KU LEUVEN

FACULTEIT PSYCHOLOGIE EN
PEDAGOGISCHE WETENSCHAPPEN

Effect of Green Tea Extracts Enriched in Epigallocattechin-3-Gallate (GTE-EGCG) on Anatomical and Behavioural Phenotype in a Mouse Model of Down Syndrome

Masterproef aangeboden tot het
verkrijgen van de graad van
Master of Science in de
psychologie
Door

Vicky Van Bulck

promotor: Dr. Zsuzsanna Vegh
copromotor: Prof. Dr. Rudi D'Hooge
m.m.v: Prof. Dr. Greetje Vande Velde
& Sergi Llambrich Ferré

Academiejaar 2020-2021

KU LEUVEN

FACULTEIT PSYCHOLOGIE EN
PEDAGOGISCHE WETENSCHAPPEN

Effect of Green Tea Extracts Enriched in Epigallocattechin-3-Gallate (GTE-EGCG) on Anatomical and Behavioural Phenotype in a Mouse Model of Down Syndrome

Masterproef aangeboden tot het
verkrijgen van de graad van
Master of Science in de
psychologie
Door
Vicky Van Bulck

promotor: Dr. Zsuzsanna Vegh
copromotor: Prof. Dr. Rudi D'Hooge
m.m.v: Prof. Dr. Greetje Vande Velde
& Sergi Llambrich Ferré

Academiejaar 2020-2021

Summary

Down syndrome (DS) is caused by the presence of an extra, third, copy of the human chromosome 21 (HSA21). Since multiple genes are affected by this triplication, the disorder affects several organs, leading to several co-morbid diagnoses. The phenotype seen in all individuals with DS typically consists of muscle weakness, craniofacial dysmorphologies and cognitive deficits. Even though a lot of research is being conducted to finding the cause of this phenotype, many questions remain, making it harder to find a therapeutic treatment. Overexpression of dual-specificity tyrosine-regulated protein kinase 1 (DYRK1A), one of the genes found in the Down Syndrome Critical Region (DSCR) of HSA21, is thought to be one of the possible genes having a key role in the cognitive and skeletal phenotype observed in DS. Since epigallocatechin-3-gallate (EGCG) is an inhibitor of DYRK1A, and is considered safe in consumption for humans, this makes EGCG a good candidate for attempting to reduce these phenotypes.

In this thesis the Ts65Dn mouse model for DS was used in order to examine the effects of green tea extracts containing EGCG (GTE-EGCG) on craniofacial malformations, as well as the cognitive profile seen in Ts65Dn mice. GTE-EGCG was administered daily with a concentration of 0.09 mg/ml by mixing it into the drinking water. The effects of the treatment on the bone structure was measured using *in vivo* µCT-scans on post-natal day 3 (PD3), PD14, PD14 and post-natal week 28 (W28). Cognitive deficits were measured using a behavioural test battery (open field, elevated plus maze, sociability/preference for social novelty, novel object recognition and passive avoidance) at W20. This was repeated again at W30, after the treatment was stopped, to investigate the long-term effects of the treatment.

Results showed differences between trisomic Ts65Dn mice (TS) and wild type mice (WT) in the shape of the mandible. GTE-EGCG treatment seemed to have an effect on TS mice at PD3, however it was not able to completely rescue the mandible shape. At PD14 and PD29, treatment had no effect on the TS mice, however it had a detrimental effect on the WT mice, making them more TS-like. At W28, treatment had no effect on TS, nor WT mice. For the behavioural testing, TS untreated mice were significantly different from the WT untreated mice in some of the tests and at some ages. Further research is required to further progress our knowledge on the phenotypes observed in DS and its underlying mechanisms.

Acknowledgements

First of all I would like to thank my promotor, dr. Zsuzsanna Vegh for her never ending support, guidance, encouragement, patience and knowledge. Without her, this thesis would not have been possible. Many thanks go to Sergi Llambrich Ferré as well, not only for trusting me to work on his project, but also for all the guidance he gave me. I would also like to thank the Laboratory of Biological Psychology for allowing me, as a student, into their lab and giving me the opportunity to work with mice.

Furthermore I would like to thank my parents, grandparents, brothers and sister, they all gave me the love of studying and wanting to know more. Without the endless support of my parents I would not have been able to study, and I am forever grateful for that.

I thank my friends for being there to support me. In particularly Julia, for proof-reading my thesis and for the support she gave me. And of course for all the wonderful conversations we have had over the last six years.

Many thanks go to the big open source community and its user base, without their countless hours spent on building tools and talking about them on fora I would not have been able to write my thesis in LaTeX or complete a horrendous amount of statistics using PSPP on my beloved Macbook that is running Linux.

And of course, a special thanks go to my partner, Marnick, for reading and rereading my thesis several times as well as all the other things he did for me. Thank you for introducing me to Linux and the values of open source software, as well as always fixing my issues with them. Thank you for constantly reminding me the deadline was close and I needed to move on to the next part of my thesis. Without you, I would have stopped writing this thesis many times. Jag älskar dig.

Clarification of approach and contribution

During the first year of this master thesis, I started by reading several articles about EGCG and behavioural testing in Ts65Dn mice. These articles were provided to me by Sergi Llambrich Ferré. In addition to immersing myself in the topic I took the course Laboratory Animal Science in order to obtain the FELASA certificate allowing me to work with lab animals. After I got this certificate I started working in the laboratory. I was responsible for the care, tail colouring and behavioural experiments of the untreated mice. Dr. Zsuzsanna Vegh explained the test set-ups and how to use the software. I collected the data of the first behavioural tests. However due to the outbreak of the COVID-19 pandemic, I was not able to go to the laboratory myself any longer, as a consequence Sergi and Zsuzsanna took over the care and data collection of the mice. They collected the data of the second round of behavioural testing as well as the passive avoidance test in the first round of behavioural testing. Whenever I had any problems in the laboratory I could contact Zsuzsanna.

Through the first and second year of this master thesis I continuously worked on the shape analysis. To start out I visited the Molecular Small Animal Imaging Center (MoSAIC), where the mice were first housed and scanned using a µCT and µMRI scanner. Here I could observe first-hand how the scanning procedures were done. During the first year, Sergi taught me how to place the landmarks on the skull and mandible. I practised this on my own and had meetings with Sergi about this regularly. As soon as I reached a good level of consistency, I started landmarking the mandibles of the treated and untreated mice of the *FUNSTRUCK* project, as well as an earlier project (*BRAINFACE*). Afterwards I performed the shape analysis of the mandibles of the *FUNSTRUCK* project. Sergi explained to me how to do this and helped me with the interpretation of the results.

The second year of this master thesis was focused on the analysis and the writing of the thesis itself. Based on the literature I received in my first year, I searched for more literature myself. The analysis of the behavioural data was done with Gnu PSPP, with the necessary feedback of Zsuzsanna. Zsuzsanna and Sergi always answered all my questions quickly and provided me with guidance and feedback during the whole thesis.

Contents

List of Figures

List of Tables

1 Introduction

Down syndrome (DS) is one of the most common genetic disorders in humans and is caused by the presence of an extra, third, copy of the human chromosome 21 (*Homo sapiens* chromosome 21, HSA21) (Dierssen et al., 2009). DS affects around 1 in 700 live births and has a global incidence of more than 200.000 cases per year (Grieco et al., 2015). However, considering the rate of fetal loss, stillbirth, and amount of spontaneous and elective abortion with DS, this is an underestimation of its real prevalence (Ethen & Canfield, 2002; Skotko, 2009; Zigman, 2013). DS was first described in an article in 1866 by John Langdon Down, yet 155 years later, we still do not have an effective treatment for this disorder. Nevertheless, a recent increase in research and social inclusion has led to a boost in the life expectancy and quality of life for individuals with DS, leading to a life expectancy of over 50 years in Western countries (Glasson et al., 2002; Wiseman et al., 2009).

1.1 Phenotype of Down Syndrome

Because of the triplication of multiple genes that are situated on HSA21, DS is a disorder that affects several organs. However, not all phenotypes are observed in every individual with DS, including congenital heart defect, eye problems, hearing loss and an increased risk for both leukemia and early-onset Alzheimer's disease (AD) (Antonarakis et al., 2004; Asim et al., 2015). Some phenotypes such as intellectual disability (ID), muscle weakness (hypotonia) and craniofacial dysmorphologies (for example a smaller chin, flatter nasal bridge, underdevelopment of the mandible, brachycephaly and an overly folded helix) are seen in all of the individuals with DS (Antonarakis et al., 2004; Delabar et al., 2006; Richtsmeier et al., 2000). However, even with the phenotypes seen in all individuals with DS, they are not seen in the same severity, for example, ID ranges from mild to moderate in DS (Patterson, 2007). This makes DS a very heterogeneous disorder with many possible outcomes.

Some individuals with DS seem to experience a faster-than-normal ageing (Zigman, 2013). Signs of early ageing are seen in DS throughout the whole body, for example: skin wrinkling, hair loss (alopecia), earlier menopause in women, cataracts, hearing loss and osteoporosis (Zigman, 2013). In addition to physiological ageing, people with DS also experience early-onset neurodegeneration, sometimes even leading to early-onset dementia (Cole et al., 2017). Individuals with DS currently have a mean life expectancy of between 50 and 60 years, which is around 20 years lower than in the average population (Max Roser & Ritchie, 2013; Zigman, 2013). Nevertheless, older cases have been observed, like a 70-year-old man described by Krinsky-McHale et al. (2008) who was healthy and not even showing any signs of AD.

The DSM-5 defines ID as a disorder with onset during the developmental period, which includes both intellectual and adaptive functioning deficits in conceptual, social and practical domains (American Psychiatric Association, 2013). All individuals with DS meet these criteria. ID is very diverse in DS with a range in IQ from 30 to 70 with an average of 50 (Contestabile et al., 2010). In contrast to normal development, IQ in DS is not constant over age, with a first decline during childhood (Dierssen et al., 2009; Vicari et al., 2004). This decline seems to be due to a slower rate of cognitive growth as well as speech and language impairments, rather than a loss of cognitive abilities (Dierssen et al., 2009; Grieco et al., 2015). IQ stabilises in adolescence and adulthood, however it declines again after the age of 40, albeit this decline is not always found in studies (Dressler et al., 2010; Esbensen et al., 2008; Grieco et al., 2015; Steingass et al., 2011). This second decline seems to be because of the increased risk for early-onset AD, as well as the early-onset neurodegeneration (Grieco et al., 2015).

Since an IQ score does not give enough information, because it is too general, many studies have investigated the cognitive profile of individuals with DS. The cognitive profile that is typically seen in DS is characterised by deficits in language, short-term memory, working memory and explicit long-term memory (Dierssen et al., 2009; Steingass et al., 2011). This cognitive profile evolves over time. In early childhood, the deficits in language development are most pronounced (Grieco et al., 2015). However, pre-language behaviour, for example babbling, seems to be normal in DS children (Contestabile et al., 2010). During late childhood and adolescence, problems with short-term and working memory become more noticeable (Grieco et al., 2015). Furthermore, the deficits are more striking if the task becomes more demanding (Contestabile et al., 2010). The cognitive profile of adults with DS is not homogeneous, resulting in some abilities being more developed than others. Relative strengths are visuospatial short-term memory, associative learning and implicit long-term memory (Dierssen et al., 2009).

Even though all individuals with DS have intellectual deficits, they still have some levels of independence. In comparison with adults with similar levels of ID, adults with DS seem to have better adaptive functioning (Esbensen et al., 2008; Steingass et al., 2011). Approximately 60% of young adults are independent with feeding and toileting, 50% can walk around in their own neighbourhood. Most help with household chores and about 50% can do some simple cooking. 30% is independent with dressing (Steingass et al., 2011).

On top of the intellectual deficits, some individuals with DS have additional behaviour and psychiatric diagnoses (Dierssen et al., 2009). The prevalence is higher than in other children, yet smaller than in other individuals with ID (Roizen & Patterson, 2003). In a study done by Myers and Pueschel (1991) 2.1% of the individuals with DS had some sort of psychiatric disorder. The most frequent disorders were disruptive disorders like attention deficit hyperactivity disorder (ADHD; 4.2%) and aggressive behaviour (7.2%). Other common disorders were repetitive behaviours, namely stereotypic behaviour (2.8%) and self-injurious behaviour (1.6%), as well as affective disorder, in particular major depressive disorders (2%). Recently, the comorbid diagnosis with autism has gained attention with a minimum prevalence of 7% (Kent et al., 1999; Roizen & Patterson, 2003).

1.2 Bone structure in Down Syndrome

Individuals with DS usually show a characteristic craniofacial phenotype (Rodrigues et al., 2019). These phenotypes include an overall reduction in skull size (microcephaly), an abnormal calvarial widening (brachycephaly), a flattened occiput, a small midface, a reduced maxima and a reduced mandible (Richtsmeier et al., 2002; Rodrigues et al., 2019). Other typical phenotypes in craniofacial morphology in DS are malocclusion with posterior crossbite and anterior open bite. These malformations are believed to be partly the result of abnormalities occurring during development of the craniofacial skeleton (Richtsmeier et al., 2002; Shukla et al., 2014). Next to the bone malformation, problems with bone mass, such as osteoporosis, have been reported in DS (Abeysekera et al., 2016).

1.3 Brain structure and cognitive profile

Brain structure and cognitive function have a close relationship. Several studies have focused on brain structures in DS, most of these studies are done by structural MRI and/or *post-mortem* analysis (Contestabile et al., 2010). The cognitive deficits that are characteristic for DS are believed to be due to the altered brain structure (Dierssen et al., 2009).

The brain of adults with DS is about 20% smaller when compared to healthy controls, even when correcting for the reduced body size in DS (Contestabile et al., 2010; Rodrigues et al., 2019). This reduction is already present prenatally and starts in 4-5 month old fetuses (Rodrigues et al., 2019). This difference becomes even more pronounced during the last trimester and in childhood (Hamner et al., 2018; Rodrigues et al., 2019). This reduction in brain size might contribute to the below average IQ seen in DS (Cheong et al., 2013; Rushton & Rushton, 2003; Willerman et al., 1991).

Morphological brain alternations have been seen in individuals with DS of various ages (Contestabile et al., 2010; Dierssen et al., 2009). Individuals with DS have disproportional smaller frontal, temporal and cerebellar regions (Contestabile et al., 2010; Dierssen et al., 2009). Research suggests that there are no significant differences between DS and euploid peers in the volume of the hippocampus in early childhood, however as children age these differences become significant, even when adjusting for total brain volume (Hamner et al., 2018; Pinter et al., 2001; Śmigielska-Kuzia et al., 2011). Since the hippocampus is central in many cognitive processes, this reduction of volume might explain the cognitive profile seen in DS (Dierssen et al., 2009). Alternations in the cerebral cortex on the other hand, might explain the deficits in language, as well as working memory and attention (Dierssen et al., 2009). Lastly, changes in the cerebellum might be responsible for deficits in motor control, as well as working memory, affect and language (Dierssen et al., 2009). Even though some brain regions are smaller than the rest of the brain, the subcortical regions seem to be preserved (Dierssen et al., 2009).

