1

Genetic Screens in Neurodegeneration

Abraham Acevedo Arozena and Silvia Corrochano

MRC Mammalian Genetics Unit, Harwell, UK

Introduction

One of the major challenges of modern biology is to further understand the molecular etiology of human diseases with the ultimate aim of providing cures or disease-ameliorating therapies. Neurodegenerative disorders are a heterogeneous group of diseases leading to the death of specific neuronal populations. Generally, neurodegenerative diseases are fatal disorders for which there are currently no effective therapies. They share common features such as the progressive nature of the disease and the association with increased age. Thus, with a growing aging population, the prevalence of neurodegenerative diseases is steadily increasing. Another common end-point is the development of unique proteinaceous inclusions and the death and/or dysfunction of particular neuronal populations that are signatures of each disorder. The appearance of these inclusions is common to the major neurodegenerative disorders, including Alzheimer’s disease (AD), Parkinson’s disease (PD), Huntington’s disease (HD) and amyotrophic lateral sclerosis (ALS). Inclusions are formed through protein misfolding and therefore these disorders can be classified as protein conformation diseases or proteinopathies.

From the genetic point of view, neurodegenerative diseases fall into two major categories: (i) familial or single gene disorders in which a single causative gene is inherited in a dominant or recessive manner, and (ii) sporadic or idiopathic cases, in which no previous family history exists and no mutations in previously identified causative genes can be generally found. For the major multigenic neurodegenerative disorders such as AD, PD, or ALS, the majority of cases are sporadic, making the identification of novel therapeutic targets more challenging.

Thus, advancing our understanding of the pathogenesis of neurodegenerative diseases is critical, aiming to find not only new causative genes but also to identify new genes and pathways that affect disease onset or progression rate. Lessons from other diseases suggest that a deeper understanding of disease pathogenesis can ultimately lead to novel drug targets and therapies. Indeed, research into neurodegeneration pathogenesis is currently at a critical turning point, starting to translate pathomechanistic findings into disease-ameliorating treatments (Fig. 1.1).

The identification of the causative mutations in familial cases represents the main entry point for investigations into neurodegeneration molecular pathogenesis. In some rare cases, the identification of causative mutations has provided immediate insight into the biology of the disease. However, moving from a causative gene to understanding disease pathogenesis, typically using a number of disease models, has proven to be an extremely difficult task for a wide
The Genetics of Neurodegenerative Disorders

The genetics of neurodegenerative disorders varies from single-gene disorders such as HD and other polyglutamine disorders, to very complex multigenic disorders such as AD, PD, or ALS.

Rarely, single-gene mutations can also cause cases of AD, PD, and ALS that in general is pathologically indistinguishable from sporadic cases, although they tend to cause earlier disease symptoms [2]. These cases are termed familial and have a previous family history of the disease. Some sporadic cases can have a concealed family history or may carry de-novo mutations that are not transmitted through the family. Understanding the genetic basis of neurodegeneration is crucial for developing targeted therapies and insights into the disease mechanisms.
mutations in familial neurodegenerative genes. Human genetics are extremely powerful in identifying genes involved in Mendelian or oligogenic disorders, with recent developments in genome sequencing technologies leading to the discovery of a number of novel genes involved in neurodegeneration [3]. Genome-wide association studies (GWAS), exome or whole-genome sequencing of large numbers of patients and controls, are currently underway and likely to lead to the identification of an ever-increasing number of genes linked to neurodegeneration. Thus, the list of causative familial neurodegeneration genes has expanded rapidly over the last decade (Table 1.1).

Table 1.1 Genes causative for familial ALS, PD, and AD.

<table>
<thead>
<tr>
<th>Amyotrophic lateral sclerosis (ALS)</th>
<th>Parkinson’s disease (PD)</th>
<th>Alzheimer’s disease (AD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOD1</td>
<td>SNCA</td>
<td>APP</td>
</tr>
<tr>
<td>VAPB</td>
<td>LRKK2</td>
<td>PSEN1</td>
</tr>
<tr>
<td>ANG</td>
<td>PKRN</td>
<td>PSEN2</td>
</tr>
<tr>
<td>TARDBP</td>
<td>PINK1</td>
<td>APOE*</td>
</tr>
<tr>
<td>FIG4</td>
<td>DJ1</td>
<td>TREM2*</td>
</tr>
<tr>
<td>FUS</td>
<td>ATP13A2</td>
<td></td>
</tr>
<tr>
<td>OPTN</td>
<td>PLA2G6</td>
<td></td>
</tr>
<tr>
<td>C9ORF72</td>
<td>FBXO7</td>
<td></td>
</tr>
<tr>
<td>UBQLN2</td>
<td>VPS35</td>
<td></td>
</tr>
<tr>
<td>VCP</td>
<td>elf4G1</td>
<td></td>
</tr>
<tr>
<td>SETX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPG11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEFH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALS2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SQSTM1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PFN1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GBA*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Asterisks (*) represent major genes whose variants predispose to AD and PD.

In the complex neurodegenerative conditions, the majority of cases occur in patients with no apparent previous family history of the disorder and are termed sporadic or idiopathic. For the three major complex disorders (AD, PD, and ALS), we still do not understand the etiology of the great majority of sporadic cases. In those cases, neurodegeneration appears to be caused by the very complex interaction of a significant number of common genetic variants that appear to be unique for each disease. These genes act in concert with environmental factors, leading to the development of neurodegeneration. There is still some debate about the strength of the involvement of genetic inheritance in sporadic cases of AD, PD, and ALS. However, the involvement of genetic heritability of a relatively large number of genes is backed by GWAS studies. So far, GWAS studies have found a number of genetic variations potentially affecting the risk of developing each disorder, but the relative contribution of each of these variations remains relatively small.

