Mouse models of neurodegenerative diseases: criteria and general methodology

Janus, C; Welzl, H

Abstract: The major symptom of Alzheimer’s disease is rapidly progressing dementia, coinciding with the formation of amyloid and tau deposits in the central nervous system, and neuronal death. At present familial cases of dementias provide the most promising foundation for modelling neurodegeneration. We describe the mnemonic and other major behavioral symptoms of tauopathies, briefly outline the genetics underlying familiar cases and discuss the arising implications for modelling the disease in mostly transgenic mouse lines. We then depict to what degree the most recent mouse models replicate pathological and cognitive characteristics observed in patients. There is no universally valid behavioral test battery to evaluate mouse models. The selection of individual tests depends on the behavioral and/or memory system in focus, the type of a model and how well it replicates the pathology of a disease and the amount of control over the genetic background of the mouse model. However it is possible to provide guidelines and criteria for modelling the neurodegeneration, setting up the experiments and choosing relevant tests. One should not adopt a “one (trans)gene, one disease” interpretation, but should try to understand how the mouse genome copes with the protein expression of the transgene in question. Further, it is not possible to recommend some mouse models over others since each model is valuable within its own constraints, and the way experiments are performed often reflects the idiosyncratic reality of specific laboratories. Our purpose is to improve bridging molecular and behavioural approaches in translational research.

DOI: https://doi.org/10.1007/978-1-60761-058-8_19

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: https://doi.org/10.5167/uzh-43160
Accepted Version

Originally published at:
DOI: https://doi.org/10.1007/978-1-60761-058-8_19
A man is but what he knoweth, Sir Francis Bacon (1).

Abstract
The major symptom of Alzheimer’s disease is rapidly progressing dementia, coinciding with the formation of amyloid and tau deposits in the central nervous system, and neuronal death. At present familial cases of dementias provide the most promising foundation for modelling neurodegeneration. We describe the mnemonic and other major behavioral symptoms of tauopathies, briefly outline the genetics underlying familiar cases and discuss the arising implications for modelling the disease in mostly transgenic mouse lines. We then depict to what degree the most recent mouse models replicate pathological and cognitive characteristics observed in patients. There is no universally valid behavioral test battery to evaluate mouse models. The selection of individual tests depends on the behavioral and/or memory system in focus, the type of a model and how well it replicates the pathology of a disease and the amount of control over the genetic background of the mouse model. However it is possible to provide guidelines and criteria for modelling the neurodegeneration, setting up the experiments and choosing relevant tests. One should not adopt a “one (trans)gene, one disease” interpretation, but should try to understand how the mouse genome copes with the protein expression of the transgene in question. Further, it is not possible to recommend some mouse models over others since each model is valuable within its own constraints, and the way experiments are performed often reflects the idiosyncratic reality of specific laboratories. Our purpose is to improve bridging molecular and behavioural approaches in translational research.

Key words: Tauopathy, Alzheimer’s disease, transgenic models, phenotype, behavioural tests.

Abbreviations:
AD Alzheimer’s disease
Aβ-amyloid peptide
FAD familial Alzheimer’s disease
FTD fronto-temporal dementia
FTDP-17 fronto-temporal dementia with parkinsonism linked to chromosome 17
GLM general linear model
MAPT microtubule-associated protein Tau
MCI mild cognitive impairment
NFT neurofibrillary tangles

1. Introduction
One hundred years ago Alois Alzheimer described in a seminal paper the behavioural symptoms of his patient, Auguste D., who suffered from a mental illness (2), (see (3) for English translation of the original paper). He observed that “[S]he developed a rapid loss of memory. She was disoriented in her home, [. . .] She is completely disoriented in time and space. Her memory is seriously impaired. If objects are shown to her, she names them correctly, but almost immediately afterwards she has forgotten everything.” After the patient died, an autopsy revealed dense deposits outside and around nerve cells and twisted
strands of fibre inside dead neurons in her brain. Today these two pathological hallmarks of Alzheimer’s disease (AD) are known to be extra-cellular plaques made largely of amyloid peptide (Aβ) and intracellular neurofibrillary tangles composed of hyper-phosphorylated microtubule-associated protein tau (4, 5).

The current view is that almost all neurodegenerative disorders can be broadly classified as disorders of protein folding (6). The accumulation of misfolded proteins may be, as in the case of AD, intra- and extra-cellular, or only intracellular, with abnormally phosphorylated protein tau aggregates being most common. The most common is tau protein which aggregates within neurons in hyper-phosphorylated form (7, 8) and causes profound loss of neurons and atrophy of the brain (9–11). Neurodegenerative diseases characterised by the presence of hyper-phosphorylated tau are collectively termed tauopathies (7, 12). The degeneration of neurons in tauopathies leads to dementia, i.e., a progressive and accelerating decline in mental function. AD is one of the most devastating tauopathies in which a patient’s memory and ability to learn is initially compromised and eventually completely destroyed. Although the behavioural pathologies of tauopathies resemble each other, there is also much variation in the clinical picture due to specific combination of neuropathological changes, variations in the form of hyper-phosphorylated tau and the individual spatio-temporal expression of neurodegeneration (reviewed by 7).