On top of the general reduced brain volume and specific morphological brain alterations, individuals with DS experience premature brain ageing (Rodrigues et al., 2019). A study done by Cole et al. (2017) shows that the patterns of brain volume seen in DS resembles that of an on average 2.49 years older healthy individual. Several other studies found changes in DS typical seen in an ageing brain, for instance increased cerebral beta amyloid depositions, neurofibrillary tau tangles, fibrillar deposits in the brain parenchyma (plaques), brain atrophy and white matter lesions (Annus et al., 2016; Cole et al., 2017; Contestabile et al., 2010). These neurodegenerative patterns resemble those of AD and might explain the general decline in IQ seen in adulthood as well as the increased risk for AD (Contestabile et al., 2010).

1.4 The role of DYRK1A

The presence of the extra copy of HSA21 is the cause of the phenotypes seen in DS. There are two hypotheses that provide an explanation of this genotype-phenotype correlation: i) the "gene-dosage effect" hypothesis claims that the phenotypes are caused by the increased dosage of some dosage-sensitive genes and their encoded proteins. This means that some genes are more important than others in the contribution to the phenotypes seen in DS. And ii) the "amplified developmental instability" hypothesis, on the other hand, is saying that HSA21 determines general alternations during homeostasis, meaning all genes equally contribute to the DS phenotypes (Dierssen et al., 2009).

The gene-dosage hypothesis claims that there is a clear correlation between the 1.5-fold overexpression of some specific genes and the brain anomalies and intellectual disability seen in DS (Antonarakis et al., 2004; Dierssen et al., 2009). Past research has described a region, the Down Syndrome Critical Region (DSCR), that is thought to contain the specific genes that cause the cognitive deficits (Antonarakis et al., 2004; Toyoda et al., 2002). Recent research has been focussed on trying to find candidate genes to study their roles in the DS phenotypes, with the goal to understand DS better (Rachidi & Lopes, 2008).

One of these candidate genes that has received a lot of attention is the dual-specificity tyrosine-regulated protein kinase 1 (DYRK1A). DYRK1A functions as a molecular switch for developmental processes, it also regulates neuronal developments as well as neurodegeneration (Becker & Sippl, 2011; Tejedor & Hämmerle, 2011; Wegiel et al., 2011). DYRK1A is one of the genes triplicated in human DS patients, as well as in the Ts65Dn mouse model and in Ts1Cje mice model (see also section 1.6 Mouse models for Down Syndrome). Both of these mouse models have the similar craniofacial phenotypes as well as the typical cognitive profile that are seen in clinical DS (Arron et al., 2006; Dierssen & de Lagrán, 2006). However, mouse models that lack this triplication (for example the Tc1 mice) do not show these phenotypes (Arron et al., 2006). Moreover, transgenic mice overexpressing DYRK1A show cognitive deficits, synaptic plasticity alternations and skeletal deficits similar to those observed in mouse models with DYRK1A overexpression (Rachidi & Lopes, 2008). This evidence suggests that DYRK1A plays an important role in brain development (Dierssen & de Lagrán, 2006). Several studies have been done on the effects of DYRK1A inhibition, resulting in a normalisation of the phenotypes linked with DYRK1A (Blazek et al., 2015; McElyea et al., 2016). This makes DYRK1A inhibitors the perfect candidate to help reduce the phenotypes seen in DS.

1.5 GTE-EGCG

In recent years, a lot of research has been done concerning green tea, indicating its health promoting effects. For instance, the prevention of cancer, anti-inflammatory, anti-arthritic and anti-bacterial effects have been shown (Pervin et al., 2018; Wolfram, 2007). The dry weight of a fresh leaf of green tea consists of approximately 30% polyphenols, mostly flavanols (one of which is epigallocatechin-3-gallate, EGCG) and flavonols (Wolfram, 2007). Because of its health promoting effects, EGCG is sold as a dietary supplement. Even though strong evidence for its health benefits in humans is missing, no negative outcomes have

been found in studies with EGCG (Pervin et al., 2018). Only at high dosage, over 600mg EGCG/person/day, a small increase in liver enzymes was found (Dekant et al., 2017). This makes EGCG a relatively safe compound to be given to humans.

Many attempts have been done to develop a DYRK1A inhibitor, however, this has not been successful, meaning that a specific inhibitor is still lacking (Jarhad et al., 2018). However, many DYRK1A inhibitors from natural and synthetic sources have been found and have been reviewed (Jarhad et al., 2018). In 2003 Bain et al. discovered EGCG as a natural inhibitor or DYRK1A. Nowadays, EGCG is a commonly used DYRK1A inhibitor. EGCG can either be given as pure EGCG, green tea or as green tea extracts containing EGCG (GTE-EGCG).

EGCG was the first DYRK1A inhibitor to show an improvement in cognition in humans, as well as in DS mouse models (Duchon & Herault, 2016). Research with transgenic mice overexpressing DYRK1A showed an improvement in brain structure as well as cognition when given a green tea diet (Guedj et al., 2009). Even though more research is needed on the duration, the dosage and the best start date of the treatment, results are promising and EGCG seems to be a very good candidate to treat DS.

1.6 Mouse models for Down Syndrome

To get a better understanding of DS, especially the cellular and molecular origins, several mouse models have been used to study the effects of DS. Approximately 80% of the murine genes are homologous to human genes, making it possible to use mouse models to approach similar issues in humans (Dierssen et al., 2009; Mouse Genome Sequencing Consortium, 2002). Around 550 genes are found in HSA21, 166 of those are orthologous to genes situated on three mouse chromosomes (*Mus musculus*, Mmu): Mmu16 (110 orthologous genes), Mmu17 (19 orthologous genes) and Mmu10 (37 orthologous genes) (Antonarakis et al., 2004; Rueda et al., 2012). Even though a lot of mouse models have been made over the past years, none of them perfectly mimic HSA21 completely.

Ts16 was the first mouse model for DS and was trisomic for the entire Mmu16, however very few similarities were found with DS, because mmu16 is orthologous with regions of HSA3, HSA8, HSA16 and HSA21; making the model a triplication for more genes than are seen in DS (Rueda et al., 2012). Moreover, since Ts16 embryos die *in utero*, it is not a good model to investigate phenotypes in young and adult mice (Rueda et al., 2012).

Currently the Ts65Dn mouse model is the most used mouse model for DS. The model was developed by Davisson and colleges in 1990 (Schmidt et al., 1995). The TS65Dn mice only have a partial trisomy of Mmu16 (including DYRK1A), making it orthologous for approximately 92 genes in HSA21 (Rueda et al., 2012). However, not all genes found in HSA21 are triplicated in the Ts65Dn mouse model and around 60 genes nonhomologous to HSA21 are triplicated in the mouse model. This makes the Ts65Dn mouse model not completely perfect for DS, nevertheless at the moment it is the best model available (Rueda et al., 2012). Ts65Dn mice share a lot of the phenotypes seen in DS, such as memory impairments, learning and memory deficits, motor dysfunction, alternations in brain structures and neurodegeneration in older mice (Dierssen et al., 2009). Another advantage of the Ts65Dn mouse model is that they live until adulthood, making it possible to investigate long term effects (Rachidi & Lopes, 2008). Yet a disadvantage is the sterility in male adult mice, making it harder to breed Ts65Dn mice.

1.7 The *FUNSTRUCK* project and this thesis

The goal of the *FUNSTRUCK* project is to investigate the long-term effects of GTE-EGCG treatment on the cognitive profile of Ts65Dn mice. To do this, we performed a longitudinal study in which Ts65Dn mice and euploid littermates were followed-up from E0 (embryonic day 0) until W34 (post-natal week 34). Treated and untreated mice as well as TS and WT mice were compared during the whole study. Starting from age E9 until W28, the treated mice received GTE-EGCG treatment with a concentration of 0.09 mg/ml every day. Cognitive deficits were measured with a battery of cognitive tests starting on W20. These tests were done once again starting on W30 after treatment was stopped to check for longitudinal effects of the EGCG treatment. In addition to the cognitive tests, we also performed µCT-scans on PD3, PD14, PD29, W28 and W34 to investigate the effects of the treatment on the bone structure. This thesis serves as a control condition for the *FUNSTRUCK* project and so the scope of this thesis is limited to the cognitive tests and the mandible shape of the untreated mice.

The hypothesis for the *FUNSTRUCK* project is that early prenatal interventions with GTE-EGCG treatment will moderate the known deficits in the cognitive profile that is typical seen in Ts65Dn mice. We also expect these effects to continue after the treatment is stopped, because of the early start of the treatment. In addition to the cognitive profile, we also expected the GTE-EGCG treatment to prevent the alternations in bone structure. As

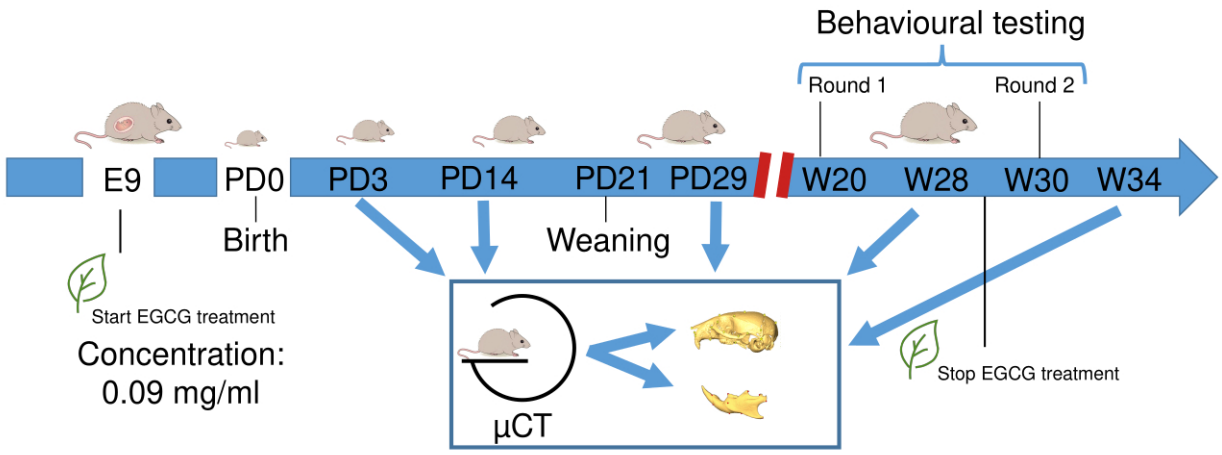
mentioned before, this thesis is limited to the control condition with the untreated mice. The hypothesis for this thesis is that TS mice are performing worse on the cognitive testing than the WT mice. On top of this, the TS mice will also have a different mandible shape than the WT mice.

2 Method

For a complete overview of the *FUNSTRUCK* project see Figure 1.

Figure 1

Timeline For the FUNSTRUCK Project.



Note. Treatment with EGCG was started on E9 and was given continuously until adulthood at W28. The mice were weaned on PD21. µCT-scans were done on PD3, PD14, PD29, W28 and W34. Behavioural testing started on W20 and W30.

2.1 Animals

Trisomic mice and euploid littermates were bred in our lab using WT male and Ts65Dn female mice imported from the Jackson Laboratory in June 2019 (Stock number 003647 and 005252 respectively, Bar Harbor, ME, USA). The genotype of the mice was confirmed by polymerase chain reaction (PCR) test of DNA extracted from tail biopsies. During the whole study the mice were group-housed with the maximum of 5 mice per cage. Before weaning the pups were group-housed together with their mother. After weaning they were grouped by their sex (male/female) and treatment group (treated/untreated, see Table 1 and 2 for the sample

size per treatment group). Some animals died during the scanning procedures, because they did not recover from the anaesthesia. The animals were housed in an animal facility of the KU Leuven with individually ventilated cages. All behavioural experiments were performed during the light phase of a 12h light/dark cycle (lights on 07:00-19:00). Temperature (22°C ± 2) and humidity (50-70%) were controlled for while food and water were supplied *ad libitum*. Cage enrichments were provided in all cages in form of shredded paper, toilet paper rolls and small cardboard boxes. Cages were cleaned every week. Tail colouring in combination with ear cuts were used to identify the mice throughout the study. All experimental procedures complied with the local, national and European regulations and were authorised by the Animal Ethics Committee of KU Leuven (ECD approval number P120/2019).

Table 1

Sample Size for the Treated Mice.

Sex	PD3		PD14		PD29		W20		W28		W30	
	WT	TS	WT	TS	WT	TS	WT	TS	WT	TS	WT	TS
Male	13	8	10	10	10	10	10	10	10	10	10	10
Female	10	15	10	8	9	7	8	7	8	7	7	7
Total	23	23	20	18	19	17	18	17	18	17	17	17

Note. Sample size is mentioned for the scanning on post-natal day 3 (PD3), PD14, PD29 and post-natal week 28 (W28) as well as for the first (W20) and second (W30) round of behavioural testing. Genotype is defined as wild type (WT) or trisomic Ts65Dn (TS). Differences in sample size over time are due to some animals not recovering from the anaesthesia during the scanning procedure.

Table 2

Sample Size for the Untreated Mice.

Sex	PD3		PD14		PD29		W20		W28		W30	
	WT	TS	WT	TS	WT	TS	WT	TS	WT	TS	WT	TS
Male	6	8	5	7	4	5	4	4	3	2	2	3
Female	11	9	8	7	8	6	8	6	8	6	6	6
Total	17	17	13	14	12	11	12	10	11	8	8	9

Note. Sample size is mentioned for the scanning on post-natal day 3 (PD3), PD14, PD29 and post-natal week 28 (W28) as well as for the first (W20) and second (W30) round of behavioural testing. Genotype is defined as wild type (WT) or trisomic Ts65Dn (TS). Differences in sample size over time are due to some animals not recovering from the anaesthesia during the scanning procedure.

2.2 GTE-EGCG treatment

The mice were randomly assigned to the treated or control group by administering GTE-EGCG to half of the pregnant dams, while the other half received untreated water. GTE-EGCG (Mega Green Tea Extract, Life Extension, USA, 45% EGCG) was administered to the treated group via drinking water that was prepared daily at a concentration of 0.09 mg/ml by dissolving it in water. Because GTE-EGCG was prone to oxidation, the drinking bottles were protected from light by covering them. Treatment was initiated prenatally at embryonic day 9 (E9) and sustained until adulthood at the end of the first round of behavioural testing (postnatal week 20; W20). During pregnancy GTE-EGCG was administered via the drinking water to the pregnant dam. Since it can cross the placenta, this administration route affected the embryos as well (Chu et al., 2006). The drinking water of the dam was continued to be provided with the GTE-EGCG treatment after birth, however now the GTE-EGCG was reaching the pups through the mother milk (Souchet et al., 2019). After weaning on postnatal day 21 (PD21) the GTE-EGCG was given directly to the mice via the drinking water. Water intake was monitored closely every day to estimate the total amount of GTE-EGCG intake per cage. On average each pup received 39 mg EGCG/kg/day at PD29, with an average body weight of 13.4 g and average water consumption of 5.8 ml/day. The behavioural testing discussed in this thesis was limited to the untreated mice. However, the shape analysis of the mandible contained both the treated and untreated mice.