Remarkably, at end-point, each neurodegenerative disorder is characterized by well-defined pathological features that are generally similar for the majority of sporadic and familial cases. As familial and sporadic cases generally lead to common pathological end-points, genes mutated in familial cases have been used as an entry point to understand the genetics and pathophysiology of both familial and sporadic neurodegeneration.
In multifactorial complex diseases, as well as in monogenic Mendelian disorders, the identification of downstream genes and pathways able to modulate disease pathobiology is critical for the understanding of the genetics of these diseases. Modifier genes, broadly defined as genes that are not disease causative but could modulate disease pathobiology, play a critical role in disease etiology and progression. Even in monogenic Mendelian diseases such as HD, a degree of clinical variation can be attributed to variation in modifier genes [4]. A number of major risk loci, not disease causative but contributing towards the lifetime risk of developing the disease (modifier genes), has already been identified. Major examples of these risk loci include \textit{APOE} alleles in AD. The e4 allele is the major genetic risk factor to develop AD, whereas the e2 allele is protective [5]. Recently, \textit{TREM2} variations have been identified as another risk factor for AD [6]. Also, variations in \textit{GBA} have been shown to contribute to the risk of developing PD [7] and non-pathogenic intermediate polyglutamine tract expansions in ataxin-2 modulate the risk of developing ALS [8]. The identification of these modifier genes and pathways, together with their molecular mechanisms of action, is the initial step towards the identification of novel therapeutic targets other than the causative genes, leading to potential therapies for both familial and sporadic cases.

**Neurodegeneration Disease Models**

The discovery of an ever-increasing number of single gene mutations leading to inherited forms of neurodegeneration has provided a strong impetus for the use of disease models to study the cellular consequences of pathogenic mutations. As discussed above, sporadic and familial cases generally share end-point pathological signatures despite having heterogeneous starting points, leading to the conclusion that by studying and understanding disease mechanisms in familial cases we are also learning about the majority of sporadic cases. Thus, a plethora of disease models based on known familial genetic variations linked to neurodegeneration has been developed over the last few decades. In this framework, genes mutated in familial neurodegenerative cases provide an entry point into disease pathobiology that is lacking for sporadic cases.

Once a new gene involved in neurodegeneration is identified, the next step is usually to make a variety of disease models to start to understand how mutations can lead to neurodegeneration. Disease models can play a very useful role in neurodegeneration research as they can model all disease stages, from early pre-symptomatic to end-stage disease. Thus, it is critical to validate the findings from the models using patient material. A deeper understanding of these early disease processes should lead to the discovery of novel therapeutic targets and also to the identification of novel disease biomarkers enabling earlier diagnosis. This could be extremely valuable, as it would allow more time for any future therapeutics to affect disease progression.

However, no disease model is perfect, and single models cannot recapitulate all aspects of the human conditions. Thus, a wide variety of disease models are required to tackle this complexity. In this sense, each model organism can be an extremely valuable tool, taking advantage of their different and unique characteristics for the study of different aspects of the disease. The limitations mentioned above also apply to patient material, including patient-derived cells like iPS cells that can be differentiated into neurons. Despite all the evident advantages of iPS cells, they are also disease models and as such are not able to recapitulate all disease aspects, particularly those involving the complexities of a mammalian brain. Therefore, each model has a role to play in our quest to try to understand the complexities of neurodegeneration and ultimately find cures.
Genetic Approaches to Discover New Genes Related to Neurodegeneration Using Disease Models

Using disease models, a number of approaches have been used to understand why mutations can lead to neurodegeneration. The genetic approach aims at identifying modifier genes and pathways that can affect the phenotypic outcomes caused by the expression of mutations linked to neurodegeneration. A modifier gene could be an enhancer, if it accelerates the phenotypes of study, or a suppressor, when the gene ameliorates the outcome of the phenotypes of study. Genetic screens have been developed using a wide variety of models including yeast, flies, and worms carrying mutations in familial neurodegeneration genes (Fig. 1.2).

Genetic screens are based on the assumption that changes in the measured phenotypic outcomes are largely dependent on the genetic variation introduced to the screen. Genetic screens can be classified as forward or reverse, depending on their starting point. Forward genetic, or phenotype-based, screens aim to identify the gene or genes responsible for a particular phenotypic trait. Therefore, forward genetic screens rely on the identification and modification of phenotypes of interest, usually from a mutagenized population. Once novel or modified phenotypes are identified, the gene or genes responsible are then found using standard genetic techniques such as positional cloning and deep sequencing. Reverse genetic screens aim at identifying the phenotypic outcome of the disruption of known gene/s in a model organism. Thus, forward genetic screens start with a mutant phenotype and then identify the gene or genes responsible, while reverse genetic screens start with a disrupted known gene to ascertain