2. Tauopathy - Hallmarks and Characteristics

2.1. Compromised Behaviour

Tauopathies are diseases characterised by a progressive and severe decline in cognitive abilities that cannot be attributed to normal aging (13). Abnormalities in other behavioural systems often precede and accompany tauopathies (14). In AD, signs of mild cognitive impairment (MCI) precede overt dementia (15). In practice, individuals will only be classified as cognitively impaired when their ability to perform everyday tasks is compromised to the point that they are no longer able to function at home and in their community (16). MCI diagnostic does not always correctly predict the development of AD-type dementia. Data derived from neuroimaging, screening for genetic risk factors (17) or detection of increased tau protein levels in cerebrospinal fluid (18,19) might provide false-positive indications of beginning dementia. Even significant hippocampal atrophy is not always a reliable marker and might be due to depression, Parkinson’s disease, or vascular dementia (17). Diagnosis of AD is also complicated since amyloid plaques and neurofibrillary tangles (NFTs) of hyperphosphorylated tau are present during normal aging. Further, depending on the type of tauopathy, patients may present a variety of other cognitive complaints including delusions, amnesia, executive dysfunction, apathy, agitation, and aggressive behaviour (8, 19–22).

2.2. Neuropathology

Tauopathies are characterised by neuronal dysfunction and loss that display a varying but overlapping spatio-temporal distribution
The neuropathological variability is the reason for the complex and also varying clinical phenotypes. In AD, progressive neuronal damage and death appear in brain regions critical for learning and memory (neocortex, hippocampus, amygdala, anterior thalamus, basal forebrain, and subcortical nuclei including the nucleus basalis of Meynert (23–28)). The highest atrophy is seen in the entorhinal cortex and hippocampus (29) which positively correlates with the degree of dementia (30). The development of neuropathology is paralleled by a decreased functionality of forebrain and brainstem monoaminergic and cholinergic systems (4, 31–33).

2.3. Intra- and Extra-cellular Protein Inclusions

The main characteristic of tauopathies is an age-progressing hyper-phosphorylation of the tau protein which accumulates in tangles with paired helical filaments, twisted ribbons, and/or straight filaments (7, 34). Excessive intra-neuronal deposition of tau protein is also a key feature of dying neurons during normal aging (35, 36). Kidd, studying brains of AD patients, was the first to show that both tangles and neuritic plaques contain dense accumulation of pathological paired helical filaments (37). In the mid-1980s, tau was identified to be the principal component of neurofibrillary lesions in the AD (34, 38, 39). Similar lesions were also found in other neurodegenerative diseases including corticobasal degeneration (40), amyotrophic lateral sclerosis/parkinsonism dementia complex of Guam (41, 42), Down syndrome (43), progressive supranuclear palsy, Pick’s disease, argyrophilic degeneration (reviewed by 7), myotonic dystrophy (44), and the family of fronto-temporal dementias (FTD) (45). AD is a special form of tauopathy additionally characterised by extra-cellular deposits of Aβ and amyloid deposits in cerebral blood vessels (46). Plaques consist mainly of a 40–42 residue amyloid peptide (Aβ40/Aβ42), cleaved from the amyloid precursor protein (APP) and are surrounded by dystrophic nerve cells. The longer Aβ42 species is normally present in small, soluble fractions in biological fluids (47), but it is elevated and early deposited in cases of familial Alzheimer’s disease (FAD) (48, 49). The increase in levels of both Aβ40 and Aβ42 correlates with progress in cognitive decline (50), and the increase in Aβ40 peptide was detected in 40% of AD patients before overt amyloid plaques could be detected. Moreover, the increase in Aβ preceded the formation of NFT, suggesting that at least in the frontal cortex of some cases, soluble species of Aβ may precipitate the formation of NFTs. However, since tau pathology can occur in the absence of Aβ other causes of NFT formation cannot be excluded (50).

2.4. Genetics

The first gene mutations related to familial neurodegeneration were identified in AD in the gene encoding amyloid β precursor protein (APP) (51–53), followed by mutations in presenilin 1 (PSEN1) (54, 55) and presenilin 2 (PSEN2) (56, 57). No mutations directly associated with NFTs have been identified in AD. However, several groups of patients with FTD inherited autosomal dominant mutations in MAPT in FTDP-17 were reported (58–60). Approximately 100 families with FTDP-17 have now
been identified, with a total of 32 unique tau gene mutations (61, 62). Moreover, in addition to MAPT, mutations in progranulin, which encodes a growth factor involved in the regulation of multiple processes including development, wound repair, and inflammation, were recently shown to be implicated in FTDP-17 (63). Finally, mutations in PSEN1 gene, implicated in AD, can produce FTD-like phenotypes with the AD neuropathology (reviewed by 64).