2.3 *In vivo* µCT imaging of bone anatomy

The mice were scanned *in vivo* using a high resolution µCT (Skyscan 1278, Bruker microCT, Belgium) on PD3 (infantile), PD14 (juvenile), PD29 (early adult), W28 (late adult, after the first round of behavioural testing) and W34 (late adult, after the second round of behavioural testing; see Figure 1). During the scanning procedure, the mice were anaesthetised using 1.5-2% isoflurane in oxygen while their temperature and breathing were monitored. Afterwards, the scans were reconstructed using NRecon version 1.7.3.1 (Bruker microCT, Belgium). In NRecon the region of interest was selected on the head, so only this would be reconstructed. Four settings were adjusted: i) the dynamic image range was set from .008 minimum to .04 maximum for PD3. For PD14, PD29 and W28 the range was set from .005 to .05. This range determines the range of intensity is shown. ii) The beam-hardening was always set to 10%. Without this setting, the edges will appear brighter than the center. iii) A fine tuning was done for post-alignment, the best fit was manually chosen. The post-alignment setting controls for possible misalignment since this would cause doubled and blurred reconstructions. iv) A fine tuning was done to reduce ring-artifacts. This is a phenomenon that might occur due various reasons such as miscalibration. These reconstructions were then loaded into Amira version 5.2.1 (ThermoFisher scientific, USA) to visualise a 3D model of the head. The correct voxel size was used to get the right resolution for the image. In order to be able to see the skull but not the soft tissue, an appropriate threshold for brightness was manually set for each mouse individually (see Figure 2). These 3D models were then used to place landmarks on the mandible. A set of 49 landmarks was used per mouse, of which 22 landmarks were placed on the mandible (see Figure 3). In a training phase, in which one person was trained to place the landmarks of a few sample mice correct and consistent, the in-trial-user-error was minimised. After placing these landmarks in Amira, the coordinates of all the landmark were all grouped into one file and loaded into MorphoJ version 1.07a (Klingenberg lab, UK).

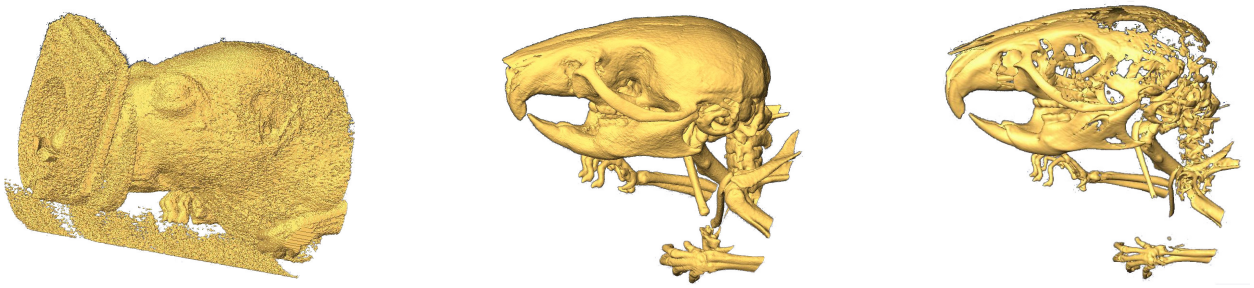
2.4 Shape analysis

We compared the shape of craniofacial structures of treated and untreated WT and TS mice using Geometric Morphometrics (GM). This statistical method was designed for the comparison of shapes based on changes in previously placed landmarks at the places of interest. GM allowed the landmarks, and thus the shape of the craniofacial structure, to be analysed apart from size, position and orientation, making morphology the only variable being compared. GM analyses were done using MorphoJ.

Figure 2

Deciding the Threshold in Amira.

A. B. C.

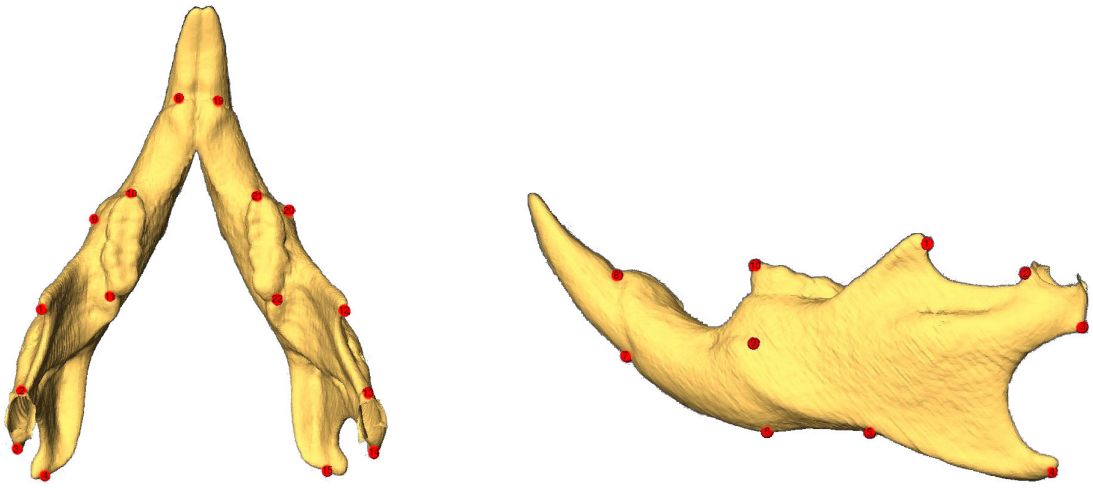


Note. A threshold was manually set for each mouse individually. Panel A shows a mouse at week 28 with threshold 0, all the soft tissue and even the scanner itself can be seen. Panel B shows the same mouse with threshold 30. This would be a good threshold for this mouse since the bone structure can clearly be seen, without showing the soft tissue. Panel C once again shows the same mouse, but this time with threshold 80. This is clearly too high for this mouse, as it also loses a lot of the bone structure.

Figure 3

Landmarks of the Mandible.

A. B.



Note. Panel A shows a top view of the mandible, while panel B shows the right view of the same mandible. Landmarks were placed on biological meaningful places that can also be repeatedly placed with enough accuracy and precision. A list of the anatomical definitions of each landmark can be found in Table A1.

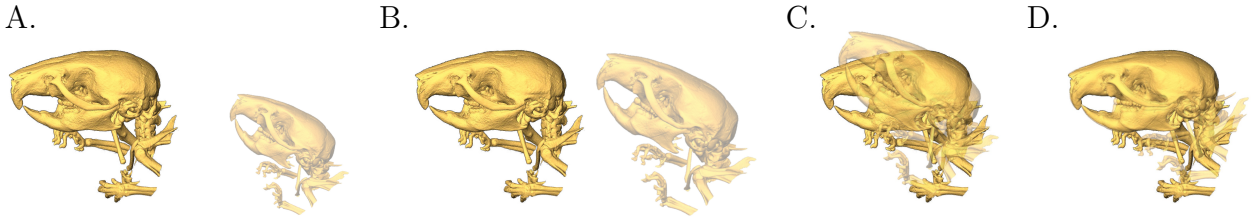
The first step in performing a GM analysis was to perform a General Procrustes Analysis (GPA, see Figure 4). This was done to reduce influences from size, position and orientation, since these were confounding variables. By performing a GPA, the landmarks were transformed to a new, common configuration called Procrustes coordinates, which is used to generate a covariance matrix. This transformation was done in three different steps. The first step was to individually rescale the landmarks to the same centroid size. The centroid was based on the average of the coordinates of the landmarks, the centroid size was measures as the squared root of the sum of squared distances of the landmarks between each landmark and the centroid. The second step was translation, the centroid was placed in the 0,0,0 point of the 3D space. In the third and last step, the landmarks were fitted into a consensus shape to minimise the sum of squared distances between homologous landmarks. The differences between the coordinates of homologous landmarks that were observed after the GPA, were the actual differences in shape between individuals.

After the GPA, a Principal Component Analysis (PCA) was done. This analysis was used to transform a covariance matrix into a smaller number of uncorrelated variables that are called Principal Components (PCs). Each PC explained a certain percentage of variance of the data, with PC1 explaining the biggest source of variance and so on until the entire variance is accounted for. The scores of PC1 and PC2 were plotted in a scatterplot. If there were no differences between the phenotypes of TS, WT and treatment group, all of these would overlap in the scatterplot. However, if there were differences in the phenotypes, the groups would be separated from each other.

The Procrustes distance was used as a quantitative indicator of the shape differences between treated WT, treated TS, untreated WT and untreated TS. The Procrustes distance was calculated as the sum of the squared distances between homologous landmarks of the average of the two groups.

Figure 4

General Procustes Analysis.



Note. Panel A shows the initial configuration. It is clear that the second mouse is smaller, rotated and in a different position from the first mouse. If nothing was done about this, the differences between the coordinates of two homologous landmarks would be very big. Panel B shows the first step of the GPA. The second mouse was scaled to the same centroid size as the other one. Panel C shows the translation step of the GPA. The mice were placed at the 0,0,0 point of the 3D space. Panel D shows the third and final step of the GPA. In this step the second mouse was rotated to minimise the differences between homologous landmarks. The mandible and skull of the two mice were overlapping greatly. Differences in the paws can still be seen, however since no landmarks were placed here this will not influence the data. Differences that could be seen in the skull and mandible between both mice were differences in morphology, without influences of size, position and orientation.

In order to statistically test the shape differences between the different groups, permutation tests were performed. The null hypothesis was that there were no differences between the four groups (WT-treated, TS-untreated, WT-untreated, TS-untreated). To obtain the initial test statistic the difference between the mean shapes of each group was calculated. The permutation tests permuted (shuffled) the treatment assignment and re-calculated the test statistic for each permutation. If the null hypothesis was true, no difference would have been seen since treatment group would not have mattered. This permutation was done 10.000 times using a computer simulation. The p-value is calculated based on how many test-statistics are more extreme than the one we calculated, divided by the total number of test-statistics.

2.5 Behavioural testing

Each behavioural test was performed twice, a first time starting at 20 weeks postnatal (W20; see Figure 5) and a second time starting at W30 after treatment was stopped to check for the longitudinal effects of the treatment (see Figure 6). In between the two rounds of testing, the mice were not tested for 3 weeks. The behavioural test battery was composed of five

Figure 5

Behavioural Testing Round 1 Timeline.

Note. Timeline for the first round of behavioural testing. In week 23 and 24 no tests were done.

Figure 6

Behavioural Testing Round 2 Timeline.

Note. Timeline for the second round of behavioural testing. The timeline is different from the first round, due to the COVID-19 pandemic.

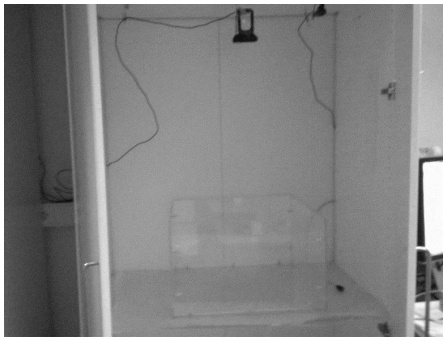
experiments: Open Field (OF), Elevated Plus Maze (EPM), Sociability/Preference for Social Novelty (SPSN), Novel Object Recognition (NOR) and Passive Avoidance (PA). Each test will be explained in more detail in the following section.

2.5.1 Open Field (OF)

The Open Field protocol allowed for evaluation of free locomotion in an open field, and assessment of anxiety-related exploration. It relied on the tendency of mice to explore new environments. For the OF protocol a 50 x 50 cm square arena was used. The arena was brightly illuminated and transparent plexiglas walls were used to induce anxiety-related behaviour (see Figure 7). Half an hour prior to the experiment, the animals were placed in a dark environment to increase the contrast with the brightly lit open field and induce anxiety related behaviour. At the start of the experiment the mouse was placed in the bottom left corner of the arena, facing the wall. After 60 s habituation the mouse's behaviour was tracked for 10 min using the software ANY-maze (Stoelting Inc., Wood Dale, IL, United States). In between each mouse the arena was cleaned using a soft tissue. Parameters such as distance travelled, walking speed and time spent in the center and periphery were extracted from tracked paths. Path length was considered as a measurement for general exploration and locomotive activity. Time spent in the periphery could be used as a measurement of anxiety.

Figure 7

Set-up of the OF.



Note. Set-up of the Open Field task. The set-up consists of a 50 x 50 x 30 cm plexiglas box, placed inside a white cupboard. Underneath small LED lights illuminate the set-up. A webcam placed above allows tracking of the animal. Image extracted from protocol Nr: 2015_027.

2.5.2 Elevated Plus Maze (EPM)

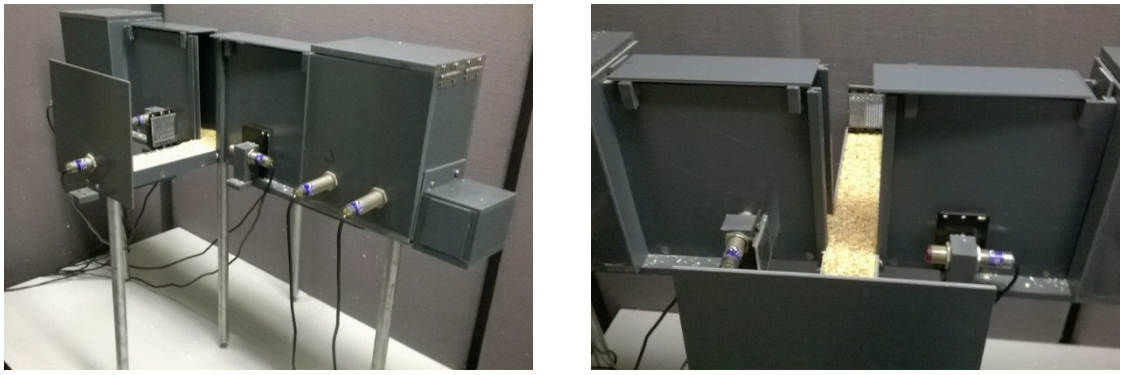
The Elevated Plus Maze protocol measured the tendency of mice to avoid open and elevated spaces by providing open and closed arms to explore. It is therefore a standard protocol to test for anxiety related behaviour. The EPM was a plus-shaped arena and consisted of two closed arms with walls and two open arms without walls. Each arm was 5 cm wide, 21 cm long and 30 cm elevated above the ground (see Figure 8). Half an hour prior to the experiment, the mice were placed in the test room to habituate and settle. At the start of the experiment, the mouse was placed in the center of the maze. During the first minute of the experiment the mouse was allowed to habituate, no data was recorded at this time. During the next 10 min the mouse freely explored the arena and data was recorded using a computerised activity logger. With the use of infra-red beams, the activity logger tracked the total number of open and closed arm visits. This was then used to calculate the percentage of open arm visits. Increased time spent in the open arm was an indication for anxiety.

2.5.3 Sociability/Preference for Social Novelty (SPSN)

The Sociability/Preference for Social Novelty protocol was a measurement for general sociability and interest in social novelty. It relied on the innate preference of mice to explore another mouse (sociability) and to spent more time with a novel mouse than a familiar mouse (social novelty). The SPSN task was performed in a dimly lit plexiglas arena consisting of three compartments: a middle compartment for the tested mouse and two flanking compartments for the stranger mice (see Figure 9). The compartments were separated by transparent

Figure 8

Set-up of the EPM.