![Figure 1.2](image-url) Genetic screens in model organisms are useful to understand disease pathogenesis and provide novel therapeutic targets.
the phenotypic consequences of known gene disruptions. The power of the genetic approach is exemplified by the work of Christiane Nüsslein-Völlhard and Eric Wieschaus using a forward genetic screen in *Drosophila* to identify a large number of loci involved in embryogenesis [9]. As exemplified by this seminal work, and a large number of subsequent genetic screens, forward genetic screens can be extremely powerful in identifying novel gene functions and genetic networks involved in particular phenotypic traits. Their major advantage is the lack of *a priori* assumptions about the functions of a gene, as mutagenic agents largely produce mutations randomly throughout the genome. Their major limitation in the genetic screens is the definition of the phenotypes of study and the ability to analyze them. Therefore, genetic screens are not completely unbiased, as inevitably only certain phenotypes are scored and the ability to identify and elaborate on particular phenotypes may introduce certain biases towards particular mutations. Also, genetic screens are inherently biased by the model organism of choice, as depending on the model certain mutations may be better tolerated than others. Overall, the success of a genetic screen critically depends on the ability to identify mutants by sorting through the mutagenized individuals to recognize phenotypes of interest.

In screens aimed at identifying genes and pathways involved in neurodegeneration, as for any other genetic screen, the most critical task is first to generate models and identify robust, quantifiable phenotypes amenable for screening. Models have been created in a number of organisms from yeast to mammals. For dominant mutations, the heterologous expression of human cDNAs encoding wild-type or mutant forms of genes mutated in familial neurodegeneration is usually the approach of choice. For mutations leading to loss of function, models can also be generated by creating null alleles of the orthologous gene in the model organism. Although human genes can be functionally equivalent to the orthologous gene, model systems could produce phenotypes that may not be related to the pathogenicity seen in human patients. Overall, it is critical to use a variety of disease models, including human patient material, to ultimately validate novel disease pathomechanisms.

We will now focus on reviewing each of the different cellular and animal models for genetic screens used in neurodegeneration, emphasizing the advantages and limitations of each of the specific models, showing their value in the study of particular phenotypes of interest.

**Saccharomyces cerevisiae**

Yeast, typically baker's yeast *Saccharomyces cerevisiae*, is a very powerful model system used for gene function discovery and drug development. One of the advantages of using yeast as a model system is that it allows assessment, in a very powerful manner, of genetic interactions leading to the discovery of novel pathways involved in cell-autonomous phenotypes very rapidly. Of course the major caveat is that yeast lacks the complexity of multicellular animals and the connection between the human brain and yeast is remote. However, a wide variety of gene functional properties are common to both yeast and humans. This can be exploited to start to understand the basis of the intracellular consequences caused by mutations affecting neurodegeneration. Modifier genes obtained in yeast screens are usually validated in other animal disease models that ultimately include disease patient material.

In yeast, model systems first need to be established, usually expressing human wild-type or mutant cDNA leading to measurable phenotypic outcomes. Over-expression of wild-type or mutant forms of human genes that have been previously involved in neurodegeneration can lead to toxicity, measured by the ability of yeast strains to grow under certain conditions. Once a model system is established, one can take advantage of the power of yeast genetics to search for enhancers and suppressors genome-wide. By using this strategy, suppressors and enhancers can be rapidly identified, establishing genetic and biochemical pathways involved in cellular toxicity.
Yeast genetic screens offer a unique entry point into disease pathogenesis. Suppressors and enhancers identified through unbiased yeast genetic screens for toxicity induced by neurodegeneration genes have yielded an ever-increasing number of modifier genes. These include modifiers for the ALS genes TDP43 and FUS, the PD genes \(\alpha\)-synuclein [10] and LRKK2 [11] and huntingtin [10] among others. The validation of these modifiers in other systems, including mammalian models and ultimately patients, is a real challenge; perhaps the most interesting feat of such screens is the identification of novel pathways involved in disease pathogenesis, potentially leading to the discovery of new and druggable targets for each disorder. For example, in the case of \(\alpha\)-synuclein, a potential dysfunction of endoplasmic reticulum in Golgi vesicular transport was first identified using a yeast heterologous model [12]. The yeast model system was generated based on the expression of human wild-type and mutant \(\alpha\)-synuclein. In this model, human \(\alpha\)-synuclein localized to the plasma membrane, but an increase in \(\alpha\)-synuclein gene dosage from one copy (leading to no growth defect) to two copies resulted in growth arrest and cell death which was accompanied by the presence of intracellular \(\alpha\)-synuclein inclusions [13]. A number of cellular defects have been previously implicated in the etiology of synucleinopathies, including mitochondrial dysfunction, reactive oxygen species (ROS), impairment of the ubiquitin-proteasome system, accumulation of lipid droplets, and ER stress. As the yeast models recapitulate many of the above cellular PD defects, they are valid models to further understand the functions of \(\alpha\)-synuclein and to try to understand which defects are likely to be disease causative. In this example, a genome-wide over-expression library expressing a large number of yeast open reading frames was used. The ability of yeast strains overexpressing one gene at a time to enhance or suppress the toxicity induced by \(\alpha\)-synuclein was measured by the ability of the yeast strains to grow when compared to the \(\alpha\)-synuclein overexpression alone. Using this strategy, a number of yeast genes that, upon over-expression, could enhance or suppress the toxicity induced by \(\alpha\)-synuclein were identified. The most effective class of suppressors identified were genes known to be involved in ER to Golgi transport, including the yeast ortholog of RAB1. Thus, an early cellular defect likely to be causative of the \(\alpha\)-synuclein cellular toxicity in yeast was first identified, by means of a genetic screen, based on the ability of genes already known to function in this pathway to modify the phenotypic outcome. This discovery was validated in fly and worm models of synucleinopathies. Moreover, deficits in this ER–Golgi transport pathway that have also been reported to affect autophagy in vitro and in mouse models [14] further validated the use of yeast model systems as a starting point leading to the discovery of novel pathways involved in disease pathogenesis.