2.5. Conclusions
The identification of gene mutations implicated in tauopathies opened a way to model the diseases in mice. At translational research level, the work is focusing on the production of experimental animal models that reproduce the essential pathology and phenotype of human tauopathies, including the formation of abundant, specific to a disease intra- and/or extra-cellular protein inclusions, neuronal degeneration, and cognitive impairment. Therefore such animal models should have the following characteristics: (1) specific to a given disorder, intra- and/or extracellular deposition of misfolded proteins, (2) age-progressing dementia, (3) disturbance in behavioural systems not related directly to cognitive function but which are observed at specific stages of a disease, and (4) coinciding neuronal loss in diseasespecific brain regions and cytoarchitecture.

3. Mouse Models of Neurodegeneration
The first successful mouse model replicating major hallmarks of neurodegenerative disease was an AD mouse model created more than a decade ago by Games and colleagues (65). Other AD mouse models followed soon and were proven extremely informative, which was chronicled in number of scholarly reviews (66–77). Mouse models expressing genes implicated in other tauopathies followed closely (reviewed by 7, 70, 78, 79).

3.1. Criteria of Mouse Models of Tauopathies
At the juncture of over a decade long history of using mouse models of neurodegeneration, it seems that a robust and good model should meet the following guidelines:

1. Replication of clinical phenotypes.
Since the cognitive decline and region-specific neuronal loss are central to neurodegenerative diseases, the model of a disease should recapitulate accurately these facets of the clinical phenotype.

2. Age-progressing phenotype.
A credible model should exhibit age-progressive neuropathology and cognitive deficits which could be evident, to various extent, in paradigms addressing different memory systems. The extent of age-progressing behavioural impairment may eventually encompass non-cognitive systems due to a significantly increasing brain pathology. Although, this may raise operational complexities associated with the interpretation of the cognitive impairment being confounded by the emergence of impairment in non-cognitive behavioural systems (80), the use of such model during drugs screens
may reveal which behavioural deficiencies can be ameliorated by a treatment at given stage of pathology.

In the models employing genetic mutations (transgenic models), phenotyping changes should be robust and correlated with the presence of familial mutations, but should be absent or less overt in the age-matched mice expressing wildtype (wt, non-mutated) gene expressed at equal (or greater) steady-state levels. These mouse models require minimum variation in their genetic background in order to be sensitive to pick up subtle changes in a phenotype. Therefore, whenever possible inbred strains, having homozygous genomes, are used almost exclusively in biomedical research. Transgenic lines, however, are often created in one strain and later backcrossed to more suitable genetic background, which complicates the control over genetic variability. Wild-derived mouse strains or recombinant strains are usually avoided in transgenic research due to unwanted genetic diversity. Similarly, outbred stocks, used predominantly in genetics, toxicology, and pharmacology, are not recommended for transgenic research due to their genetic variability caused by genetic drift, directional selection, and genetic contamination during breeding (81–83).

4. Validation of a model.
The independent confirmation and replication of the key facets of the phenotype in independent transgenic lines harbouring the same construct should be carried out in independent laboratories (84), including standardisation of expertise of technical personnel and differences in handling methods (85).

