diagnosis and management in Pitt-Hopkins syndrome: first international consensus statement. *Clin. Genet.* 95, 462–478. doi: 10.1111/cge.13506

Zweier, C., Peippo, M. M., Hoyer, J., Sousa, S., Bottani, A., Clayton-Smith, J., et al. (2007). Haploinsufficiency of TCF4 causes Syndromal mental retardation with intermittent hyperventilation (Pitt-Hopkins syndrome). *Am. J. Hum. Genet.* 80, 994–1001. doi: 10.1086/515583

Curriculum vitae

Personal data

Name	Laura Tamberg
Date of birth	22.08.1988
Place of birth	Tallinn, Estonia
Citizenship	Estonia

Contact data

E-mail	laura.tamberg@taltech.ee

Education

2014–2023	Tallinn University of Technology, PhD
2012–2014	Tallinn University of Technology, MSc (gene technology)
2007–2010	University of Tartu, BSc (gene technology)
2004–2007	Tallinn Technical High School, graduated with honours

Language competence

Estonian	Native
English	Fluent
Russian	Beginner

Professional employment

2022–2023	University of Tartu, Faculty of Medicine, Institute of Clinical Medicine, Clinical Medicine Specialist (0,75)
2022–...	Tallinn University of Technology, School of Science, Department of Chemistry and Biotechnology, Engineer (0,20)
2022–...	Tallinn University of Technology, School of Science, Department of Chemistry and Biotechnology, Engineer (0,40)
2020–2022	Tallinn University of Technology, School of Science, Department of Chemistry and Biotechnology, Engineer (0,50)
2017–2020	Tallinn University of Technology, School of Science, Department of Chemistry and Biotechnology, Engineer (0,10)
2014–2016	Tallinn University of Technology, Faculty of Science, Department of Gene Technology, Engineer (1,00)

Scholarships and awards

2016	Dora Pluss T1.1 mobility scholarship to attend „Neurofly 2016“ conference
2014	The best poster award at the annual conference of of Estonian Human Genetics Society
2014	Estonian Academy of Sciences Students’ Research Prize

Courses and conferences

September 2022	Estonian Human Genetics Society 24th annual conference, participation, Pärnu, Estoniae
November 2021	Estonian Human Genetics Society 23rd annual conference, participation, Haapsalu, Estonia
February 2020	Doctoral Winter School of Chemistry and Biotechnology, participation
November 2019	Doctoral Science Day, oral presentation, Jõgeva, Estonia

November 2019	Estonian Human Genetics Society 21st annual conference, participation, Pärnu, Estonia
August 2019	Participation in course „Baltic Summer School on Rodent Behavioural Models“, Pühajärve, Estonia
June 2019	„Nordic Neuroscience“ conference, poster presentation „Silencing of *Drosophila* ortholog of TCF4 leads to behavioural impairment“, Helsinki, Finland
November 2018	„Neuroscience 2018“ conference, poster presentation „Silencing of *Drosophila* ortholog of TCF4 leads to behavioural
October 2018	Estonian Human Genetics Society 20th annual conference, participation, Viljandi, Estonia impairment“, San Diego, USA
September 2018	„Neurofly 2018“ conference, poster presentation „Silencing of *Drosophila* ortholog of TCF4 leads to behavioural impairment“, Krakov, Poland
September 2018	Participation in course „Developmental Biology Minisymposium“, Tallinn, Estonia
Mai 2018	Doctoral Science Day, poster presentation, Vehendi, Estonia
November 2017	Estonian Human Genetics Society 19th annual conference, oral presentation „Pitt-Hopkinsi sündroomi mudeldamine äädikakärbses“, Rakvere, Estonia
September 2016	„Neurofly 2016“ conference, poster presentation „Silencing of *Drosophila* ortholog of TCF4 leads to behavioural impairment“, Platanias, Greece
October 2014	„Neurofly 2014“ conference, poster presentation „Introducing Pitt-Hopkins-associated mutations to *Drosophila daughterless*“, Hersonissos, Greece
October 2014	Estonian Human Genetics Society 16th annual conference, participation, Otepää, Estonia
October 2013	Estonian Human Genetics Society 15th annual conference, participation, Narva, Estonia
October 2012	Estonian Human Genetics Society 14th annual conference, participation, Haapsalu, Estonia

Supervised dissertations

Loviisa Pihlas, Master´s Degree, 2023, (sup) Laura Tamberg, Alex Sirp,
Interaction studies of *Drosophila* transcription factors Daughterless and Schnurri, Tallinn University of Technology, School of Science, Department of Chemistry and Biotechnology

Käthy Rannaste, Master's Degree, 2021, (sup) Laura Tamberg; Mari Palgi,
Creating tools for investigating the transcriptional activity of Daughterless *in vitro* and *in vivo*, Tallinn University of Technology, School of Science, Department of Chemistry and Biotechnology

Loviisa Pihlas, Bachelor's Degree, 2021, (sup) Laura Tamberg; Alex Sirp,
Transkriptsioonfaktori Schnurri vaigistamise ja üleekspressiooni mõju uurimine äädikakärbse närvisüsteemis, Tallinn University of Technology, School of Science, Department of Chemistry and Biotechnology