In a number of cases, modifiers initially identified through unbiased yeast genetic screening have been found to be important in disease pathobiology in human patients. An excellent example is the identification of non-pathogenic polyglutamine (PolyQ) length expansions of ataxin-2 (between 27 and 33 polyQ) as a risk factor to develop ALS [8]. Ataxin-2 is a polyQ protein mutated in the neurodegenerative disorder spinocerebellar ataxia type 2 (SCA2). As in HD and other polyQ disorders, the length of the ataxin-2 PolyQ tract correlates with SCA2 disease onset, with lengths over 34 PolyQ leading to the development of SCA2, and longer repeats leading to earlier onset. The link between ataxin-2 and ALS was first identified when the yeast ortholog of ataxin-2, \(pbp1\), was found as a dose-sensitive modifier of toxicity induced by the heterologous expression of the familial ALS and frontotemporal dementia gene TDP43 in yeast. This functional interaction was validated in vivo in a Drosophila TDP43 proteinopathy model. This prompted the authors to assess a possible link between ALS and ataxin-2, finding that non-SCA2 causative, intermediate PolyQ repeat expansions (between 27 and 33 PolyQ) in the human ataxin-2 gene are associated with an increased risk to develop ALS. Since the initial discovery, multiple references have confirmed this link between intermediate length ataxin-2...
PolyQ expansions and the risk of developing ALS in different human populations [15] [16]. However, we still do not understand the connection between ataxin-2 intermediate repeat expansions and ALS pathogenesis.

Overall, yeast provides an excellent high-throughput starting point model system to understand the biological processes perturbed by mutations linked to neurodegeneration, allowing for the rapid and cost-effective identification of multiple candidate modifier genes and pathways. Also, the unsurpassed knowledge of yeast genetic and biochemical pathways provides the opportunity to use systems biology approaches to understand the biology of cellular pathways perturbed by neurodegeneration. Their intrinsic limitations relate to the nature of the genes that can be identified and the number of phenotypic outcomes that can be assessed.

**Caenorhabditis elegans**

The worm *Caenorhabditis elegans* (*C. elegans*) is a very accessible and easily manipulated invertebrate with a relatively simple central nervous system (CNS) composed of 302 neurons. The possibility of specifically labelling different neuronal pools and the ability to reproduce the aging process make *C. elegans* an excellent model to study neurodegenerative diseases. *C. elegans* has been extensively used to model neurodegeneration and to identify modifiers of known familial neurodegeneration genes and elaborate on their possible functions. Forward genetics or phenotype-based approaches have also been carried out in *C. elegans* to identify genes required for normal function of the nervous system, including neuronal outgrowth, synaptic transmission, mechanotransduction, or locomotion [17]. Sydney Brenner carried out the first *C. elegans* mutagenesis genetic screens, isolating a large number of mutants showing viable but visible phenotypes. In a study published in 2013, *C. elegans* was used in a forward genetics screen to establish a crucial role for glia location in synapse formation during growth, establishing for the first time that glia position plays a critical role in the establishment and location of synapses [18]. Strategies have also been developed to avoid possible pleiotropic effects of gene knockdown in whole animal survival. This allows the silencing of genes in specific neuronal populations to further understand their functions [19]. This approach could be further exploited to systematically search for genes required for the function of specific neuronal populations.

One of the advantages of the worm over yeast or cellular models is that it allows the study of non-cell-autonomous disease mechanisms, taking into account the complexities of multicellular organisms and the interactions between different cellular populations. These include differing susceptibility to neurodegeneration in different neuronal populations, but also the effects of other CNS (or non-CNS) cellular populations, such as glia or muscle, on neuronal degeneration. There is now increasing evidence that proteotoxicity can be communicated between cells and tissues, leading to disease spreading. Due to its relative simplicity and multicellularity, *C. elegans* has emerged as an ideal animal model system to study these non-cell-autonomous disease mechanisms. These include prion-like propagation of proteotoxicity and the relation between the organismal regulation of the proteostasis network and/or stress responses with neurodegeneration [20].

The ability to manipulate the worm genome enables genetic screens using interference RNA (RNAi) to down-regulate specific mRNAs in a genome-wide manner. Many *C. elegans* models have been produced [21] and are used to screen for disease enhancers and suppressors, including HD [22], AD [23], PD [24], and ALS [25] models. As an example, transgenic expression of mutant human Tau in the worm leads to an uncoordinated phenotype. This phenotype could be exploited in a suppressor phenotype-based screen to uncover a novel modifier of Tau toxicity, named Sut-2. Expression levels of the orthologous gene for Sut-2 in humans (*MSUT2*) are reduced in post-mortem tissue from AD cases in brain regions affected by Tau pathology,
suggesting MSUT2 as a novel therapeutic target for tauopathies such as AD [26]. Moreover, *C. elegans* has also been used to identify common modifiers for other neurodegenerative disorders. For example, genetic screens have been used to identify common modifiers for HD and ALS (SOD1) in *C. elegans* models. Interestingly, a number of common modifiers have been discovered affecting the worm proteostasis network, suggesting that common treatments for proteinopathies such as HD and ALS are possible [27].