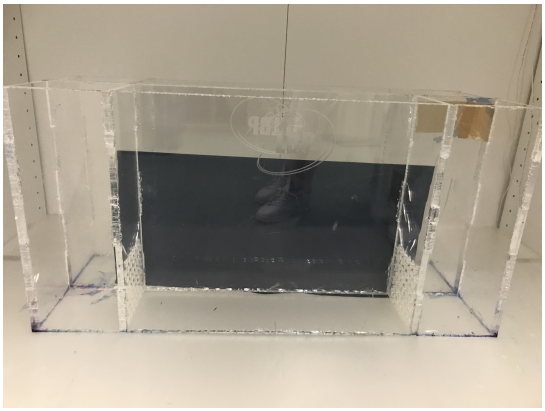


Note. Set-up of the Elevated Plus Maze. The set-up contains two open and two arms enclosed by high walls and a removable ceiling. Images extracted from protocol Nr: 2015_026.

plexiglas walls with small holes to allow the mice to see and smell each other. A virtual zone (2.5 cm in front of the separation walls) was defined as approach zone, and entries or time spent in these zones were scored for social approach behaviour. Specific mice were bred solely to use as stranger mice for all SPSN experiments. The mice were placed in the experimental room to acclimate for at least half an hour prior to the experiment. The protocol consisted of 3 phases: habituation phase, sociability (S) phase and preference for social novelty (PSN) phase. During the habituation phase, the tested mouse was placed in the middle chamber for 5 min. After 5 min, during the S phase, one stranger mouse of the same sex (stranger 1, S1) was placed in one of the flanking chambers, while the other side remained empty. Whether S1 was placed in the compartment on the left or right side of the tested mouse was alternated to control for potential side preference. Approach and exploration of S1 compared to the empty (E) side was recorded for 10 min. The preference for S1 over E was considered as a measurement of sociability. After 10 min, during the PSN phase, a second stranger mouse of the same sex (S2) was placed in the empty chamber while approach and exploration to the novel mouse were recorded for 10 min. The preference for S2 over S1 was considered as a preference for social novelty and is depending on social memory. In between every tested mouse the arena was cleaned using a 70% ethanol solution to reduce olfactory contamination. The behaviour was recorded using a webcam and ANY-maze software. This software tracked the path length and the time spent by each compartment.

Figure 9

Set-up of the SPSN.



Note. Set-up of the Sociability/Preference for Social Novelty task. A middle chamber is flanked by two smaller chambers that are connected by a wall with 52 x 8 mm round holes. Image extracted from protocol Nr: 2015_007.

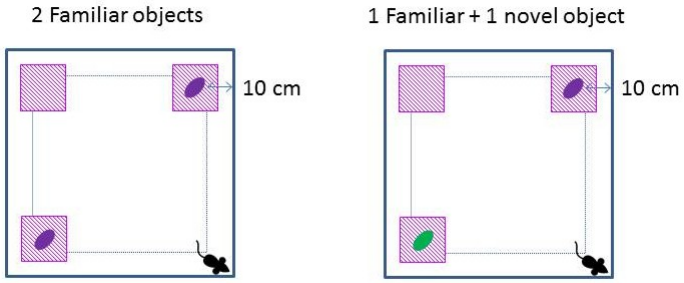
2.5.4 Novel Object Recognition (NOR)

The Novel Object Recognition was a protocol used to test episodic memory and learning. It relied on the innate behaviour of mice to explore novel objects and to remember familiar objects. The NOR task utilised a dimly lit version of the arena used in the OF protocol. The NOR task consisted of three phases with a duration of 10 min each: habituation, training and testing. The task was performed over three consecutive days, with one day for each phase. On the first day, during the habituation phase, the mouse was placed in the empty arena to track exploration. On the second day, during the training phase, two of the same objects (50 ml Falcon tubes filled with darkly coloured liquid) were presented approximately 10 cm from the walls in opposing corners of the arena (see Figure 10). The arena was cleaned using a soft tissue in between the mice, the objects were cleaned using a 70% ethanol solution to avoid olfactory contamination. On the last day, during the testing phase, a novel object (Lego blocks) and one of the already familiar object (Falcon tubes) were placed in the same two opposing corners as in the training phase. The placement of the novel and familiar object were counterbalanced across the mice to prevent their placement from influencing any results. The behaviour was recorded using the same ANY-maze software as was used for the OF. This software tracked the time spent by each object and total path length. Path length was used as a measurement of general locomotion. The time spent by each object was used as a measurement of memory.

Figure 10

Set-up of the NOR.

A. B.



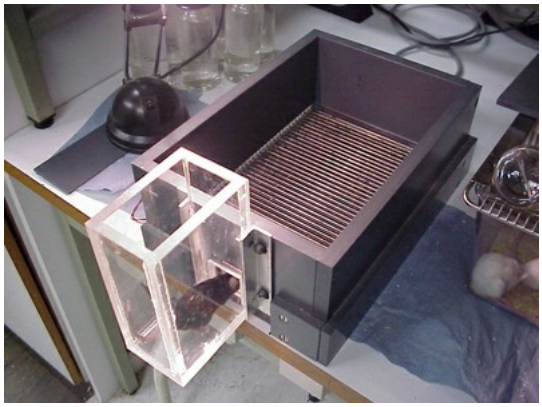
Note. Set-up of the Novel Object Recognition task. Objects are placed in the two small pink virtual zones. The objects are close to the wall but allowing a 10 cm corridor for the mouse to pass. During the training phase two of the same objects are used (50 ml Falcon tubes filled with darkly coloured liquid, panel A). During the testing phase one of the familiar objects and one novel objects was used (Lego blocks, panel B). Image extracted from protocol Nr: 2015_014.

2.5.5 Passive Avoidance (PA)

In the Passive Avoidance protocol a specific context was paired with an unpleasant foot shock and was therefore used to measure the ability for associative learning. It relied on the innate behaviour of mice to go to a darker environment because they prefer darkness over a well-lit environment. The PA task arena consisted of a box with two compartments (see Figure 11). A brightly lit plexiglas compartment shared a small door with a darkened compartment which was equipped with a conductive grid floor. This experiment was performed over two consecutively days. On day one the animals were fear conditioned, and on day two memory was tested. The procedure was as follows: Half an hour prior to the experiment, the animals were placed in a dark environment to dark habituate. After these 30 min, the mouse was placed in the lit compartment with their snout away from the door. After 5 s the door was opened, allowing the mouse to enter the dark compartment. As soon as the mouse touched the grid with its four paws, the door was closed and an electric shock (0.5 mA) was delivered for 2 s. Latency to enter was recorded manually. After 20 s, the mouse was removed from the box and was placed back in its cage. The same procedure was repeated on the second day, but no shock was delivered. If the mouse remembered the foot shock from the day before, they would be more hesitant to enter the darkened compartment again. A cut-off period of 300 s was used, in case the mouse did not enter the dark compartment at all. The time to

Figure 11

Set-up of the PA.



Note. Set-up of the Passive Avoidance task. The arena a box with two compartments. A brightly lit plexiglas compartment shared a small door with a darkened compartment which was equipped with a conductive grid floor, allowing the experimenter to give a foot shock (0.5 mA) to the mouse. Image extracted from protocol Nr: 2015_022.

enter the dark compartment was recorded on both days and was used as a measurement of memory.

2.6 Statistical analysis

All statistical analyses were done using Gnu PSPP software version 1.4.1. Differences within and between conditions were analysed using independent t-tests and two-way analysis of variance (ANOVA). All analyses were done with a statistical significance of $\alpha = 0.05$.

3 Results

3.1 Shape analysis

In all the groups (TS treated, TS untreated, WT treated and WT untreated) there were some mice left out of the analysis due to artifacts and incomplete reconstructions (see Table 3 for a complete overview).

A) PD3

At birth (PD3) no difference in mandible shape were found with the PCA, as indicated by the overlapping groups in the scatterplot with PC1 and PC2, which together explain 32.35% of

Table 3

Number of Mice Left Out of the Shape Analysis.

	PD3		PD14		PD29		W28	
	WT	TS	WT	TS	WT	TS	WT	TS
Treated	6	9	8	1	1	1	0	0
Untreated	1	2	3	5	4	1	1	0
Total	7	11	11	6	5	2	1	0

Note. The number of mice that are left out of the shape analysis is mentioned for the analysis of post-natal day 3 (PD3), PD14, PD29 and post-natal week 28 (W28). Genotype is defined as wild type (WT) or trisomic Ts65Dn (TS). Animals were left out due to artifacts or incomplete reconstructions.

the variation (see Figure 12A). Even though the PCA showed no differentiation, permutation tests revealed significant differences between the groups. There was an effect of genotype, the TS treated and untreated mice were significantly different from the WT treated and untreated mice (see Table A2 for all the p-values). In the TS mice there was an effect of treatment, the TS treated mice were significantly different from TS untreated mice ($p = .014$). However, it looks like the effect of treatment did not rescue the mandible shape, since TS treated mice are still significantly different from the WT treated ($p < .001$) and WT untreated mice ($p < .001$). In the WT mice, on the other hand, there seems to be no effect of treatment, showing no significant differences between WT treated and WT untreated mice ($p = .180$).

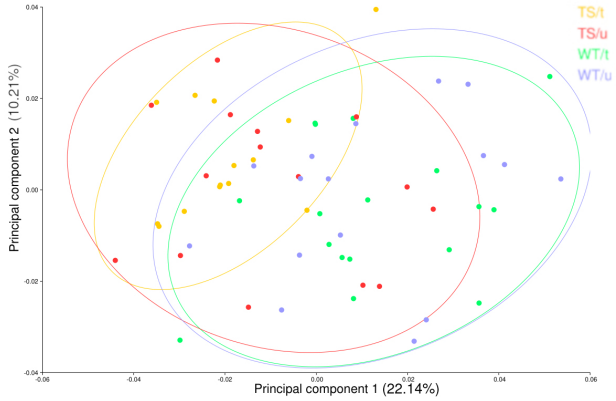
B) PD14

At PD14 there were still no differences found with the PCA, as the groups still mainly overlapped in the scatterplot with PC1 and PC2 which together explain 38.68% of the variation (see Figure 12B). However, as at PD3, the permutation tests showed significant differences in the groups (see Table A2 for all the p-values). The significantly differences between genotypes, can be seen again at PD14, with the TS treated and untreated mice being statistical significant different from the WT treated and untreated mice. However, unlike at PD3, this time it appears that there was no effect of treatment in the TS mice, but in the WT mice. There was a significant difference between the WT treated and the WT untreated mice ($p = .004$). The PCA showed the WT treated mice were moving closer to the TS mice and further away from the WT untreated mice on PC1, suggesting a negative effect of the treatment.

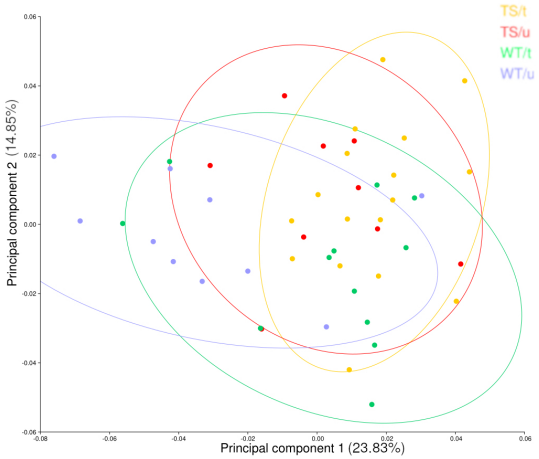
Figure 12

Results of the Principal Component Analysis.

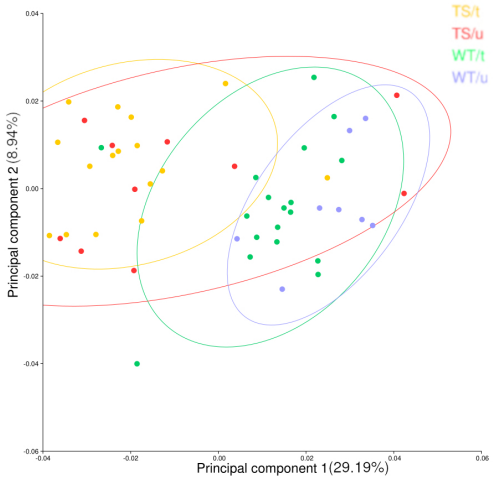
A. PCA of PD3



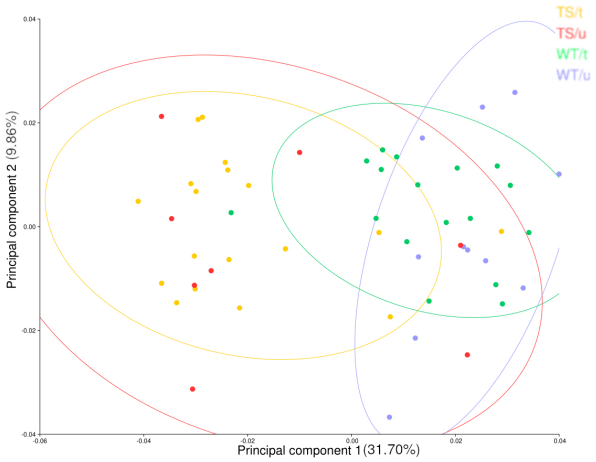
B. PCA of PD14



C. PCA of PD29



D. PCA of W28



Note. The results of the PCA are based on the 22 Landmarks on the mandible for each mouse in different development stages. The PCA was performed using a covariance matrix generated from the Procustes distances. The four different colors represent the different groups: trisomic Ts65Dn treated (TS/t), TS untreated (TS/u), wild type treated (WT/t) and wild type untreated (WT/u). Panel A shows the PCA results of post-natal day 3 (PD3), panel B shows the results of PD14, panel C shows the results of PD29 and panel D shows the results of post-natal week 28 (W28).

C) PD29

At PD29 a differentiation between TS and WT mice started to appear, as can be seen in the scatterplot of PC1 and PC2, which together explain 38.94% of the variation (see Figure 12C). This differentiation was mostly confirmed by the permutation tests. Once again, the TS treated and untreated mice were significantly different from the WT treated and untreated mice (see Table A2 for all the p-values). Like at PD14, it appears there was no effect of treatment in the TS mice. In the WT mice, there was a significant difference between the WT treated and WT untreated mice ($p = .007$). As at PD14 the PCA showed the WT treated mice were moving closer to the TS mice and further away from the WT untreated mice on PC1. This is, again, suggesting a negative effect of the treatment.

D) W28

At W28 the differentiation between TS and WT mice is seen again in the scatterplot of PC1 and PC2, which together explain 41.56% of the variation (see Figure 12D). The permutation tests showed significant differences between the four groups (TS treated, TS untreated, WT treated and WT untreated; see Table A2 for all the p-values). However, there is no significant difference between TS treated and untreated mice ($p = .336$), nor between the WT treated and untreated mice ($p = .205$). The significant differences between WT treated and TS untreated mice ($p < .001$), suggest that the negative effect we were seeing at PD14 and PD29, did not change the phenotype enough to make them TS-like at W28.