Lisette Marleen Mikk, Bachelor's Degree, 2021, (sup) Laura Tamberg; Mari Palgi,

HA-märgistatud transkriptsioonifaktor Schnurri ekspressiooni kirjeldamine *Drosophila melanogaster*-i embrüotes, Tallinn University of Technology, School of Science, Department of Chemistry and Biotechnology

Anastassia Šubina, Bachelor's Degree, 2018, (sup) Laura Tamberg; Tõnis Timmusk, FLAG-märgistatud Daughterless valgu ekspressiooni ja transaktivatsioonivõime uurimine, Tallinn University of Technology, School of Science, Department of Chemistry and Biotechnology

Publications

Esvald, E.-E.*, Tuvikene, J.*, Kiir, C. S.*, Avarlaid, A., Tamberg, L., Sirp, A., Shubina, A., Cabrera-Cabrera, F., Pihlak, A., Koppel, I., Palm, K. and Timmusk, T. (2023). Revisiting the expression of BDNF and its receptors in mammalian development. Frontiers in Molecular Neuroscience 16. doi: 10.3389/fnmol.2023.1182499.

Sirp, A.*, Shubina, A.*, Tuvikene, J., Tamberg, L., Kiir, C. S., Kranich, L. and Timmusk, T. (2022). Expression of alternative transcription factor 4 mRNAs and protein isoforms in the developing and adult rodent and human tissues. Frontiers in Molecular Neuroscience 15. DOI 10.3389/fnmol.2022.1033224.

Tamberg, L., Jaago, M., Säälik, K., Sirp, A., Tuvikene, J., Shubina, A., Kiir, C. S., Nurm, K., Sepp, M., Timmusk, T and Palgi, M. (2020). Daughterless, the *Drosophila* orthologue of TCF4, is required for associative learning and maintenance of the synaptic proteome. Disease Models and Mechanisms 13, dmm042747. doi:10.1242/dmm.042747.

Tamberg, L., Sepp, M., Timmusk, T. and Palgi, M. (2015) Introducing Pitt-Hopkins syndrome-associated mutations of TCF4 to *Drosophila* daughterless. Biology Open 4, 1762–1771. doi:10.1242/bio.014696.

Elulookirjeldus

Isikuandmed

Nimi	Laura Tamberg
Sünniaeg	22.08.1988
Sünnikoht	Tallinn, Eesti
Kodakondsus	Eesti

Kontaktandmed

E-post	laura.tamberg@taltech.ee

Hariduskäik

2014–2023	Tallinna Tehnikaülikool, PhD
2012–2014	Tallinna Tehnikaülikool, MSc (geenitehnoloogia)
2007–2010	Tartu Ülikool, BSc (geenitehnoloogia)
2004–2007	Tallinna Tehnikagümnaasium, lõpetatud kuldmedaliga

Keelteoskus

Eesti keel	emakeel
Inglise keel	kõrgtase
Vene keel	algtase

Teenistuskäik

2022–2023	Tartu Ülikool, Meditsiiniteaduste valdkond, kliinilise meditsiini instituut, kliinilise meditsiini assistent (0,75)
2022–...	Tallinna Tehnikaülikool, Loodusteaduskond, Keemia ja biotehnoloogia instituut, insener (0,20)
2022–...	Tallinna Tehnikaülikool, Loodusteaduskond, Keemia ja biotehnoloogia instituut, insener (0,40)
2020–2022	Tallinna Tehnikaülikool, Loodusteaduskond, Keemia ja biotehnoloogia instituut, insener (0,50)
2017–2020	Tallinna Tehnikaülikool, Loodusteaduskond, Keemia ja biotehnoloogia instituut, insener (0,10)
2014–2016	Tallinna Tehnikaülikool, Matemaatika-loodusteaduskond, Geenitehnoloogia instituut, insener (1,00)

Stipendiumid ja tunnustused

2016	Dora Pluss T1.1 mobiilsustipendium „Neurofly 2016“ konverentsil osalemiseks
2014	Parima posteri preemia Eesti Inimesegeneetika Ühingu aastakonverentsil
2014	Eesti Teaduste Akadeemia üliõpilaste teadustööde preemia

Kursused ja konverentsid

September 2022	Eesti Inimesegeneetikaühingu 24. aastakonverentsil osalemine, Pärnu, Eesti
November 2021	Eesti Inimesegeneetikaühingu 23. aastakonverentsil osalemine, Haapsalu, Eesti
Veebruar 2020	Keemia ja botehnoloogia doktorikooli talvekool, osalemine
November 2019	Doktorantide teaduspäev, suuline ettekanne, Jõgeva, Eesti
November 2019	Eesti Inimesegeneetikaühingu 21. aastakonverentsil osalemine, Pärnu, Eesti