*C. elegans* has been extensively used to model neurodegeneration and to identify modifiers of known familial neurodegeneration genes and elaborate on their possible functions. Models for dominant mutations usually rely on heterologous neuronal expression of wild-type or mutant human cDNA transgenes. Transgenic animals usually develop motor or pathological phenotypes that could be potentially exploited in genetics screens. Obviously, one of the big advantages of the worm over models is that the worm is a multicellular animal with a relatively simple and very well studied CNS, allowing for the identification of genes specifically involved in neuronal function [21]. Another advantage of using *C. elegans* for high-throughput genetic screens is the low cost and its high generation rate. However, its drawbacks include the relative lack of complexity of its CNS, as well as some intrinsic limitations for the manipulation of the worm genome.

**Drosophila melanogaster**

The fruit fly *Drosophila melanogaster* has also been very successfully used in genetic screens and to model neurodegenerative disorders including ALS, PD, AD, HD, and many others. The powerful genetic techniques developed for *Drosophila* over the last century of its use as a model for genetic research allow the expression of human genes in a tissue-specific manner as well as the study of loss or gain of function mutations in the *Drosophila* orthologous genes. The fly brain is much more complex than that of *C. elegans*, allowing for the characterization of behavioral phenotypes. Seminal work by Seymour Benzer proved that mutations in single genes could affect complex behaviors in *Drosophila*. This opened an avenue for the use of forward genetic screens using mutagenesis aimed at the identification of genes affecting behavior, including genes involved in learning and memory [28]. The combination of the possibility of mutagenesis and the use of behavioral phenotypes in genetic screens makes *Drosophila* a formidable model organism to further understand how mutations in particular genes lead to complex motor and cognitive phenotypes, and ultimately neurodegeneration. Indeed, *Drosophila* has been used very successfully over the last few decades to identify novel pathways involved in neurodegeneration downstream of the causative mutations.

As with all other organisms, disease models first need to be established. *Drosophila* has classically allowed for the generation of null alleles that could be informative for loss of function mutations in human patients. An excellent example of the use of null alleles to understand neurodegenerative disease pathogenesis in *Drosophila* comes from the characterization of phenotypes arising in flies carrying pink1 and parkin null alleles. Recessive *PINK1* and *PARKIN* mutations are causative of PD (Table 1.1), and thus loss of function of *PINK1* and *PARKIN* are likely to contribute to PD in these patients. Pink1 and parkin loss of function lead to very similar deficits in *Drosophila*, including muscle degeneration and abnormal mitochondrial morphology. This observation led to the discovery that pink1 and parkin function within the same pathway controlling mitochondrial integrity. Moreover, genetic analysis in *Drosophila* places pink1 upstream of parkin, further establishing mitochondrial dysfunction as a major player in PD pathogenesis [29].

For dominant mutations, models usually rely on the expression of transgenes driven by Gal4/upstream activating sequence (UAS) systems, allowing time- and tissue-specific expression of
any gene of interest in *Drosophila*. Expression of mutant or wild-type forms of human cDNA in the CNS could be achieved using CNS-specific Gal4 drivers [30]. A wide variety of models have been made using this system, including a plethora of neurodegeneration models. However, expression of human genes in all neurons could compromise the overall survival of the flies. Thus, particularly for the study of neurodegeneration, a number of models have been made expressing the neurodegenerative genes in the fly compound eye, using eye-specific Gal4 drivers. The advantage of this approach is that it bypasses any possible compromised survival due to CNS expression, while allowing for the rapid quantification of neuronal degeneration in the context of the compound eye. This degeneration leads to rough eye phenotypes. These can be easily quantified using standard microscopy techniques, allowing for the use of these phenotypes in high-throughput genetic screens. Indeed, screens for enhancers and suppressors of rough eye phenotypes as indicators of neuronal toxicity induced by expression of mutations linked to neurodegeneration genes have been very fruitful in identifying novel genes and pathways involved in neurodegeneration. Usually, modifiers obtained using this approach are validated in other models, including *Drosophila* pan-neuronal models, to avoid the possibility of modifiers affecting eye phenotypes specifically. An excellent example of a *Drosophila* genetic screen using this approach is the initial identification of modifiers of the polyglutamine disorder spinocerebellar ataxia type 1 (SCA1) mediated by polyQ expansion in the gene ataxin-1. In this study, transgenic flies expressing human cDNA carrying a pathogenic ataxin-1 polyQ expansion expressed in the fly retina produced a dose-dependent eye degeneration. This model was used to screen for enhancers and suppressors of disease using both loss of function and over-expression alleles. The screen uncovered novel modifiers involved in glutathione-mediated cellular detoxification, transcriptional regulation, and RNA processing, revealing additional pathogenic mechanisms in SCA1 [31]. More recently, the development of *Drosophila* RNAi libraries has allowed for the genome-wide screening of loss of function alleles in enhancer and suppressor screens.