3.2 Behavioural testing

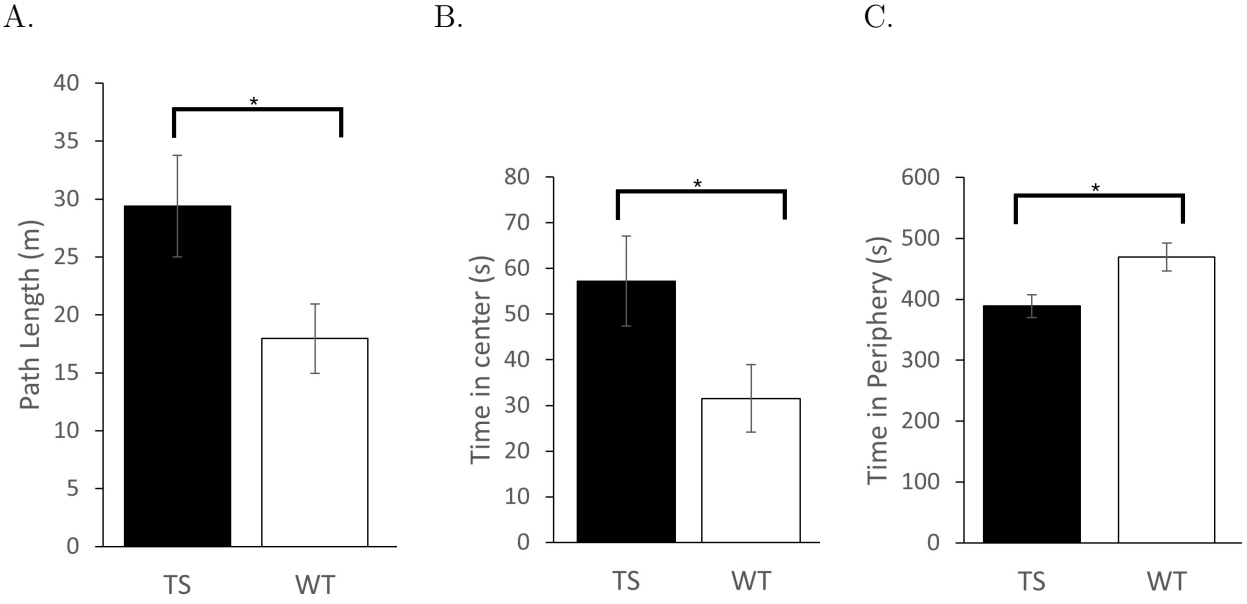
3.2.1 Open field (OF)

A) Testing at week 20

In the first round of behavioural testing (W20) significant differences were found between TS and WT when comparing free locomotion and exploration behaviour using the OF task (see Figure 13). An independent t-test for path length showed a significant difference [$t(20) = 2.21$, $p = .039$] between TS ($M = 29.4$, $SEM = 4.38$) and WT ($M = 17.9$, $SEM = 3.00$). Since the TS mice were having a longer path length, this suggest TS mice showed more locomotory activity. The TS mice spent more time in the center ($M_{TS} = 57.20$, $SEM_{TS} = 9.83$, $M_{WT} = 31.52$, $SEM_{WT} = 7.38$), this was also statistically significant [$t(20) = 2.13$, $p = .046$]. Time spent in the periphery ($M_{TS} = 388.76$, $SEM_{TS} = 18.91$, $M_{WT} = 469.39$, $SEM_{WT} = 23.29$) was also significant different [$t(20) = -2.62$, $p = .017$], with the TS mice spending less time in the periphery. The reduced time spent in the periphery and the increased time in the

Figure 13

Results of the Open Field task at W20.



Note. Exploration on the open field of trisomic (TS, black bars, n =10) or wild type (WT, white bars, n = 12) animals. Analysis of path length (panel A), time in center (panel B) and time in the periphery (panel C) showed significant differences between the TS and WT animals. Data is presented as means ± SEM. (Comparison by independent t-test). * $p < .05$

center in combination with the longer path length suggests the TS experienced less anxiety, as well as more general exploration than the WT mice.

B) Testing at week 30

Unlike at W20, in the second round of behavioural testing (W30) no significant differences were found between TS and WT for the OF task (see Figure 14). An independent t-test showed no significant differences between TS and WT mice for path length [$t(215) = 1.56$, $p = .140$], nor for the time spent in the center [$t(8) = 1.81$, $p = .107$] or periphery [$t(11) = 11.136$, $p = .189$]. This suggest that there is no effect of genotype.

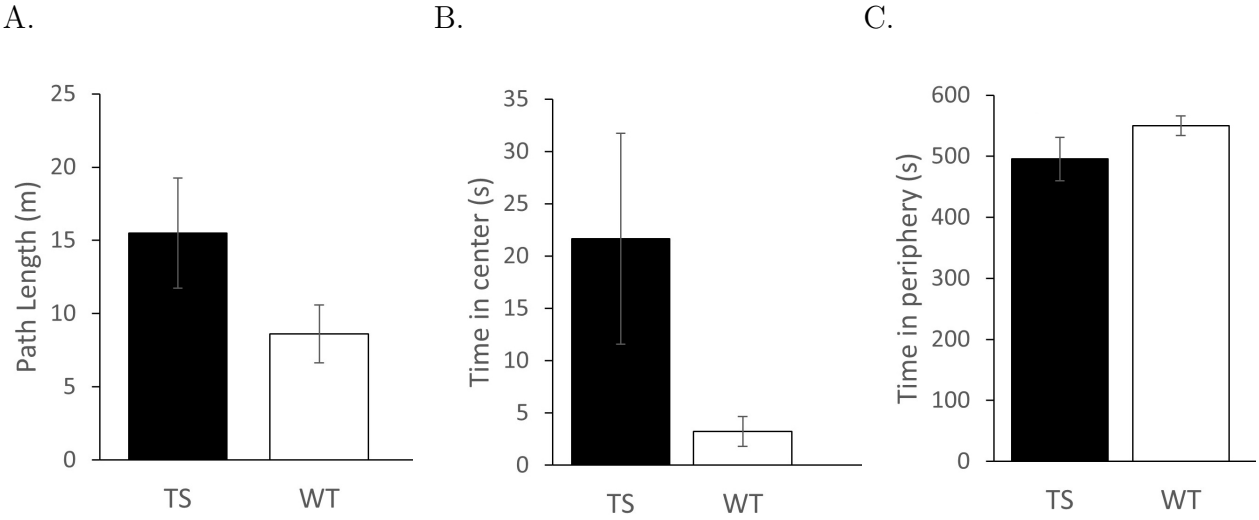
3.2.2 Elevated Plus Maze (EPM)

A) Testing at week 20

In the first round of testing (W20) no significant differences were found between TS and WT when comparing exploration behaviour in the EPM (see Figure 15). We did not observe significant differences in the total number of arm visits [$t(19) = 0.59$, $p = .563$] nor did we observe a significant difference in the percentage of open arm visits [$t(19) = 0.24$, $p = .816$].

Figure 14

Results of the Open Field task at W30.



Note. Exploration in the open field of trisomic (TS, black bars, n =9) and wild type (WT, white bars, n = 8) animals. Analysis of path length (panel A), time in center (panel B) and time in the periphery (panel C) showed no significant differences between the TS and WT animals. Data is presented as means ± SEM. (Comparison by independent t-test).

These results suggest no difference in anxiety levels between TS and WT mice.

B) Testing at week 30

As at W20, in the second round of behavioural testing (W30) no significant differences were found between TS and WT in the EPM (see Figure 16). We did not observe differences in the total number of arm visits [$t(15) = 1.49$, $p = .156$] nor in the percentage of open arm visits [$t(15) = 0.08$, $p = .936$]. This suggest no difference in anxiety levels between genotypes.

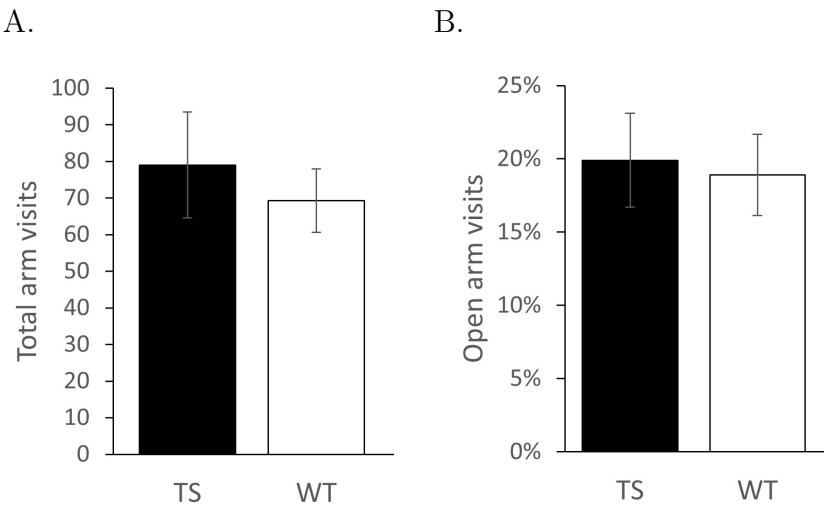
3.2.3 Sociability/preference for social novelty test (SPSN)

A) Testing at week 20

Sociability and social recognition were measured using the SPSN test (see Figure 17). There was no significant interaction effect [$F(2,40) = 1.79$, $p = .180$] in path length between phase and genotype. We did find evidence for a main effect of phase [$F(2,40) = 11.17$, $p < .001$], with the distance travelled being reduced in the S1 and S2 phase. We found no evidence for a main effect of genotype [$F(1,20) = 1.59$, $p = .223$], suggesting both groups were equally active. An independent t-test of the path length during the habituation phase showed no significant differences [$t(20) = 0.25$, $p = .807$] between TS and WT mice in the distance travelled.

Figure 15

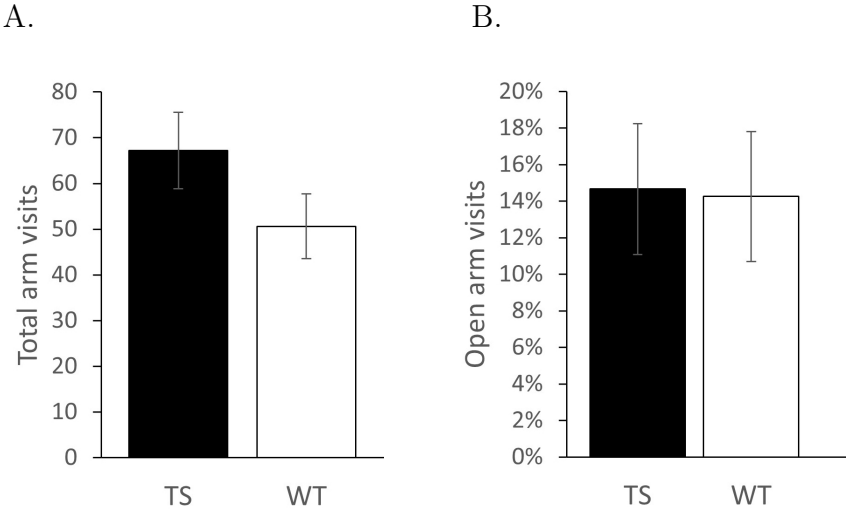
Results of the Elevated Plus Maze at W20.



Note. Elevated plus maze activity as measured by infrared beams of trisomic (TS, black bars, n =10) and wild type (WT, white bars, n = 11) animals. Total number of arm visits (panel A) and percentage of open arm visits (panel B) showed no significant differences between the TS and WT animals. Data is presented as means ± SEM. (Comparison by independent t-test).

Figure 16

Results of the Elevated Plus Maze at W30.

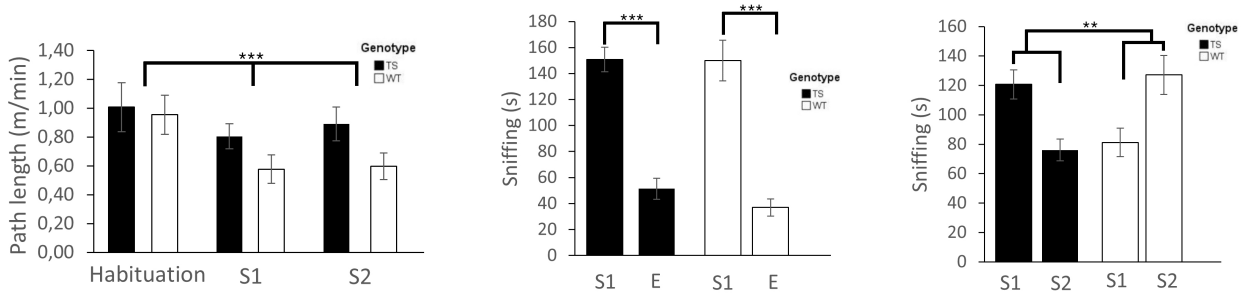


Note. Elevated plus maze activity as measured by infrared beams of trisomic (TS, black bars, n =9) and wild type (WT, white bars, n = 8) animals. Total number of arm visits (panel A) and percentage of open arm visits (panel B) showed no significant differences between the TS and WT animals. Data is presented as means ± SEM. (Comparison by independent t-test).

Figure 17

Results of the Sociability/Preference for Social Novelty Test at W20.

A. B. C.



Note. Sociability and social memory is measured for the trisomic (TS, black bars, n =10) and the wild type (WT, white bars, n = 12) mice. Path length per min (panel A) showed all animals were equally active and walked more in the habituation phase. The sociability phase (panel B) showed that both groups preferred spending time with the mouse over the empty side, showing no difference between the two groups. The social novelty phase (panel C) showed TS mice preferred spending time with S1, while WT mice preferred spending time with S2. Data is presented as means ± SEM. (Comparison by a two-way ANOVA). ** p < .01, *** p < .001

When checking for sociability (S1 phase), no significant interaction effect between side and genotype was found [$F(1,20) = 0.25$, $p = .623$]. We found evidence for a main effect of side [$F(1,20) = 62.55$, $p < .001$], with both groups spending more time with the new stranger mouse (S1) than the empty side. However, we did not find evidence for a main effect of genotype [$F(1,20) = 1.02$, $p = .324$]. These results suggest TS and WT showed the same amounts of sociability behaviour, with both groups showing interest in the new animal.

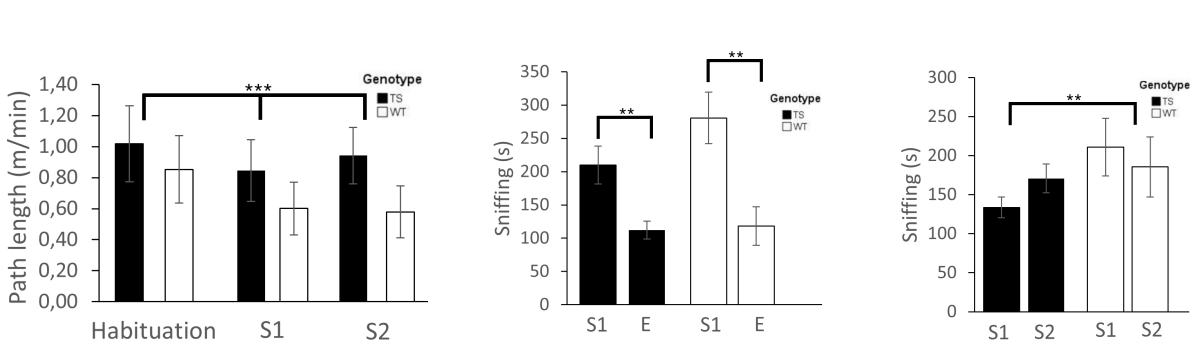
When checking for social memory (S2 phase), a significant interaction effect [$F(1,20) = 11.95$, $p = .002$] was observed, with the TS mice spending more time with S1 and the WT mice spending more time with S2. We found no evidence for a main effect of side [$F(1,20) = 0.135$, $p = .717$], nor for a main effect of genotype [$F(1,20) = 0.624$, $p = .439$]. These result suggest the TS mice showed no social memory for the S1 mouse.