3.2. Caveats and Pitfalls of Using Mouse Strains
How close the mouse model replicates the pathology observed in a disease may often depend on the design of the model and behavioural systems in focus. The choice of a mouse strain is crucial, since many strains suitable for genetic manipulation are not particularly suitable for behavioural studies (86, 87). Genetic background of a mouse strain can be used as a tool in the analysis of a mutation (88), and the use of mapping and cloning strategies allows the identification of modifier genes existing in different mouse strains (89, 90). Noteworthy is that gene targeted or transgenic mice are usually initially created on 129 or FVB strain backgrounds. The reason for the latter is that the oocytes of these strains are large, thus increasing the probability of successful injection of the transgene construct. These strains, however, are not particularly suitable for behavioural testing (91) and it is often necessary to transfer the mutation to a more suitable background, usually C57BL/6 strain, by backcrossing for at least 10 generations. While the strategy largely results in the replacement of the donor background with the recipient background, the region flanking the selected gene remains likely of donor origin. Thus, the genes in this flanking region travel with the selected transgene, and when comparing transgenic (or knockout) mice with non-transgenic littermates, allelic differences in
the flanking genes can conceivably influence the traits of interest. The problem of genetic background was partially remedied in the case of knockout models in which genes are targeted directly in C57BL/6-derived embryonic stem (ES) cells (see Knockout Mouse Project (KOMP) http://www.komp.org/ for further information). Many neurodegenerative mouse models, however, are still maintained on mixed and segregating genetic backgrounds, thus even inbred littermates are not genetically identical. Also, one has to bear in mind that many strains, like C3H/HeJ, SJL/J, FVB/NJ, MOLF/E, PL/J, SWR/J, BUB/BnJ, CBA/J, or NON/LtJ, are not suitable for behavioural testing when visual cues are of importance due to the presence of retinal degeneration (rd) http://www.jax.org Retinal degeneration (rd) is caused by an autosomal recessive mutation resulting in a rapid age-progressing degeneration of rods and cones (92, 93). About 20% of all inbred mouse strains carry the rd causing PDE6B gene (94), and above listed strains are particularly prone to degeneration since they are homozygous for the PDE6B gene. Other strains, like A/J, BALB/cByJ, AKR/J, KK/H1J, to mention a few, are albino and can be expected to have mild defects in their vision (95, 96). Other deficits include age-progressing hearing loss in A/J, BALB/cByJ, C57BLKS/J, C57L/J, and C57BR/cdJ (but not C57BL/10 J) or other strains (129S1/SvJ, BALB/cByJ, or I/LnJ) may have partially developed corpus callosum (source, http://www.jax.org). In conclusion, taking into account the genetic strain background effect on behaviour, the breeding scheme of a transgenic line or the generation of multiple transgenic lines, the presence of retinal degeneration or other possible mutations expressed in homozygous state, the design of the transgenic mouse model has to be carefully planned to avoid serious confounding variables in most learning tasks which depend on visual acuity of animals. The observed pathology in a transgenic mouse model may also depend on the choice of the promoter used to drive a transgene expression. The most common promoters include APP promoter (AD mouse models) (97), brain-enriched prion protein promoter (98–100), the platelet-derived growth factor b-chain (PDGFb) promoter (65) (both PrP and PDGF promoters resulting in a transgene expression also outside of the CNS), and neuronal specific Thy-1 promoter (101). Another problem with transgenic models relates to spontaneous genetic changes which may affect the phenotype of a model. Mice engineered to overexpress a transgene can potentially with time change the number of disease-causing transgene copies, leading to possible loss of a phenotype. Without routine checks of a transgene expression within and between the laboratories, the differences in the transgene copy number can prevent replication of the results between the laboratories. Awareness of this issue should prompt researchers to check periodically the genetic constitution of their transgenic stocks.

3.3. Modelling Human Dementia

Given the disparities between species, it can be challenging to draw definitive conclusions about the association of cognitive function between humans and transgenic mouse models. In order to make appropriate comparisons, tests of memory in rodent models of neurodegeneration should target cognitive systems that
are found and conserved across species, including humans, and have a clearly delineated function and a defined neuroanatomy. Assessment of spatial navigation and its dependence on the hippocampus fulfills the above assumption, since this memory system is highly conserved in mammals (102). The neuroanatomical structure of hippocampus, together with changes in its synaptic plasticity during memory formation (103–110), serves as a well-defined model of memory that has been frequently employed in studies using rodent species (106, 109–112). Humans with temporal lobe damage also have severe impairments in learning and memory, including the recall of spatial locations and solving spatial maze tasks (113–115), confirming the involvement of the hippocampus in spatial memory in humans. Similar findings are seen in AD patients, who have significantly increased atrophy of the hippocampus (116, 117) and impaired performance in spatial navigation tests (118–123).

4. Methods and Experimental Design

4.1. Evaluation of Phenotype

The evaluation of the phenotype of a new mouse model of neurodegeneration should be based on a battery of tests characterising the physical and motor development of mice, their response to the array of basic stimuli, as well as the characterisation of the targeted by the model behavioural system(s). The detailed description of each test is beyond the scope of this chapter. The reader should consult existing sources, from textbooks (124, 125), general articles related to behavioural phenotyping (87, 126), specific articles related to analysis of strategies (127), memory (128, 129) in a water maze test, or methodological procedures related to spatial orientation tests (130). More specialised articles describe experimental approaches which can enhance learning in strains of mice known as poor performers in a specific test (131), or articles comparing performance of different species (e.g., rats and mice) in a test (132) and also Chapters 17 and 19 of this volume. Above all, the existence of the recently launched Nature Protocols journal http://www.nature.com/nprot/index.html which publishes detailed experimental protocols, defies any attempt here to provide complete guidance to the plethora of behavioural paradigms available. Nature Protocols articles not only present theoretical background underlying each testing paradigm, detailed procedures, examples of collected data, but also suggest type of equipment, analyses, including trouble-shooting section. Instead, we have provided a list of most frequently used tests accompanied with appropriate information which should enable behavioural and molecular researchers to use this as a compendium facilitating the interpretation of the results found in scientific literature or as a starting point in establishing relevant testing paradigms of their models.