Overall, the advantages of using *Drosophila* as a model system are the unique genetic tools available, coupled with the breadth of phenotypic outcomes that can be measured. Flies are also relatively cheap to maintain and results from screens can be obtained within weeks. It is estimated that approximately 75% of human disease genes have orthologs in *Drosophila* [32], which is slightly more than the estimated number in *C. elegans* (~65%). This suggests that many gene functions are conserved between humans and these invertebrate organisms. However, despite its genomic similarities, *Drosophila* is still an invertebrate with a very different life cycle to humans and other mammals, including a developmental process with other developmental stages: embryo, larva, pupa, and adult. Taken together, these differences between flies and humans can lead to phenotypic outcomes in disease models that may not reflect the human condition. Thus, it is critical to acknowledge that *Drosophila* are not perfect models for human diseases, and as in all other models, the fly is only able to model certain aspects of the complexity of human diseases. In this context, fly models can be extremely useful in the elucidation of pathogenic mechanisms leading to neurodegeneration, but a thorough understanding of the limitations of the models is also critical to maximize their potential.

**Danio rerio**

The zebrafish, *Danio rerio*, has classically been used to study developmental processes as the embryos are small, transparent, and undergo rapid development *ex utero*, allowing for *in vivo* analysis of embryo- and organogenesis. Their generation time (2–3 months) and the ability to produce large numbers of eggs make them an excellent model for high-throughput functional studies.
Forward genetics screens in zebrafish usually use the chemical mutagenic agent N-ethyl-N-nitrosourea (ENU). ENU is a potent mutagenic agent that produces mainly point mutations genome-wide. Thus, phenotype-driven screens rely on the mutagenic actions of ENU to produce fish carrying genome-wide mutations that may affect prospective phenotypes. These screens have yielded an impressive array of new mutants in a variety of disease areas [33]. Over the last decade, zebrafish has also emerged as an attractive vertebrate organism to model human disease, including AD, PD, HD, and ALS [34]. Examples of its use include the discovery that TDP43 and FUS, two RNA-binding proteins mutated in ALS, function within the same genetic pathway, suggesting that TDP43 lies upstream of FUS [35].

Currently, disease modeling in zebrafish is particularly advantageous when phenotypes can be assessed during embryogenesis or larval stages, including high-throughput drug screening. The development of novel genome-editing technologies such as zinc finger nucleases (ZFN, TALENs) and more recently CRISPR/Cas9 systems, together with established techniques such as targeting induced local lesions in genomes (TILLING), will allow for the rapid production of fish lines carrying a wider variety of alleles. These could include loss of function as well as engineered point mutations in multiple genes together with endogenous reporter lines. These novel technologies are likely to make zebrafish an even more attractive organism in which to model neurodegeneration in the near future.

Mus musculus

Due to the nature of the high-throughput design of genetic screens in cellular systems and small animals, it is challenging to test other disease phenotypes apart from aggregation potential and toxicity. Therefore, newly identified modifiers from cellular or non-mammalian model organisms are usually validated in mammalian models, particularly in mice. Using these models, it is possible to assess not only cellular aggregation and toxicity, but also the complex interaction between neurodegeneration and its consequences on different neuronal populations together with other possible non-neuronal effects. The similarities between the rodent and human brain allow for the direct comparison of most neuropathology findings, making possible the direct mapping of complex behaviors, including motor and cognitive functions, to the equivalent human neuronal populations. Moreover, they also allow for the assessment of systemic interactions between the brain and all other organs, such as muscle or the endocrine system.

Mice are not commonly used in genetic screens to find modifier genes, mainly due to cost and time constraints. They are largely used for the in-depth analysis of pathology, as well as in pre-clinical studies. However, there are a few examples of phenotype-driven screens in mice, in particular using ENU mutagenesis, in order to find modifiers for a particular disease [36].

Mice and humans share the great majority of their genome, with over 95% of human genes estimated to have a murine counterpart. This commonality leads to very similar organogenesis, including a mouse brain that shares its major anatomical architecture with the human brain. Despite its obvious complexity, the mouse brain is less complex, with an estimated 70 million neurons compared to the 86 billion estimated in the human brain. All this, together with the wide variety of alleles available [knock out (KO), transgenics, point mutants (knock in [KI], ENU-induced), insertions, deletions], makes the mouse the current model organism of choice to understand neurodegeneration pathobiology. The mouse is also widely used as a pre-clinical model to assess possible disease therapeutics.

The mouse is used for disease modeling due to its relatively quick generational rate (3 months) and the unique ability to manipulate its genome. This includes genetic manipulations of embryonic stem (ES) cells as well as the possibility of producing transgenic animals via pronuclear
injection. Mice are widely used to model all aspects of disease biology with large numbers of laboratory mice being bred around the world. This allows for the non-systematic identification of spontaneous mutants leading to obvious phenotypes such as skin and hair mutations or locomotor abnormalities. Examples of these spontaneous mutations include the wobbler mouse model of ALS carrying a mutation in \(\text{Vps54}\) [37] and the tremble mutants carrying \(\text{Pmp22}\) mutations [38]. Systematic efforts aiming at the identification of novel genes in a wide variety of disorders have taken advantage of mutagenic agents such as ionic radiation or ENU mutagenesis to produce genome-wide mutations [39]. These forward genetics screens have been very successful in identifying novel genes involved in a wide variety of human diseases, including CNS disorders [40]. These include a plethora of novel mutants harboring mutations in genes affecting motor function, including locomotion phenotypes, free wheel running activity, and tremor development. However, these screens have not been as successful in identifying novel genes involved in learning and memory and cognition. This is likely due to the intrinsic high degree of behavioral variability between individual mice. Also, tests assessing cognition in mice are time consuming and usually require extensive mice training, making them less amenable for high-throughput phenotypic analysis. Normally, these phenotype-based screens do not leave mice to age, and thus mutations likely producing progressive phenotypes such as neurodegeneration are probably under-represented. However, to our knowledge, there is at least one large-scale ENU mutagenesis program, ongoing at UK MRC Harwell, with an aging phenotyping pipeline.