B) Testing at week 30

The mice were tested again for sociability and social memory at W30 (see Figure 18). There was no significant interaction effect [$F(2,26) = 1.293$, $p = .291$] in path length between phase and genotype. We did find evidence for a main effect of phase [$F(2,26) = 6.69$, $p = .005$], with the distance travelled being reduced in the S1 and S2 phase. We found no evidence for a main effect of genotype [$F(1,13) = 0.88$, $p = .365$], suggesting both groups were equally active. An independent t-test of the path length during the habituation phase showed no

Figure 18

Results of the Sociability/Preference for Social Novelty Test at W30.



Note. Sociability and social memory is measured for the trisomic (TS, black bars, n = 9) and the wild type (WT, white bars, n = 8) mice. Path length per min (panel A) showed all animals were equally active and walked more in the habituation phase. The sociability phase (panel B) showed that both groups preferred spending time with the mouse over the empty side, showing no difference between the two groups. The social novelty phase (panel C) showed TS mice spent less time than the WT mice with both S1 and S2. Data is presented as means ± SEM. (Comparison by a two-way ANOVA). ** p < .01, *** p < .001

significant differences [$t(15) = 0.43$, $p = .674$] between TS and WT mice in the distance travelled.

When checking for sociability (S1 phase), no significant interaction effect between side and genotype was found [$F(1,13) = 0.819$, $p = .382$]. We found evidence for a main effect of side [$F(1,13) = 13.45$, $p = .003$], with both groups spending more time with the new stranger mouse (S1) than the empty side. However, we did not find evidence for a main effect of genotype [$F(1,13) = 4.57$, $p = .052$]. These results suggest TS and WT showed the same amounts of sociability behaviour, with both groups showing interest in the new animal.

When checking for social memory (S2 phase), no significant interaction effect [$F(1,15) = 0.74$, $p = .405$] was observed. We found no evidence for a main effect of side [$F(1,15) = 0.03$, $p = .872$]. However we did find evidence for a main effect of genotype [$F(1,15) = 9.537$, $p = .007$], with the TS mice spending less time than the WT mice with both S1 and S2.

3.2.4 Novel Object Recognition (NOR)

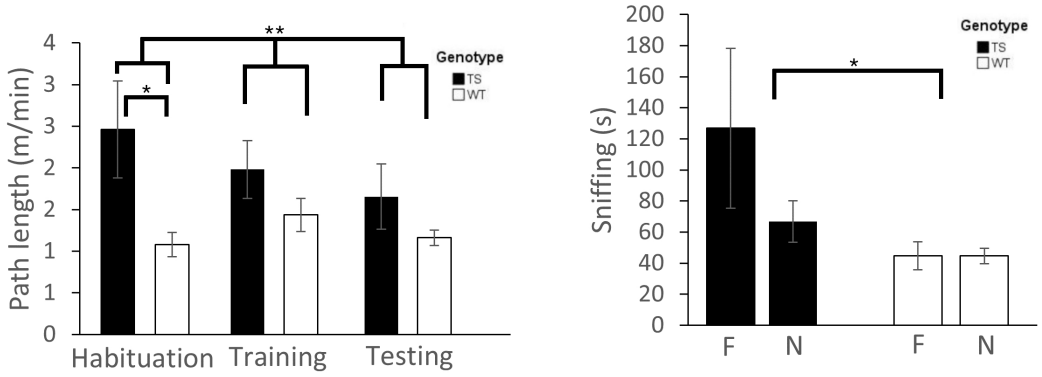
A) Testing at week 20

General memory impairment was tested in the novel object recognition test using a two-way ANOVA (see Figure 19). To ensure that the animals form an information trace of the object during the training phase, animals should at least spend 15 seconds exploring both objects

Figure 19

Results of the Novel Object Recognition Test at W20.

A. B.



Note. General memory impairment was tested in the novel object recognition test for the trisomic (TS, black bars, n =10) and the wild type (WT, white bars, n = 12) mice. Path length per min (panel A) showed a significant interaction between phase and genotype. Panel B shows the time spent sniffing the familiar object (F) and the novel object (N). The testing phase showed that TS mice spent more time sniffing both objects, than the WT mice. Data is presented as means ± SEM. (Comparison by a two-way ANOVA). * p < .05, ** p < .01

together. All animals met this criteria. There was a significant interaction effect [$F(2,40) = 5.25$, $p = .009$] for path length between phase and genotype. We found no evidence for a main effect of phase [$F(2,40) = 3.15$, $p = .054$], nor for a main effect of genotype [$F(1,20) = 3.93$, $p = .061$], suggesting both groups were equally active and there was no difference in walking in the different phases. An independent t-test of the path length during the habituation phase showed a significant difference [$t(20) = 2.50$, $p = .021$] between TS and WT mice in the distance travelled.

To evaluate memory, sniffing time for both objects was compared during the testing phase using a two-way ANOVA. There was no significant interaction effect [$F(1,20) = 1.25$, $p = .276$] for sniffing time between the objects and genotype. We found no evidence for a main effect of object [$F(1,20) = 1.05$, $p = .318$]. However, we did find a main effect of genotype [$F(1,20) = 5.32$, $p = .032$], with the TS mice sniffing longer on both objects than the WT mice. These results suggest there was no memory for the familiar object in both groups.

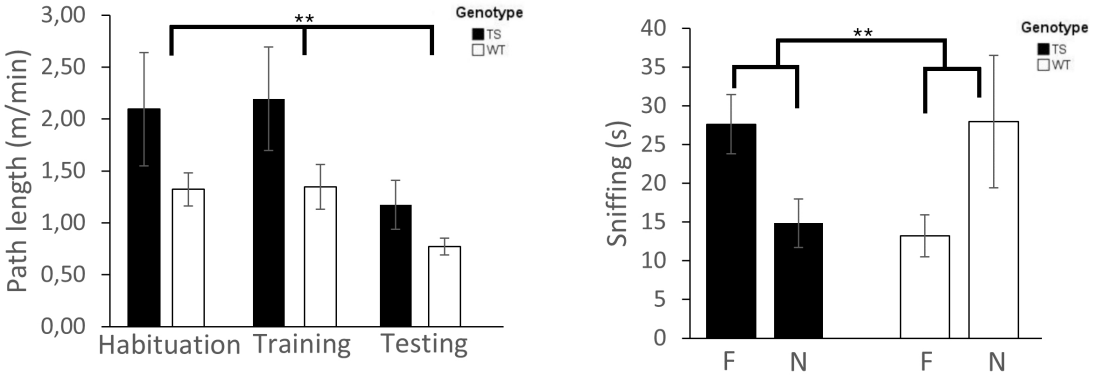
B) Testing at week 30

General memory impairment was tested in the novel object recognition test using a two-way ANOVA (see Figure 20). To ensure that the animals form an information trace of the object during the training phase, animals should at least spend 15 seconds exploring both objects

Figure 20

Results of the Novel Object Recognition Test at W30.

A. B.



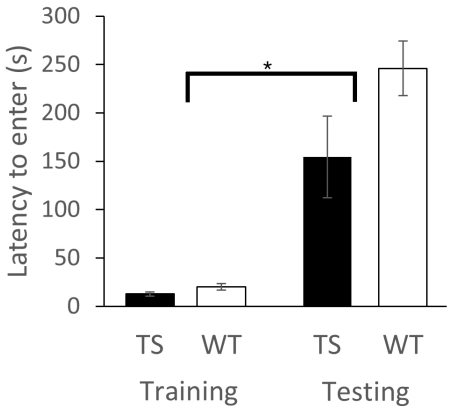
Note. General memory impairment was tested in the novel object recognition test for the trisomic (TS, black bars, n =9) and the wild type (WT, white bars, n = 7) mice. Path length per min (panel A) showed all animals were equally active and walked more in the habituation phase. Panel B shows the time spent sniffing the familiar object (F) and the novel object (N). The testing phase showed that there was a significant interaction between object and genotype. Data is presented as means ± SEM. (Comparison by a two-way ANOVA). ** $p < .01$

together. One animal (WT, male) did not meet this criteria and was excluded from the analyses. There was no significant interaction effect [$F(2,28) = 0.65$, $p = .530$] for path length between phase and genotype. We found evidence for a main effect of phase [$F(2,28) = 9.00$, $p = .001$]. However, we did not find evidence for a main effect of genotype [$F(1,14) = 2.03$, $p = .176$], suggesting both groups were equally active and all mice walked more in the habituation phase than in the training and the testing phase. An independent t-test of the path length during the habituation phase showed no significant differences [$t(14) = 1.21$, $p = .247$] between TS and WT mice in the distance travelled.

To evaluate memory, sniffing time for both objects was compared during the testing phase using a two-way ANOVA. There was a significant interaction effect [$F(1,14) = 11.22$, $p = .005$] for sniffing time between the objects and genotype, with the TS mice sniffing the familiar object longer and the WT mice sniffing the new object longer. We found no evidence for a main effect of object [$F(1,14) = 0.03$, $p = .860$], nor did we find a main effect of genotype [$F(1,14) = 0.01$, $p = .908$]. These results suggest there was no memory for the familiar object in the TS mice, however the WT mice seemed to have a memory for the familiar object.

Figure 21

Results of the Passive Avoidance at W20.



Note. Latency to enter (s) the dark compartment of the Passive Avoidance set-up on day one (training) and day two (testing) for the trisomic (TS, black bars, n =10) and wild type (WT, white bars, n = 12) mice. An analysis showed significant differences between the training and the testing phase, however no interaction effect, nor significant differences between genotype were found. Data is presented as means ± SEM. (Comparison by two-way ANOVA). * $p < .05$

3.2.5 Passive Avoidance (PA)

A) Testing at week 20

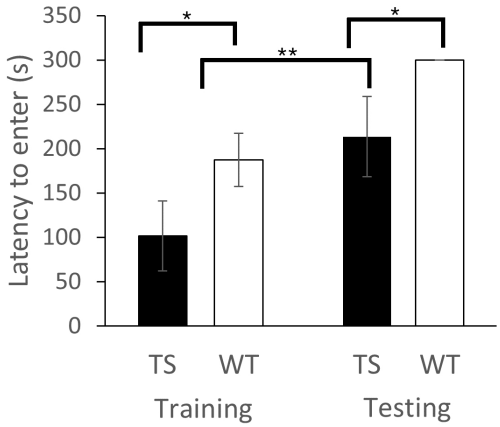
Latency to enter the dark compartment in the first round of behavioural testing was analysed with a two-way ANOVA. There is no evidence found for an interaction effect [$F(1,20) = 2.94$, $p = .102$] between day and genotype (see Figure 21). We observed a significantly difference [$F(1,20) = 55.70$, $p < .001$] between the training phase and the testing phase (main effect of day), all the animals took longer to enter the dark compartment in the testing phase. However, we did not find a main effect of genotype [$F(1,20) = 3.93$, $p = .061$] in the first round of behavioural testing. This is suggesting there was no significant difference in the latency to enter between the TS and the WT mice, and thus there was no significant difference in memory between TS and WT.

B) Testing at week 30

Latency to enter was measured again at W30 (see Figure 22). Like at W20 there is no evidence found for an interaction effect [$F(1,15) = 0.00$, $p = .990$] between day and genotype. We also observed the main effect of day again [$F(1,15) = 18.58$, $p = .001$], with all the animals taking longer to enter the dark compartment on the second day. However, unlike at W20, this time there was also a main effect of genotype found [$F(1,15) = 5.189$, $p = .038$], with the

Figure 22

Results of the Passive Avoidance at W30.



Note. Latency to enter (s) the dark compartment of the Passive Avoidance set-up on day one (training) and day two (testing) for the trisomic (TS, black bars, n =8) and wild type (WT, white bars, n = 9) mice. An analysis showed significant differences between the training and the testing phase. No interaction effect was found. Significant differences between genotype were found. Data is presented as means ± SEM. (Comparison by two-way ANOVA). * $p < .05$, ** $p < .01$

TS mice having shorter latencies to enter than WT mice. This suggest the TS mice always went faster into the dark compartment, an thus had a worse memory than the WT mice.

4 Discussion

The aim of this thesis was to determine if GTE-EGCG treatment starting during embryonic development and continuing throughout adulthood would improve the craniofacial malformation that is typically seen in the Ts65Dn mouse model for DS, as well as to check if we could find differences in the behavioural profile of untreated TS and WT mice. We tested the hypotheses that i) untreated TS mice will have a different mandible shape than the untreated WT mice. ii) EGCG will rescue the mandible shape of the treated TS mice, making them more WT-like. And iii) untreated TS mice will perform worse on the cognitive testing than the untreated WT mice. To examine these hypotheses, a longitudinal study was done in which Ts65Dn mice and euploid littermates were followed-up from embryonic day 0 (E0) until post-natal week 34 (W34).

4.1 The shape analysis showing mixed results for the effect of GTE-EGCG on the development of the mandible

The results of the shape analysis showed mixed results: the effect of the treatment was depending on the developmental stage. At PD3, there was an effect of treatment in the TS mice, making them distinguishable from untreated TS mice. However, this effect did not rescue the mandible shape completely, since the treated TS mice were still different from the treated WT and untreated WT mice. In this age, the treated WT mice showed no effect of treatment. The effect seen in TS mice, in combination with the lack of effect in WT mice, are suggesting that the pharmacological intervention at PD3 reduced the severity in TS animals, but had no detrimental effect in WT animals. This effect was not longer detectable at PD14 and PD29 in TS animals, indicating that the EGCG treatment is no longer affecting the the mandible shape/growth. We did find, however, an effect of treatment in treated WT mice, making them significantly different from untreated WT mice. We can see the treated WT mice moving closer to TS mice on the scatterplot of PC1 and PC2. These results are suggesting there is a negative effect of treatment on the WT mice at this age. These results suggest that the timeline for bone formation is differential in TS and WT animals and EGCG might be contraindicated in non-trisomic individuals at a certain age. At W28, there is no significant difference between the treated and untreated TS mice, indicating that the initial beneficial effect at PD3 was not consistent and only short lived. Similarly the negative effect of EGCG in WT mice detected at PD14 and PD29, was also no longer detectable, showing no significant difference with untreated WT mice. Furthermore, the observed differences in mandible formation between WT and TS animals were still present at W28 indicating that the effect of EGCG was limited to subtle changes at PD3 that were no longer observed at later ages.

Previous research has also found differences between the shape of the mandible in TS and WT mice (Richtsmeier et al., 2000). This research found most of the linear distances between landmarks smaller in Ts65Dn mice, with the biggest differences between the coronoid and the angular process as an endpoint. Further analyses consistently found the Ts65Dn mandible to be significantly smaller than their euploid littermates.

Treatment with EGCG has shown mixed results in previous research. For instance research done by Stringer et al. (2017) has showed detrimental effects in skeleton formation in both TS and WT mice, including a decrease in the cortical thickness, yield force and ultimate stress. As in our study, almost no therapeutic benefits of EGCG treatment was found

in this study, however it is important to note Stringer et al. did not look at mandible shape and used pure EGCG whereas we used GTE-EGCG treatment. Another inportant difference with our study is the age at which the treatment began. Stringer et al. started the treatment in adolescence, whereas we already started the treatment prenatally.

Other research, such as De la Torre et al. (2014) found that GTE-EGCG treatment at a low dose (30 mg/kg/day, which is the same dose as our study) could rescue facial development in TS mice when treatment was started prenatally. However, a high dose (100 mg/kg/day) could have detrimental effects.