August 2019	Osalemine kursusel „Baltic Summer School on Rodent Behavioural Models“, Pühajärve, Eesti
Juuni 2019	„Nordic Neuroscience“ konverents, posterettekanne „Silencing of *Drosophila* ortholog of TCF4 leads to behavioural impairment“, Helsingi, Soome
November 2018	„Neuroscience 2018“ konverents, posterettekanne „Silencing of *Drosophila* ortholog of TCF4 leads to behavioural impairment“, San Diego, Ameerika Ühendriigid
Oktoober 2018	Eesti Inimesegeneetikaühingu 20. aastakonverentsil osalemine, Viljandi, Eesti
September 2018	„Neurofly 2018“ konverents, posterettekanne „Silencing of *Drosophila* ortholog of TCF4 leads to behavioural impairment“, Krakov, Poola
September 2018	Osalemine kursusel „Developmental Biology Minisymposium“, Tallinn, Eesti
Mai 2018	Doktorantide teaduspäev, posterettekanne, Vehendi, Eesti
November 2017	Eesti Inimesegeneetikaühingu 19. aastakonverents, suuline ettekanne „Pitt-Hopkinsi sündroomi mudeldamine äädikakärbses“, Rakvere, Eesti
September 2016	„Neurofly 2016“ konverents, posterettekanne „Silencing of *Drosophila* ortholog of TCF4 leads to behavioural impairment“, Platanias, Kreeka
Oktoober 2014	„Neurofly 2014“ konverents, posterettekanne „Introducing Pitt-Hopkins-associated mutations to *Drosophila daughterless*“, Hersonissos, Kreeka
Oktoober 2014	Eesti Inimesegeneetikaühingu 16. aastakonverents, posterettekanne „Pitt-Hopkinsi sündroomi mudeldamine äädikakärbses, Otepää, Eesti
Oktoober 2013	Eesti Inimesegeneetikaühingu 15. aastakonverentsil osalemine, Narva, Eesti
Oktoober 2012	Eesti Inimesegeneetikaühingu 14. aastakonverentsil osalemine, Haapsalu, Eesti

Juhendatud väitekirjad

Loviisa Pihlas, magistrikraad, 2023, (sup) Laura Tamberg, Alex Sirp,
Interaction studies of *Drosophila* transcription factors Daughterless and Schnurri, Tallinna Tehnikaülikool, Loodusteaduskond, Keemia ja biotehnoloogia instituut

Käthy Rannaste, magistrikraad, 2021, (juh) Laura Tamberg; Mari Palgi,
Creating tools for investigating the transcriptional activity of Daughterless *in vitro* and *in vivo*, Tallinna Tehnikaülikool, Loodusteaduskond, Keemia ja biotehnoloogia instituut

Loviisa Pihlas, bakalaureusekraad, 2021, (juh) Laura Tamberg; Alex Sirp,
Transkriptsioonfaktori Schnurri vaigistamise ja üleekspressiooni mõju uurimine äädikakärbse närvisüsteemis, Tallinna Tehnikaülikool, Loodusteaduskond, Keemia ja biotehnoloogia instituut

Lisette Marleen Mikk, bakalaureusekraad, 2021, (juh) Laura Tamberg; Mari Palgi,
HA-märgistatud transkriptsioonifaktor Schnurri ekspressiooni kirjeldamine *Drosophila melanogaster*-i embrüotes, Tallinna Tehnikaülikool, Loodusteaduskond, Keemia ja biotehnoloogia instituut

Anastassia Šubina, bakalaureusekraad, 2018, (juh) Laura Tamberg; Tõnis Timmusk, FLAG-märgistatud Daughterless valgu ekspressiooni ja transaktivatsioonivõime uurimine, Tallinna Tehnikaülikool, Loodusteaduskond, Keemia ja biotehnoloogia instituut

Publikatsioonid

Esvald, E.-E.*, Tuvikene, J.*, Kiir, C. S.*, Avarlaid, A., Tamberg, L., Sirp, A., Shubina, A., Cabrera-Cabrera, F., Pihlak, A., Koppel, I., Palm, K. and Timmusk, T. (2023). Revisiting the expression of BDNF and its receptors in mammalian development. Frontiers in Molecular Neuroscience 16. doi: 10.3389/fnmol.2023.1182499.

Sirp, A.*, Shubina, A.*, Tuvikene, J., Tamberg, L., Kiir, C. S., Kranich, L. and Timmusk, T. (2022). Expression of alternative transcription factor 4 mRNAs and protein isoforms in the developing and adult rodent and human tissues. Frontiers in Molecular Neuroscience 15. DOI 10.3389/fnmol.2022.1033224.

Tamberg, L., Jaago, M., Säälik, K., Sirp, A., Tuvikene, J., Shubina, A., Kiir, C. S., Nurm, K., Sepp, M., Timmusk, T and Palgi, M. (2020). Daughterless, the *Drosophila* orthologue of TCF4, is required for associative learning and maintenance of the synaptic proteome. Disease Models and Mechanisms 13, dmm042747. doi:10.1242/dmm.042747.

Tamberg, L., Sepp, M., Timmusk, T. and Palgi, M. (2015) Introducing Pitt-Hopkins syndrome-associated mutations of TCF4 to *Drosophila* daughterless. Biology Open 4, 1762–1771. doi:10.1242/bio.014696.

ISSN 2585-6901 (PDF)
ISBN 978-9916-80-036-2 (PDF)