4.2. Evaluation of Mnemonic Function

The prevalence of spatial memory tasks in the characterisation of mouse models of neurodegeneration (Table 19.1) is justified by high evolutionary conservation of spatial memory across mammalian species. The main question asked during such characterisation
relates to the abilities of a mouse genome to cope with the presence of an expressed transgene protein. The existence of possible effects due to strain genetic background, modifier genes, compensatory effects, and/or subtle differences in the experimental paradigms, including the strains’ different response to handling (133), can yield different, often contradictory results. Therefore, a broad characterization of mouse behaviour, including both hippocampus-independent memory systems (134–136) and other non-cognitive behavioural systems, such as changes in agitation and aggression levels (137), locomotor, exploratory, or stereotypic activity (138) can be very useful in evaluating transgenic mice. The results of such studies not only extend our understanding of the effect of these transgene on behaviour but also allow us to identify potential confounds in memory tests (139). Moreover, studying hippocampus-dependent memory in different testing paradigms may sometimes provide interesting additional information regarding a particular mouse model. For example, the APP Tg2576 mice were tested in T-maze alternation and contextual fear conditioning tasks (134). Investigators reported a significant impairment in T-maze alternation but, surprisingly, they found that the animals were unimpaired in both contextual (hippocampus-dependent) and auditory fear conditioning tests (hippocampus-independent task). The mice showed attenuated contextual discrimination only with a decrease in the salience of the context and without changes in tone conditioning discrimination. Such detailed validation of the existing mouse models is necessary in order to provide a more powerful experimental framework for behavioural characterisation of future models and to increase the effectiveness of screening of potential therapeutics. A pragmatic approach would dictate that robust phenotypes obtained in less labile tests (in which data collection is based on motor or strong sensory inputs) would be replicable within tolerable margins, while more labile phenotypes based on emotional or social behaviours may be strongly affected by differences in laboratory practice (140), especially in poorly managed animal colonies.

Table 19.1
Evaluation of cognitive phenotypes of mouse models of neurodegeneration. These testing paradigms represent some of the most commonly used tests employed in behavioral evaluation of mouse models of neurodegenerative disease (Reprinted from (70) with permission)

SHIRPA protocol (147)
In most cognitive tests, learning rate and memory strength is inferred from measures of locomotor behavior, therefore any possible effect of a transgene on motor and/or perceptual systems can yield false-positive (impaired learning) results. To this end, a general phenotypic assessment of transgenic mice along with non-trangenic littermates must precede specific cognitive tests in order to eliminate these possible confounds. SHIRPA protocol provides a comprehensive evaluation of mice behavior ranging from the assessment of exploration and activity levels to thermal nociception. This battery of simple tests begins with procedures most sensitive to physical manipulation, like anxiety tests performed in the open-field or elevated plus or zero mazes. Other screens focus on gross phenotyping abnormalities, assessment of sensorimotor deficits (rota-rod), holeboard exploratory activity, and thermal analgesia. Although application of the full battery can be time consuming and requires a well-equipped lab, a subset of simple tests can be carried out and is highly recommended
for initial characterization (126).

**MORRIS WATERMAZE (MWM) (148, 149)**
The MWM test has been the most widely used testing paradigm to study hippocampus-dependent spatial memory in rodent species. Reference memory or place discrimination version of MWM requires mice, trained with repeated trials over several days, to use external visual cues around the testing room to search for the hidden (barely submerged) escape platform in the water maze. Spatial navigation encompasses the development of different search strategies with spatial strategy (reflected by a direct swim to a platform) taking place at the end of this complex learning process (127, 150). The main dependent variable reflecting learning acquisition is escape latency - the time it takes a mouse to find a platform, or search path, which is less biased by the differences in swim speed. Memory bias is evaluated in trials where a mouse searches a pool where the hidden platform has been removed. Spatial learning is reflected by decreased escape latency or search path, while spatial memory by increased search in areas or quadrants of the pool containing a platform during training. An annulus-crossing index (a number of swimming over former platform location adjusted for swims in other 3 quadrants of a pool (127, 151, 152)) represents an alternative, more stringent measure of memory bias. In cases when more than one probe trial is carried out during training, a mean probe score (the mean percent of time spent in target quadrant during all probes (153)) can be used as a reliable memory evaluation index. Correspondingly, learning impairment is reflected by longer escape latency or search path during training, and memory impairment by displaced or random search, which is reflected by about 25% of time, spent in each of quadrant of the pool during a probe trial. To address episodic-like memory in mice, a more complex version of the MWM test was developed (154). In this test, numerous locations of the platform were used and the number of new locations learned during the whole training reflects learning capacity of an individual mouse. In a cued or visible platform version of the MWM test a platform location is marked by a visible cue that mice associate with an escape from water. This version of the test is often implemented as a control for normal visual acuity, an unimpaired learning of simple association between a proximal cue and an escape platform, or as demonstration of a comparable swim speed between studied genotypes. These controls should be used with caution, however, because in contrast to rats, some strains of mice with hippocampal lesions often show also partial impairment in the cue navigation task (155).