These mixed results show that much is still unknown about EGCG and its possible effect in DS. More research should be done, especially about when to give the treatment, at which dose and how long it should be given.

4.2 Behavioural testing showing mixed results in the differences between untreated TS and WT mice

The behavioural comparison was only done in untreated animals, since the cohorts were split between two master theses. Untreated animals were compared in their exploratory behaviour and in their cognitive function. We found some subtle effects between TS and WT animals, taking also the rather small group size and mixed sex in consideration. When compared to WT, TS animals were hyperactive in the Novel Object Recognition, SPSN and Open Field. Furthermore, we observed more time spent in the center in TS animals, while their behaviour in the elevated plus maze was similar to WT. For cognitive performance, TS animals had impaired social recognition, novel object recognition, and impaired passive avoidance compared to WT animals.

4.2.1 Hyperactivity in TS animals

When placed in a novel arena, TS animals show distinct hyperactivity. This was observed in the open field and in the NOR arena at W20 (also a trend in the SPSN). Hyperactivity has been reported earlier in TS65Dn animals and has been linked to arousal and emotionality when placed in a novel environment (Escorihuela et al., 1995; Faizi et al., 2011; Illouz et al., 2019). Indeed, in the SPSN protocol, when presented with stranger mice, animals usually reduce their locomotion due to extend exploration of the stranger mouse, however this was less obvious in TS animals at W20 and at W30, indicating that hyperactivity is very dominant and not affected much by other stimuli. This effect could be the result of increased arousal due to

a monoamine dysfunction (Dekker et al., 2017). Interestingly, we did not observe an increase in arm visits in the EPM, which is also a readout for locomotion and thus hyperactivity, meaning all animals spent the majority of time in the closed arms. This could indicate that the provision of a safe zone (i.e. closed arms in the EPM) would regulate arousal observed in TS animals. In this regard, the increased visits to the centre observed in TS mice in the OF, might be more related to hyperactivity than to reduced anxiety. These findings are in contrast with earlier research that has been conducted on the EPM in Ts65Dn mice, where TS mice spent more time in the open arms as well as more total arm visits (Coussons-Read & Crnic, 1996; Illouz et al., 2019).

4.2.2 Ambiguous results in terms of memory

Memory was assessed in the SPSN, NOR and PA tests. When tested for social memory, animals are presented with a familiar and a novel mouse and usually show a preference for the novel mouse. This innate preference can be tested in the SPSN protocol by presenting a novel stranger mouse (S2) to a test mouse. Indeed at W20, in contrast to WT animals, TS mice showed no preference for S2 indicating impaired social memory. At W30, we could not detect a preference for a novel stranger in either genotype, making a conclusion with regards to the genotype difficult. Previous research has also found impairments of social memory, showing it selective and is dependent on the age of the mice as well as trial length, time between trials and the context (Faizi et al., 2011). The NOR is similar in concept to the SPSN, but instead of stranger mouse, objects were presented to the test mouse (Albasser et al., 2010). NOR protocols are less robust than SPSN because objects are much less salient and attractive to mice than conspecifics. We observed no object recognition memory in WT nor in TS at W20, however, when we repeated the test at W30, WT animals showed a clear preference for a novel object over a familiar object, while TS did not. This impairment in long term memory is in line with results from previous research and has been linked with the changes in the hippocampus of Ts65Dn mice (Garner et al., 2007; Kleschevnikov et al., 2012).

In contrast to SPSN and NOR, the passive avoidance test investigates contextual fear conditioning and avoiding a threatening situation. At W20, both genotypes were clearly learning the association of entering a dark box with a foot shock and showed avoidance behaviour. Conceptually, PA is considered a very salient and easy to learn test. Furthermore, a single exposure to a foot shock creates a very strong memory trace evidenced by the persistence of avoidance behaviour in the mice at W30. However, in TS this memory was less consistent than in WT, indeed the latency to enter the dark box was shorter in TS mice at W30, indicating that they did not remember the association as well. Previous research

showed mixed results in the PA in Ts65Dn mice, significant differences with WT mice were not always found (Day et al., 2019; Holtzman et al., 1996).

4.2.3 No differences in social behaviour

Social behaviour was tested using the SPSN. When testing sociability, animals are presented with a novel mouse (S1) and an empty side. Usually mice show a preference to the side with the novel mouse. Indeed at W20 and W30 both groups showed the same preference for the novel mouse (S1) over the empty side, suggesting there are no differences in sociability between TS and WT mice. These results are in line with previous research (Faizi et al., 2011)

4.3 Limitations

An important limitation of this thesis is the small number of animals we had per condition. The interpretation of the results should therefore be done carefully, since a bigger sample size is needed to confirm the results we found. The smaller group size was due to difficulties breeding Ts65Dn mice. We also lost a number of animals during anaesthesia and scanning. On top of the animals that died, some were also excluded from the shape analysis due to the nature of CT-scans being sensitive to movements from the animals during the scanning procedure. For these reasons it is recommended to breed a higher number of mice in further research.

Even though all the tests were carried out to the best of our abilities, the tests were not performed, nor scored blinded due to the nature of the study. Due to the COVID-19 pandemic the tests of the second round of behaviour testing were not done by the same researcher as the first round. This could possibly have led to small differences in the procedures, and thus might have influenced the results. Moreover, the tests were done in a much shorter time frame than the first round of behavioural testing. Since the mice were being handled more and had less time to recover in between the different tests, this could have led to differences between the first and second round of behavioural testing.

GTE-EGCG was administered in three different ways over the course of their life span: via the mother through means of the placenta, via the mother through means of breastfeeding and via their own drinking water. Research done by Chu et al. (2006) shows that the concentration of catechins was around 10 times higher in the plasma of pregnant rats than in the placenta and around 50-100 times higher than in the fetus. Research done by Isbrucker et al. (2006) shows pups had high plasma EGCG levels at the end of lactation. This could

be because pups consume more quantity of food per body weight than adults and/or because they might have a slower elimination of EGCG. This study also shows pregnant dams consumed more food during the lactation period, which could maybe have led to a greater level of EGCG in the breast milk. Because the treatment was administered to the mother and not the pups itself during the first two stages of treatment, it is impossible to know the real concentration the pups received during this time. Since early development is a critical period for the bone formation, as well as the formation of the brain, this period might be critical to administer the right dose of EGCG (Chai & Maxson Jr, 2006; Dierssen & de Lagrán, 2006). Furthermore, since GTE-EGCG treatment was not given to each mouse individually, but rather to the whole cage by administrating it to the drinking water of the whole cage, it is not possible to say how much exactly each pup got. It is possible some pups might have received more treatment than others.

Lastly the focus of this thesis is on the inhibition of DYRK1A. Past research has suggested DYRK1A has a key role in the phenotypes of DS (Dierssen & de Lagrán, 2006; Dierssen et al., 2009). Nevertheless, its role and mechanisms are still not completely clear. Moreover, since EGCG is known to have several other effects like the prevention of cancer, anti-inflammatory, anti-arthritic and anti-bacterial effects, it is impossible to evaluate its effect in DS as only caused by the inhibition of DYRK1A activity (Pervin et al., 2018; Wolfram, 2007). Especially since we did not measure the DYRK1A levels in the mice, it is important to keep in mind we tested the effect of GTE-EGCG on TS mice, not on the levels of DYRK1A. On top of that, we did not administered pure EGCG but GTE-EGCG, only containing 45% EGCG, meaning there might also have been an effect of the other substances present in green tea. More research should be conducted to understand the mechanisms of DS better.

5 Conclusion

In conclusion the first hypothesis was confirmed, TS untreated mice showed a different mandible shape than the WT untreated mice. For the second hypothesis we only found some confirmation. GTE-EGCG treatment seemed to have an effect on TS mice at PD3, however it was able to completely rescue the mandible shape. In this age there was no (negative) effect of the treatment in WT mice. However, at PD14 and PD29, treatment had a negative effect on the WT mice, making them more TS-like. Moreover, treatment had no (good) effect in the TS mice at PD14 and PD29. At W28, the negative effect on WT mice is not seen any longer. There is still no effect seen on the TS mice. Further research should focus on the timing of the treatment, and with the possibility of having to stop the treatment during adolescence.

We could also not find conclusive evidence for the third hypothesis. TS untreated mice were only significantly different from WT untreated mice in some behavioural tests and at some ages. Further research is required to further progress our knowledge on the behavioural phenotype seen in DS.

6 References

Abeysekera, I., Thomas, J., Georgiadis, T. M., Berman, A. G., Hammond, M. A., Dria, K. J., Wallace, J. M., & Roper, R. J. (2016). Differential effects of epigallocatechin-3-gallate containing supplements on correcting skeletal defects in a down syndrome mouse model. *Molecular nutrition & food research*, *60*(4), 717–726.

Albasser, M. M., Chapman, R. J., Amin, E., Iordanova, M. D., Vann, S. D., & Aggleton, J. P. (2010). New behavioral protocols to extend our knowledge of rodent object recognition memory. *Learning & Memory*, *17*(8), 407–419.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Annus, T., Wilson, L. R., Hong, Y. T., Acosta–Cabronero, J., Fryer, T. D., Cardenas–Blanco, A., Smith, R., Boros, I., Coles, J. P., Aigbirhio, F. I., Menon, D. K., Zaman, S. H., Nestor, P. J., & Holland, A. J. (2016). The pattern of amyloid accumulation in the brains of adults with down syndrome. *Alzheimer's & dementia*, *12*(5), 538–545.

Antonarakis, S. E., Lyle, R., Dermitzakis, E. T., Reymond, A., & Deutsch, S. (2004). Chromosome 21 and down syndrome: From genomics to pathophysiology. *Nature reviews. Genetics*, *5*(10), 725–738.

Arron, J. R., Winslow, M. M., Polleri, A., Chang, C.-P., Wu, H., Gao, X., Neilson, J. R., Chen, L., Heit, J. J., Kim, S. K., Nobuyuki, Y., Miyakawa, T., & Crabtree, G. R. (2006). NFAT dysregulation by increased dosage of DSCR1 and DYRK1A on chromosome 21. *Nature*, *441*(7093), 595–600.

Asim, A., Kumar, A., Muthuswamy, S., Jain, S., & Agarwal, S. (2015). "down syndrome: An insight of the disease". *Journal of biomedical science*, *22*(1), 41–41.

Bain, J., McLauchlan, H., Elliott, M., & Cohen, P. (2003). The specificities of protein kinase inhibitors: An update. *Biochemical journal*, *371*(1), 199–204.

Becker, W., & Sippl, W. (2011). Activation, regulation, and inhibition of dyrk1a. *The FEBS Journal*, *278*(2), 246–256. https://doi.org/10.1111/j.1742-4658.2010.07956.x

Blazek, J. D., Abeysekera, I., Li, J., & Roper, R. J. (2015). Rescue of the abnormal skeletal phenotype in ts65dn down syndrome mice using genetic and therapeutic modulation of trisomic dyrk1a. *Human Molecular Genetics*, *24*(20), 5687–5696. https://doi.org/10.1093/hmg/ddv284

Chai, Y., & Maxson Jr, R. E. (2006). Recent advances in craniofacial morphogenesis. *Developmental Dynamics*, *235*(9), 2353–2375. https://doi.org/10.1002/dvdy.20833

Cheong, J. L., Anderson, P. J., Roberts, G., Burnett, A. C., Lee, K. J., Thompson, D. K., Molloy, C., Wilson-Ching, M., Connelly, A., Seal, M. L., et al. (2013). Contribution of

brain size to iq and educational underperformance in extremely preterm adolescents. *PloS one*, *8*(10), e77475.

Chu, K., Wang, C., Chu, C., Choy, K., Pang, C., & Rogers, M. (2006). Uptake and distribution of catechins in fetal organs following in utero exposure in rats. *Human Reproduction*, *22*(1), 280–287. https://doi.org/10.1093/humrep/del353

Cole, J. H., Annus, T., Wilson, L. R., Remtulla, R., Hong, Y. T., Fryer, T. D., Acosta-Cabronero, J., Cardenas-Blanco, A., Smith, R., Menon, D. K., Zaman, S. H., Nestor, P. J., & Holland, A. J. (2017). Brain-predicted age in down syndrome is associated with beta amyloid deposition and cognitive decline. *Neurobiology of Aging*, *56*, 41–49. https://doi.org/10.1016/j.neurobiolaging.2017.04.006

Contestabile, A., Benfenati, F., & Gasparini, L. (2010). Communication breaks-down: From neurodevelopment defects to cognitive disabilities in down syndrome. *Progress in Neurobiology*, *91*(1), 1–22. https://doi.org/10.1016/j.pneurobio.2010.01.003

Coussons-Read, M. E., & Crnic, L. S. (1996). Behavioral assessment of the ts65dn mouse, a model for down syndrome: Altered behavior in the elevated plus maze and open field. *Behavior genetics*, *26*(1), 7–13.

Day, S. M., Yang, W., Wang, X., Stern, J. E., Zhou, X., Macauley, S. L., & Ma, T. (2019). Glucagon-like peptide-1 cleavage product improves cognitive function in a mouse model of down syndrome. *eNeuro*, *6*(2), ENEURO.0031–19.2019.

De la Torre, R., De Sola, S., Pons, M., Duchon, A., de Lagran, M. M., Farré, M., Fitó, M., Benejam, B., Langohr, K., Rodriguez, J., Pujadas, M., Bizot, J. C., Cuenca, A., Janel, N., Catuara, S., Covas, M. I., Blehaut, H., Herault, Y., Delabar, J. M., & Dierssen, M. (2014). Epigallocatechin-3-gallate, a dyrk1a inhibitor, rescues cognitive deficits in down syndrome mouse models and in humans. *Molecular Nutrition & Food Research*, *58*(2), 278–288. https://doi.org/10.1002/mnfr.201300325

Dekant, W., Fujii, K., Shibata, E., Morita, O., & Shimotoyodome, A. (2017). Safety assessment of green tea based beverages and dried green tea extracts as nutritional supplements. *Toxicology Letters*, *277*, 104–108. https://doi.org/10.1016/j.toxlet.2017.06.008

Dekker, A. D., Vermeiren, Y., Albac, C., Lana-Elola, E., Watson-Scales, S., Gibbins, D., Aerts, T., Van Dam, D., Fisher, E. M., Tybulewicz, V. L., Potier, M.-C., & De Deyn, P. P. (2017). Aging rather than aneuploidy affects monoamine neurotransmitters in brain regions of down syndrome mouse models. *Neurobiology of Disease*, *105*, 235–244. https://doi.org/10.1016/j.nbd.2017.06.007

Delabar, J. M., Aflalo-Rattenbac, R., & Créau, N. (2006). Developmental defects in trisomy 21 and mouse models. *The Scientific World Journal*, *6*, 1945–1964.

Dierssen, M., & de Lagrán, M. M. (2006). Dyrk1a (dual-specificity tyrosine-phosphorylated and -regulated kinase 1a): A gene with dosage effect during development and neurogenesis. *The Scientific World Journal*, *6*, 1911–1922.

Dierssen, M., Herault, Y., & Estivill, X. (2009). Aneuploidy: From a physiological mechanism of variance to down syndrome. *Physiological Reviews*, *89*(3), 887–920.