Also, a systematic effort aimed at the generation and characterization of null mutations for all mouse genes within the next decade is already underway. The international mouse phenotyping consortium (IMPC) is an international collaboration aiming at assigning phenotypic functions to all genes in the mouse genome. As a result of this project, null conditional alleles for all mouse genes would be available by 2020, giving the research community unlimited access to mice or ES cells harboring conditional null mutations for any particular gene. Moreover, the phenotypic characterization of all these null lines will allow, for the first time in a mammalian organism, phenotype-based characterization of an entire genome.

There is a plethora of mouse models for a wide variety of neurodegenerative conditions. As is the case for other animal models, the most commonly used approach to model dominant mutations in the mouse entails the ectopic expression of human transgenes. These usually express wild-type or mutant forms of neurodegeneration genes under the control of endogenous or neuronal enriched promoters. These ectopic expression systems have led to the production of the most used models, including a variety of SOD1 transgenic models for ALS [41]; APP, Tau, and PS1 transgenics, and their combinations to model AD [42]; and N-terminal polyglutamine expanded huntingtin transgenic models of HD [43]. As with other model organisms, these transgenic mice usually express wild-type or mutant forms of human-derived cDNA, therefore lacking introns and their possible regulatory sequences, as well as the endogenous 3′ untranslated region. This is corrected with transgenic models expressing bacterial artificial chromosomes (BAC), which include the entire gene structure. However, both cDNA and BAC transgenes are usually integrated randomly in the mouse genome, with their expression largely dependent upon integration sites and number of transgene copies present. They also lack the possible regulatory elements acting in the endogenous locus due to its physical position. Moreover, they both express wild-type or mutant forms of neurodegeneration genes in the context of the wild-type allele of the mouse endogenous gene that could potentially influence disease pathobiology. This can be corrected by the use of a null mouse background, leading to a humanized model in which the human wild-type or mutant form of a neurodegeneration gene is the only form of the gene present. However, neurodegenerative genes are usually ubiquitously expressed and have critical roles
outside the CNS, and thus the use of null backgrounds may also have consequences for disease pathobiology.

Most human mutational events are point mutations. Mice carrying point mutations in the mouse endogenous genes are genetically and biochemically the closest models to human conditions. Mice carrying point mutations can be obtained either by homologous recombination (KI) or by ENU mutagenesis [40]. The majority of the relatively few examples of mice carrying point mutations in neurodegeneration genes usually lead to late disease symptoms, making them less attractive for disease modeling. Emerging genome-editing technologies, such as CRISPR/Cas9 technology, are paving the way for the humanization of entire genes, replacing an entire mouse gene with its human counterpart [44], leading to even closer models, biochemically, to human conditions.

The adoption of the mouse as a model for genetic screens has been relatively slow. This has particularly been the case because of expense and timeframe constraints, making such screens impractical for most laboratory settings. Although some phenotype-based modifier screens using ENU mutagenesis over transgenic models of neurodegeneration are ongoing, including, to our knowledge, HD, ALS, and AD models, none has yet been published. However, mouse modifier screens have been published in other disease areas, including other CNS disorders. An excellent example is a dominant ENU suppressor screen using Mecp2 KO mice that established a novel role for cholesterol metabolism in Rett syndrome, an autistic spectrum disorder [36]. The findings from this screen exemplify the power of using the mouse as a model organism for genetic screens, as the mutants obtained can inform of disease pathogenesis in a mammalian system with direct correlation with the human condition. Mecp2 null males are normal at birth and weaning but develop limb clasping, tremors, lethargy, and abnormal breathing, which progressively worsen until death occurs between 6 and 16 weeks of age. In this example, male Mecp2 KO mice carrying dominant ENU mutations were screened for disease suppression using a health scoring system that included assessments of limb clasping, tremors, body size, cage activity, and development of skin and/or eye inflammation. Using this phenotyping system, a number of mutant mice were identified as suppressors. After inheritance testing and positional cloning, one of the suppressors was identified as a nonsense mutation in Sdhe, which encodes a critical enzyme for cholesterol metabolism. This finding suggested that cholesterol metabolism may be compromised in models of Rett syndrome, which was indeed confirmed in another Rett syndrome mouse model. Moreover, the finding that cholesterol metabolism was compromised in Mecp2 null mice directly suggested the use of drugs aimed at inhibiting cholesterol synthesis, such as statins, to treat Rett syndrome. Indeed, statins were beneficial in Mecp2 null mice when administered in low doses, but detrimental in higher doses.