**OBJECT RECOGNITION (OR) (156, 157)**
This test exploits a natural tendency of rodents to explore novel objects and to show an exploratory preference for replaced or displaced objects. The dependence of object recognition memory on the hippocampus is related to the protocol of a test. Short delays between initial exploration phase and a memory test make OR test independent from the hippocampus (136), however, when longer delays (hours) are implemented, OR memory depends on hippocampus function (158, 159). Object memory impairment is demonstrated when an animal shows no preference in exploration (close proximity, nose contact) of a new or displaced object.

**FEAR CONDITIONING (FC) (160, 161)**
The FC paradigm which is an example of classical Pavlovian associative learning and involves an association of a neutral tone (conditioned stimulus, CS) paired with a brief electric foot-shock (unconditional stimulus, US) delivered in a novel context. Mice trained in that manner develop a fear response (conditioned response, CR), expressed as defensive or anti-predatory behavior in a form of freezing (complete cessation of movement) which coincides with autonomic and endocrine response (increased heart beat rate and blood pressure), and sensory alteration (analgesia, potentiated startle). The paradigm may involve two types of conditioning that can be performed simultaneously or independently during a training phase: contextual (CFC), when an animal develops an association between shock and training context (conditioning chamber), and tone fear conditioning when shock (US) is
associated with a neutral tone (CS). The tone conditioning is performed either as delay conditioning paradigm when there is a temporal overlap between CS and US (a foot-shock is delivered within the last 1–2 s of tone duration), or more demanding trace fear conditioning which requires the association of a CS with an US across an interval of time known as trace interval (a foot-shock [US] is delivered after the tone [CS] is turned off). The time between CS and US can vary and an additional temporal processing is required because CS and US are separated therefore an animal has to retain a trace of CS across this time interval in order to associate it with the US. While delay tone conditioning is hippocampus independent but requires intact amygdale (162, 163), the trace and contextual fear conditioning are sensitive to hippocampal lesions (161, 164). The sensitivity of the mice to foot-shock can be established empirically recording the current thresholds that elicit specific response like flinch, jump (165). The lowest current eliciting learning (for mice a current of 0.35–0.4 mA is appropriate) should be used. Impairment in FC is evaluated during test phase on the following day after training, and is reflected by reduced freezing time when an animal is placed in familiar chamber (context conditioning) or when the animal is exposed to a conditioned tone in a new environment.

CONDITIONED TASTE AVERSISON (CTA) (166–168)
CTA is a special form of classical Pavlovian conditioning, representing an adaptive specialization which defends an organism against repeated ingestion of toxic foods (166–169). CTA is well conserved in many different species including humans (169, 170). When acquiring a CTA response, an animal learns to associate the specific taste of a novel food, usually a saccharine solution (conditioned stimulus, CS) with experimentally induced through i.p. injection of lithium chloride after saccharine intake (unconditioned stimulus, US), nausea. Because of one trial pairing between CS and US, a long-lasting avoidance of food with this specific taste develops. The brain areas implicated in the CTA include the agranular insular cortex, the parvicellular thalamic ventral posteromedial nucleus, and the parabriachial nucleus of the pons, which are part of the gustatory pathway (171, 172) and the amygdala (173, 174). Impairment in CTA is reflected by increased saccharine intake as compared to control mice in choice tests (usually two bottles test, one containing water, one saccharine).

4.3. Experimental Design
The purpose of doing experiments is to distinguish between alternative hypotheses or explanations. However, even a perfectly designed experiment might lack sufficient power, if sample sizes of animals in experimental groups are small. Since it is often not known what effect on behaviour a mutation exerts, it is advisable to properly characterize a non-transgenic or wild-type control mice, thus establishing a yardstick for robust characterization of the phenotype in question, avoiding floor or ceiling effects in the data recording which can cause a skewed data distribution. Larger sample sizes are desirable (n = 8–12, or more (141) depending how robust a focal behaviour is), but attention has to be paid to ensure that the mice are not tested too long during the day, which may result in their fatigue or may span over different phases of the circadian cycle. Also, one has to be aware that a change in behaviour of mice may sometimes be caused not by the experimental treatment, but merely by the handling or attention paid to them by the experimenter. The effect, known in psychology as the Hawthorne effect (142), is often a cause of differences between the obtained results in various laboratories (85). A common error, which occurs less often in behavioural research but crops up frequently in physiological experiments, refers to treating repeated data points coming from the same subjects as independent from each other measures. This approach, called pooling fallacy (143), leads to an inappropriate increase in the sample size of mixed,
dependent, and independent data points, thus violating many assumptions of experimental design and parametric data analysis. Problems with independence of data may arise in less obvious situations, when the obtained data correlate closely between mice coming from the same litter or between mice housed in the same cage. These litter- or cage-effects can be the result of, for example, differential maternal care, highly variable housing conditions (mice in cages placed at the bottom of a rack in a densely populated with racks rooms are kept in constant semi-darkness), singly housed animals are known to perform worse in learning and memory tasks, etc., and can introduce confounding factors often impossible to overcome due to the small number available and often difficult to derive mutant mice. Awareness of these issues, however, may help during the inspection and first steps in the interpretation of the raw data. If the data generated by mice from the same litter or cage are on one spectrum of the distribution, one should consider the replication of the experiment using more careful and balanced assignment of mice to experimental groups.