Down, J. L. H. (1866). Observations on an ethnic classification of idiots. *London Hospital Reports*, (3), 259–262.

Dressler, A., Perelli, V., Feucht, M., & Bargagna, S. (2010). Adaptive behaviour in down syndrome: A cross-sectional study from childhood to adulthood. *Wiener Klinische Wochenschrift*, *122*(23), 673–680.

Duchon, A., & Herault, Y. (2016). Dyrk1a, a dosage-sensitive gene involved in neurodevelopmental disorders, is a target for drug development in down syndrome. *Frontiers in behavioral neuroscience*, *10*, 104–104.

Esbensen, A. J., Seltzer, M. M., & Krauss, M. W. (2008). Stability and change in health, functional abilities, and behavior problems among adults with and without down syndrome. *American Journal on Mental Retardation*, *113*(4), 263–277.

Escorihuela, R. M., Fernández-Teruel, A., Vallina, I. F., Baamonde, C., Lumbreras, M. A., Dierssen, M., Tobeña, A., & Flórez, J. (1995). A behavioral assessment of ts65dn mice: A putative down syndrome model. *Neuroscience Letters*, *199*(2), 143–146. https://doi.org/10.1016/0304-3940(95)12052-6

Ethen, M. K., & Canfield, M. A. (2002). Impact of including elective pregnancy terminations before 20 weeks gestation on birth defect rates: Congenital malformations: Surveillance report. *Teratology (Philadelphia)*, *66*, S32–S35.

Faizi, M., Bader, P. L., Tun, C., Encarnacion, A., Kleschevnikov, A., Belichenko, P., Saw, N., Priestley, M., Tsien, R. W., Mobley, W. C., & Shamloo, M. (2011). Comprehensive behavioral phenotyping of ts65dn mouse model of down syndrome: Activation of b1-adrenergic receptor by xamoterol as a potential cognitive enhancer. *Neurobiology of Disease*, *43*(2), 397–413. https://doi.org/10.1016/j.nbd.2011.04.011

Garner, C. C., Zuniga, E., Malenka, R. C., Nguyen, J., Fernandez, F., Morishita, W., & Blank, M. (2007). Pharmacotherapy for cognitive impairment in a mouse model of down syndrome. *Nature neuroscience*, *10*(4), 411–413.

Glasson, E., Sullivan, S., Hussain, R., Petterson, B., Montgomery, P., & Bittles, A. (2002). The changing survival profile of people with down's syndrome: Implications for genetic counselling. *Clinical genetics*, *62*(5), 390–393.

Grieco, J., Pulsifer, M., Seligsohn, K., Skotko, B., & Schwartz, A. (2015). Down syndrome: Cognitive and behavioral functioning across the lifespan. *American journal of medical genetics. Part C, Seminars in medical genetics*, *169C*(2), 135–149.

Guedj, F., Sébrié, C., Rivals, I., Ledru, A., Paly, E., Bizot, J. C., Smith, D., Rubin, E., Gillet, B., Arbones, M., & Delabar, J. M. (2009). Green tea polyphenols rescue of brain defects induced by overexpression of dyrk1a. *PloS one*, *4*(2), e4606–e4606.

Hamner, T., Udhnani, M. D., Osipowicz, K. Z., & Lee, N. R. (2018). Pediatric brain development in down syndrome: A field in its infancy. *Journal of the International Neuropsychological Society*, *24*(9), 966–976. https://doi.org/10.1017/S1355617718000206

Holtzman, D. M., Santucci, D., Kilbridge, J., Chua-Couzens, J., Fontana, D. J., Daniels, S. E., Johnson, R. M., Chen, K., Sun, Y., Carlson, E., Alleva, E., Epstein, C. J., & Mobley, W. C. (1996). Developmental abnormalities and age-related neurodegeneration in a mouse model of down syndrome. *Proceedings of the National Academy of Sciences - PNAS*, *93*(23), 13333–13338.

Illouz, T., Madar, R., Biragyn, A., & Okun, E. (2019). Restoring microglial and astroglial homeostasis using dna immunization in a down syndrome mouse model. *Brain, Behavior, and Immunity*, *75*, 163–180. https://doi.org/10.1016/j.bbi.2018.10.004

Isbrucker, R., Edwards, J., Wolz, E., Davidovich, A., & Bausch, J. (2006). Safety studies on epigallocatechin gallate (egcg) preparations. part 3: Teratogenicity and reproductive toxicity studies in rats. *Food and Chemical Toxicology*, *44*(5), 651–661. https://doi.org/10.1016/j.fct.2005.11.002

Jarhad, D. B., Mashelkar, K. K., Kim, H.-R., Noh, M., & Jeong, L. S. (2018). Dual-specificity tyrosine phosphorylation-regulated kinase 1a (dyrk1a) inhibitors as potential therapeutics. *Journal of medicinal chemistry*, *61*(22), 9791–9810.

Kent, L., Evans, J., Paul, M., & Sharp, M. (1999). Comorbidity of autistic spectrum disorders in children with down syndrome. *Developmental Medicine & Child Neurology*, *41*(3), 153–158. https://doi.org/10.1017/S001216229900033X

Kleschevnikov, A. M., Belichenko, P. V., Faizi, M., Jacobs, L. F., Htun, K., Shamloo, M., & Mobley, W. C. (2012). Deficits in cognition and synaptic plasticity in a mouse model of down syndrome ameliorated by gabab receptor antagonists. *The Journal of neuroscience*, *32*(27), 9217–9227.

Krinsky-McHale, S. J., Devenny, D. A., Gu, H., Jenkins, E. C., Kittler, P., Murty, V. V., Schupf, N., Scotto, L., Tycko, B., Urv, T. K., et al. (2008). Successful aging in a 70-year-old man with down syndrome: A case study. *Intellectual and Developmental Disabilities*, *46*(3), 215–228.

Max Roser, E. O.-O., & Ritchie, H. (2013). Life expectancy [https://ourworldindata.org/life-expectancy]. *Our World in Data*.

McElyea, S. D., Starbuck, J. M., Tumbleson-Brink, D. M., Harrington, E., Blazek, J. D., Ghoneima, A., Kula, K., & Roper, R. J. (2016). Influence of prenatal egcg treatment and dyrk1a dosage reduction on craniofacial features associated with down syndrome. *Human molecular genetics*, *25*(22), 4856–4869.

Mouse Genome Sequencing Consortium. (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature (London)*, *420*(6915), 520–562.

Myers, B. A., & Pueschel, S. M. (1991). Psychiatric disorders in persons with down syndrome. *The journal of nervous and mental disease*, *179*(10), 609–613.

Patterson, D. (2007). Genetic mechanisms involved in the phenotype of down syndrome. *Mental retardation and developmental disabilities research reviews*, *13*(3), 199–206.

Pervin, M., Unno, K., Ohishi, T., Tanabe, H., Miyoshi, N., & Nakamura, Y. (2018). Beneficial effects of green tea catechins on neurodegenerative diseases. *Molecules (Basel, Switzerland)*, *23*(6), 1297.

Pinter, J., Brown, W., Eliez, S., Schmitt, J., Capone, G., & Reiss, A. (2001). Amygdala and hippocampal volumes in children with down syndrome: A high-resolution mri study. *Neurology*, *56*(7), 972–974.

Rachidi, M., & Lopes, C. (2008). Mental retardation and associated neurological dysfunctions in down syndrome: A consequence of dysregulation in critical chromosome 21 genes and associated molecular pathways. *European Journal of Paediatric Neurology*, *12*(3), 168–182. https://doi.org/10.1016/j.ejpn.2007.08.010

Richtsmeier, J. T., Baxter, L. L., & Reeves, R. H. (2000). Parallels of craniofacial maldevelopment in down syndrome and ts65dn mice. *Developmental dynamics*, *217*(2), 137–145.

Richtsmeier, J. T., Zumwalt, A., Carlson, E. J., Epstein, C. J., & Reeves, R. H. (2002). Craniofacial phenotypes in segmentally trisomic mouse models for down syndrome. *American Journal of Medical Genetics*, *107*(4), 317–324. https://doi.org/10.1002/ajmg.10175

Rodrigues, M., Nunes, J., Figueiredo, S., Martins de Campos, A., & Geraldo, A. F. (2019). Neuroimaging assessment in down syndrome: A pictorial review. *Insights into imaging*, *10*(1), 1–13.

Roizen, N. J., & Patterson, D. (2003). Down's syndrome. *The Lancet*, *361*(9365), 1281–1289. https://doi.org/10.1016/S0140-6736(03)12987-X

Rueda, N., Flórez, J., & Martínez-Cué, C. (2012). Mouse models of down syndrome as a tool to unravel the causes of mental disabilities. *Neural plasticity*, *2012*, 584071–26.

Rushton, J. P., & Rushton, E. W. (2003). Brain size, iq, and racial-group differences: Evidence from musculoskeletal traits. *Intelligence*, *31*(2), 139–155.

Schmidt, C., Moran, T. H., Wohn, A., Davisson, M. T., Reeves, R. H., Irving, N. G., Sisodia, S. S., Bronson, R. T., & Kitt, C. (1995). A mouse model for down syndrome exhibits learning and behaviour deficits. *Nature genetics*, *11*(2), 177–184.

Shukla, D., Bablani, D., Chowdhry, A., Thapar, R., Gupta, P., & Mishra, S. (2014). Dentofacial and cranial changes in down syndrome. *Osong Public Health and Research Perspectives*, *5*(6), 339–344. https://doi.org/10.1016/j.phrp.2014.09.004

Skotko, B. G. (2009). With new prenatal testing, will babies with down syndrome slowly disappear? *Archives of Disease in Childhood*, *94*(11), 823–826. https://doi.org/10.1136/adc.2009.166017

Śmigielska-Kuzia, J., Boćkowski, L., Sobaniec, W., Sendrowski, K., Olchowik, B., Cholewa, M., łukasiewicz, A., & łebkowska, U. (2011). A volumetric magnetic resonance imaging study of brain structures in children with down syndrome. *Neurologia i neurochirurgia polska*, *45*(4), 363–369.

Souchet, B., Duchon, A., Gu, Y., Dairou, J., Chevalier, C., Daubigney, F., Nalesso, V., Créau, N., Yu, Y., Janel, N., et al. (2019). Prenatal treatment with egcg enriched green tea extract rescues gad67 related developmental and cognitive defects in down syndrome mouse models. *Scientific reports*, *9*(1), 1–13.

Steingass, K. J., Chicoine, B., McGuire, D., & Roizen, N. J. (2011). Developmental disabilities grown up: Down syndrome. *Journal of developmental and behavioral pediatrics*, *32*(7), 548–558.

Stringer, M., Abeysekera, I., Thomas, J., LaCombe, J., Stancombe, K., Stewart, R. J., Dria, K. J., Wallace, J. M., Goodlett, C. R., & Roper, R. J. (2017). Epigallocatechin-3-gallate (egcg) consumption in the ts65dn model of down syndrome fails to improve behavioral deficits and is detrimental to skeletal phenotypes. *Physiology & Behavior*, *177*, 230–241. https://doi.org/10.1016/j.physbeh.2017.05.003

Tejedor, F. J., & Hämmerle, B. (2011). Mnb/dyrk1a as a multiple regulator of neuronal development. *The FEBS Journal*, *278*(2), 223–235. https://doi.org/10.1111/j.1742-4658.2010.07954.x

Toyoda, A., Noguchi, H., Taylor, T. D., Ito, T., Pletcher, M. T., Sakaki, Y., Reeves, R. H., & Hattori, M. (2002). Comparative genomic sequence analysis of the human chromosome 21 down syndrome critical region. *Genome research*, *12*(9), 1323–1332.

Vicari, S., Bates, E., Caselli, M. C., Pasqualetti, P., Gagliardi, C., Tonucci, F., & Volterra, V. (2004). Neuropsychological profile of italians with williams syndrome: An exam-

ple of a dissociation between language and cognition? *Journal of the International Neuropsychological Society*, *10*(6), 862–876.

Wegiel, J., Gong, C.-X., & Hwang, Y.-W. (2011). The role of dyrk1a in neurodegenerative diseases. *The FEBS Journal*, *278*(2), 236–245. https://doi.org/10.1111/j.1742-4658.2010.07955.x

Willerman, L., Schultz, R., Neal Rutledge, J., & Bigler, E. D. (1991). In vivo brain size and intelligence. *Intelligence*, *15*(2), 223–228. https://doi.org/10.1016/0160-2896(91)90031-8

Wiseman, F. K., Alford, K. A., Tybulewicz, V. L., & Fisher, E. M. (2009). Down syndrome—recent progress and future prospects. *Human molecular genetics*, *18*(R1), R75–R83.

Wolfram, S. (2007). Effects of green tea and egcg on cardiovascular and metabolic health. *Journal of the American College of Nutrition*, *26*(4), 373S–388S.

Zigman, W. B. (2013). Atypical aging in down syndrome. *Developmental disabilities research reviews*, *18*(1), 51–67.

7 Appendix

Table A1

Anatomical Definition of Mandible Landmarks.

Landmark	Anatomical definition
1	Tip of the coronoid process (left)
2	Tip of the coronoid process (left)
3	Posterior-most point on mandibular condyle (left)
4	Mandible angle (left)
5	Superior-most point on inferior border of mandibular ramus (left)
6	Inferior-most point on border of ramus inferior to incisor alveolar (left)
7	Inferior-most point on border of ramus inferior to incisor alveolar (left)
8	Superior-most point on incisor alveolar rim (left)
9	Superior-most point on incisor alveolar rim (left)
10	Anterior point on molar alveolar rim (left)
11	Intersection of molar alveolar rim and base of coronoid process (left)
12	Tip of the coronoid process (right)
13	Anterior-most point on mandibular condyle (right)
14	Posterior-most point on mandibular condyle (right)
15	Mandible angle (right)
16	Superior-most point on inferior border of mandibular ramus (right)
17	Inferior-most point on border of ramus inferior to incisor alveolar (right)
18	Inferior-most point on incisor alveolar rim (right)
19	Inferior-most point on incisor alveolar rim (right)
20	Mandibular foramen (right)
21	Anterior point on molar alveolar rim (right)
22	Intersection of molar alveolar rim and base of coronoid process (right)

Note. The placement of the landmarks can be seen in Figure 3. This set of landmarks was previously determined by the Richtsmeier laboratory and can be found on: https://getahead.la.psu.edu/landmarks/.

Table A2

P-values of the Permutation Tests.

	PD3		
	TS/t	TS/u	WT/t
TS/u	.0137*		
WT/t	<.001***	.0026**	
WT/u	<.001***	.0034**	.1799
	PD14		
	TS/t	TS/u	WT/t
TS/u	.2837		
WT/t	.0026**	.0308*	
WT/u	<.001***	.0076**	.0044**
	PD29		
	TS/t	TS/u	WT/t
TS/u	.1993		
WT/t	<.001***	.0058**	
WT/u	<.001***	.0093**	.0066**
	W28		
	TS/t	TS/u	WT/t
TS/u	.3355		
WT/t	<.001***	< .001***	
WT/u	<.001***	< .001***	.2049

Note. P-values for the permutation tests (10.000 permutations) for the Procrustes distances among groups for the mandible shape. The different groups are defined as trisomic Ts65Dn treated (TS/t), TS untreated (TS/u), wild type treated (WT/t) and WT untreated (WT/u).

* p < .05. ** p <.01. *** p<.001.