Thus, modifier screens in the mouse have the potential to identify unrecognized pathways of pathogenesis in a mammalian system, potentially providing novel genes and pathways that could be targeted for the design of novel therapeutic strategies. On the downside, the cost and time constraints of genetic screens in mice make them impractical for most laboratory settings. Also, despite commonalities, there are also clear differences in neuroanatomy between mice and humans, for example in the frontal cortex. Other major differences include lifespan and size. Neurons in humans need to survive for decades and usually only show signs of degeneration in mid-life, after 50 years of age. However, in mice, this could translate to approximately 1.5–2 years of age. Also, the size difference could be critical. For example, axons of human motor neurons can be as long as 2 meters, potentially leading to specific problems associated with their maintenance. Despite its drawbacks, the mouse is the mammalian model of choice to validate previous findings from other models or screens, and pre-clinical studies of drug testing.
Human Cellular Models and Post-mortem Material

The overarching objective of studying diseases using model organisms is to provide mechanistic insight into disease pathobiology that is not possible with human material alone with the ultimate aim of providing etiology-based therapies. This relies on a detailed understanding of the molecular pathways involved in the disease processes in different model organisms. These genes and pathways, identified using different disease models, ultimately need to be validated using human material. Until recently, for neurodegeneration research, these patient-derived materials had been limited to post-mortem tissue together with cerebrospinal fluid (CSF) and easily accessed cell lines such as lymphoblastoid and fibroblast cell lines. These limitations have led to significant problems. Work using neuronal human material has traditionally been limited to end-stage disease using post-mortem samples. We have not been able to systematically study the early pathological changes in human patients that lead to disease that would be key to understanding how the disease starts and progresses, and to the design of novel etiology-based therapies. The emergence of non-invasive imaging techniques such as bio-markers for disease progression in neurodegeneration is likely to have a huge impact. However, in the great majority of sporadic cases of the major neurodegenerative disorders patients are only diagnosed when the disease process is already well underway. Therefore, longitudinal studies focusing on the early molecular events leading to disease onset and progression are generally only possible in model organisms.

With the development of techniques using iPS cells, aimed at differentiating cellular populations including human skin fibroblasts into neurons, it is now possible to produce patient-derived neuronal populations. These iPS cells have a huge potential, allowing for the first time the study of cellular disease mechanisms using differentiated human neuronal populations. Moreover, as they are patient-specific, they are excellent models to study the heterogeneity of disease processes leading to neuronal toxicity. As with other cellular disease models, their utility relies on how the findings using human derived iPS cells are finally related to human neurodegeneration. Although human in origin, iPS cells are living not in the context of a fabulously complex mammalian brain, but in the confines of a laboratory setting. Despite this limitation, their potential to further our understanding of disease pathobiology is immense. This includes their potential use in cellular phenotype-based genetic screens using siRNA libraries to identify enhancers and suppressors of prospective cellular phenotypes due to patient-specific mutations.

The Future

Genetic screens have contributed greatly to our current understanding of neurodegenerative disease pathobiology. However, the knowledge gained by this and all other approaches has generally not yet led to the discovery of disease-ameliorating drugs in human neurodegeneration. In this sense, neurodegeneration research is lagging behind other fields, such as cancer, where a deeper understanding of disease pathogenesis has led to the development of new therapeutics. The gap between the clinic and the laboratory is generally wider for neurodegeneration research. Still, there are some aspects to consider when moving from small organisms to mammals and humans. Among them are the differences in metabolism, whether the drug might pass the blood–brain barrier, the dose, side effects and, also important, the genetic variability found in humans that is not reproduced in those studies. This perhaps reflects the fact that we still fundamentally do not understand how neurodegeneration occurs. Thus, a deeper understanding of disease pathogenesis is probably required to better guide the rational design of novel therapeutics.
Model organisms play a critical role in this translational research cycle, and therefore understanding disease pathogenesis requires the continued development of novel models that better reflect the human condition. A much deeper understanding of disease pathogenesis in the already established models is also needed. The emergence of novel genome-editing technologies, such as CRISPR/Cas9, that could be used in a wide variety of model organisms is likely to provide an invigorating emphasis to the development of novel disease models that better mimic particular aspects of the human conditions. In this light, genetic screens are likely to continue to contribute towards a deeper disease understanding. However, it is critical to acknowledge each model’s limitations and to combine their use with research using human material to try to maximize the strengths of each model organism.

The use of an integrated approach, focusing on the strengths of different models to identify and further validate novel genes and pathways involved in disease pathobiology, can be exemplified by an excellent example using models of SCA1. Using this approach, parallel cell-based and Drosophila genetic screens were employed to identify novel, druggable targets for SCA1 [45]. The authors found that down-regulation of different components of the RAS–MAPK–MSK1 pathway modulated ataxin-1 expression levels in SCA1 cellular and Drosophila models, which directly correlated with disease onset. They further validated their findings in the mouse by crossing SCA1 mouse models with Msk1 and Msk2 heterozygote null mice. Moreover, pharmacological inhibitors of the pathway also lead to a down-regulation of ataxin-1 levels in cellular and organotypic SCA1 models. This integrated approach starting with different model systems, in this case parallel cell-based and fly genetic screens, aimed at identifying common pathways is likely to identify common toxicity mechanisms that may be better translated to the human conditions.

The potential use of iPS cells in this kind of integrated approach coupled with the strengths of genetic screens in different model organisms, from yeast to mice, is likely to increase our understanding of disease pathogenesis while continuing to provide novel druggable targets in the future. While the path towards new disease treatments is certainly very complex, we are clearly a lot further along the road to potential cures than a couple of decades ago. The continued identification of new genes and the use of genetic screens to further elucidate disease pathobiology can only take us closer to this elusive goal.

Acknowledgments

We are grateful to the MRC for funding, Steve Thomas for figure development, and Angela Hoslin for critical reading.

References