4.4. Data Analysis Most experiments, especially those evaluating learning and memory involve repeated tests or training sessions. Results are usually presented in blocks of training trials or days. Plotting data in that fashion usually reflects adequately the learning process, but blocking data over repeated trials reduces variance and may result in some unusual patterns of behaviour escaping attention. It is advisable, therefore to inspect the raw data, especially the data generated from first training trials to check if mutant mice are free from subtle motor or sensory deficits. It can be assumed that control and mutant mice, which are well habituated to handling and lab conditions but naïve to a particular behavioural test, should show comparable performance during first training trial(s). Any cognitive impairment, if not confounded by compromised locomotor or sensory deficiency, should become apparent as training progresses, but should not be present at the beginning of training, unless of course the severity of cognitive decline impairs the interaction of an animal with the surrounding environment. Data are generally analyzed by analysis of variance (ANOVA), with genotype and/or treatment as between subject factors and training days and/or trials as within-subject factor(s). One issue to remember is that the data must meet the criteria of parametric statistics and in the case of repeated measure or within-subject factor, an assumption of compound symmetry must be met in order to avoid bias in the interpretation of the results of a test involving within-subject factor. The assumption of compound symmetry refers to a pattern of constant variances on the diagonal and constant covariances off the diagonal in the variance–covariance matrix. In practice this means that the correlations within the matrix of the repeated factor (days or trials) have to be the same at all distances between measurements. This assumption, however, is hardly met in the analysis of learning data, since as animals learn over time and improve their performance in a task, thus the variance decreases as learning progresses. A departure from the assumption of compound symmetry is usually evaluated by slightly the more stringent sphericity test (Mauchly sphericity test, SPSS GLM (Statistical Package for Social Sciences, SPSS Inc.)
Chicago)), and in cases of severe departures, degrees of freedom should be adjusted either by Greenhouse–Geisser ε-correction (tends to underestimate, especially when ε is close to 1) or by Huynh–Feldt estimator (which tends to overestimate ε) to avoid false-positive results (144).

4.5. Animal Facility and Behavioural Tests
Replication of the results are at the core of the falsification process of hypotheses or theories (145). However, even the best guidance or step-by-step description of procedures and methods pertaining to a specific testing paradigm may often yield unpredictable and different results between laboratories, despite very careful execution by experienced in the field researchers (84). Therefore, in this section we would like to alert readers to some potential problems and issues which may affect the execution of behavioural experiments and resulting data. The list of presented issues, by no means exhaustive, includes problems which are often not formally documented, or issues which are considered trivial or too obvious to mention in many the Material and Methods sections but are highly relevant.

The characterisation of behavioural phenotypes starts with housing conditions in the animal facility. Animal rooms located in the vicinity of noisy, heavy traffic areas, like cage washing areas, are less desirable and in some cases may even lower the breeding rate, induce cannibalism, and increase hyper-reactivity in animals. The housing conditions, including the number of mice in a cage (not more than four mice over extended period), having minimum and consistent level of enrichment in a form of pressed cotton nesting material (nestlets), mouse huts or igloos, or pup tents (source http://www.bio-serv.com/) can significantly improve breeding and reduce anxiety of animals. Housing mice singly in cages is not recommended due to heightened rate of developing stereotyping behaviour, obesity, and decreased performance in learning ability. When breeding transgenic lines, it is not uncommon that the newly born transgenic pups tend to be smaller then their non-transgenic littermates (for example in the case of APP TgCRND8 mice, personal observation). Supplementing lactating females and their pups, especially at the pups age when they start to consume solid food (14–15 days of age), with easily available and more palatable moister powdered mouse Purina chow in a Petri dish, facilitates pups’ growth and can reduce the weight difference between transgenic TgCRND8 and their non-Tg littermates (unpublished data). The distance of the housing room to the behavioural testing room(s) is relevant and the transportation of mice between the two locations can be stressful, therefore adding appropriate time for acclimation to new testing conditions should be done. Last but not least, husbandry practices including care and feeding of animals, cleaning of equipment, physical surroundings, and routine checks of the stock health by experienced, well-trained, and well-managed facility staff guarantee good health, growth, reproduction, survival of mice. Personnel with poor management and/or inexperience in mouse handling and husbandry may adversely affect animal stress level and behaviour. Excessive noise produced during cage changing (for example changing cages under a hood and putting or stacking metal cage lids on the metal surface produces
extremely noisy conditions, including high levels of ultrasounds which mice are sensitive to), undetected leaking water bottles or wet cleaning equipment (mops and buckets) left in animal rooms produce impossible to pinpoint confounding variables all of which increase stress and anxiety levels of mice, consequently negatively and variably affecting their performance in the behavioural tests.

As an example, in Fig. 19.1 we provide the results of training two cohorts of same-age C3B6 mice (mixed genetic background of C3 (C3H) and B6 (C57BL/6)) in the spatial reference memory version of a water maze test, in two different animal facilities. Both cohorts were trained by the same laboratory assistant, highly experienced with the behavioural procedure, mouse husbandry, and certified in laboratory animal medicine. Mice in animal facility A were maintained in a quiet room and highly qualified and well-managed personnel provided high quality husbandry care. The environment in colony B was more stressful and mice were exposed to noisy conditions. The comparison of mice performance between colonies (main between-subject factor) and the analysis of their learning (days as repeated measure or within-subject factor) revealed no significant difference in the average performance between the colonies \((F(1,23) = 1.2, \text{NS})\), but it also revealed a significant interaction between the colony conditions and learning rate of mice \((F(4,92) = 4.2, p < 0.01, \text{colony by training days interaction})\). Mice in colony A showed a significant learning through a rapid improvement in their search for a hidden escape platform over days (Fig. 19.1, \(p < 0.001\) – simple effects ANOVA with days as a repeated measure), whereas the mice in colony B, however, did not show any signs of improvement over training period (Fig. 19.1).

Fig. 19.1. Learning acquisition in the spatial reference memory version of a water maze test of mixed background C3B6 mice in two different animal colonies. Mice in the colony A were kept in a quiet room with appropriate for behavioral experiments husbandry practices, while the colony B had increased noise level and sub-optimal for behavioral studies conditions. The mice trained in the colony A showed a significant improvement in finding a hidden platform location during training (their search path was on average about 5 m shorter at the end of training as compared to at the beginning of training). On the other hand, the mice in colony B showed no improvement (the rate of improvement between day 1 and day 5 was about 1 m). See text for further details.

5. Conclusions
The main goal of the generation of animal models of human diseases is to better understand their underlying pathology, which should lead to the discovery and tests of potential therapeutics. In our review we focused on neurodegenerative diseases which present a complex, age-progressing dementia with rapidly progressing neuronal death. It might be unrealistic to think that the full complexity of human brain disorder can be modelled in a mouse using a crude, single genetic modification. However, as we tried to outline in this chapter, the interpretation of the results coming from mouse models of neurodegeneration should not follow “one gene – one disease” paradigm. At present, none of the existing mouse models of tauopathy fully replicate the characteristics of the modelled disease. Using these mouse models to study
well-conserved signalling pathways in vivo may be better warranted than replicating fully the complexity of dementia. Genetically modified mouse models are integral part of modern drug discovery, but the interpretation of the obtained results must be carried out within constraints of a model and mouse biology. The intensive screens of many compounds would require systematic, well-controlled, standardized phenotyping approach which is presented by Sacca and colleagues in Chapter 3 of this volume. The initial characterisation of new models or detailed characterisation of specific aspects of existing models should be based on careful experimental design, including larger number of mice in completely randomized experimental designs in animal facilities which promote maximal expression of mouse natural behaviour. Rigor of the experimental design will ensure replicability of the results across the labs and between different models. Even seemingly good and robust models of a human disease can yield many false-positive results due to differences in methodology or less rigorously carried out experiments (146). Our intention is to highlight important aspects of experimental design which are not always identified a priori and which may often generate confounding factors seriously biasing obtained data. We argue that our success in the endeavour of modelling human cognitive impairment may often depend on how well we understand the behaviour of a mouse. Detailed analysis of the potential and limitation of a model and the interpretation of the results within the framework of mouse biology should improve considerably detailed evaluation of potential therapeutics. Testing a specific hypothesis, negative results should be as valuable as positive ones and should be made available to the scientific community.

Acknowledgements
We would like to thank Amanda Hanna, Heather Melrose, Douglas Wahlsten, and Alexander Gaukhman for valuable comments on earlier drafts of the manuscript.

References
signaling is impaired in Alzheimer’s disease brain. *Neurobiol Aging* 18, 111–120.


