CHAPTER 6

Development of an iPSC-based model of ataxin-3 aggregation

M. Minoia\textsuperscript{1*}, A. Thiruvalluvan\textsuperscript{2*}, H.W. Boddeke\textsuperscript{2}, H.H. Kampinga\textsuperscript{1}, J.C. Copray\textsuperscript{2}

* shared first authors

Manuscript in preparation
Abstract

SCA3 is a neurodegenerative disorder caused by expansion of polyglutamine (polyQ)-encoding CAG repeats in the ATXN3 gene. Like in all polyQ diseases, the CAG repeat length is inversely associated with the age at onset (AO) of the disease. Likewise, CAG-repeat length is proportionally related to the aggregation propensity of the ataxin3 protein, strongly suggesting that aggregation formation is a main disease-inducing factor. However, in patients with the same CAG repeat length, AO can vary up to more than 40 years, suggesting that other genetic or environmental factors must be involved. Studying such factors requires patient representative models. The recently discovery of the ability to reprogram fibroblast to pluripotent stem cells (iPSC) and the subsequent strategies to develop neuronal-like cells from them allows the studying the aggregation behaviour of the polyQ expanded ataxin3 protein within the patient’s own genomic context. As a proof-of-concept, we here generated iPSC-derived neuronal lines from 3 SCA3 patients with similar CAG repeat length but with different AO and analysed ataxin3 aggregation behaviour and measure HSP levels. We show that iPSC and neuronal generation is possible and that HSP expression levels seem to differ between the different SCA3 iPSC-derived neurons.
Introduction

Spinocerebellar ataxia type 3 (SCA3), also known as Machado–Joseph disease (MJD), is the most common spinocerebellar ataxia and the second most common polyglutamine (polyQ) disease after Huntington’s disease (HD) [1,2]. SCA3 is an autosomal dominant neurodegenerative disorder characterized by neuronal loss in the cerebellum, thalamus, midbrain, pons, medulla oblongata and spinal cord of patients [3,4]. SCA3 is clinically heterogeneous, but the main feature is the progressive ataxia, which affects balance, gait and speech. Current therapeutic strategies are only able to provide symptomatic relief [5]. SCA3 is caused by an expanded stretch of CAG triplets in the coding region of the ATXN3 gene. Healthy individuals have up to 44 CAG repeats, whilst affected individuals have between 52 and 86 glutamine repeats [6]. There is a clear correlation between CAG repeat size and age of onset [7]. Likewise, aggregation propensity is strongly related to the CAG repeat length for all polyQ diseases irrespective of the gene affected, strongly suggesting that a toxic gain of function aggregation due to the polyQ expansion is the driving force initiating disease. However, more and more data suggest that the full-length protein is not aggregating until a triggering event activates proteases (caspases, calpains) that lead to cleavage of the full-length protein into smaller, aggregation prone fragments [8–11].

Importantly, whilst CAG repeat length accounts for 50-65% of the variance in AO, disease onset, especially for shorter repeat lengths, can vary up to 30 years for SCA3 patients with the same CAG repeat [7]. Similar observations are found in other polyQ disease, which suggests that AO in the polyQ diseases might be modified by common environmental (disease triggering) as well as genetics (fragility) factors. Over the last decade several potential modifiers for SCA3 pathology were identified using cell- and animal models [3,12–14]. Many of these modifiers are components of the cellular protein quality control (PQC) system such as heat shock protein (HSP) expression, proteasomal activity or autophagosomal activity [12–15]. These findings led to the hypothesis that differences in HSP expression between SCA3 patients could be a genetic co-factor that modulates the AO independently of the CAG repeat length [16,17]. However, to more directly test whether or not PQC components are bona fide modifiers of AO, current model systems are not well suited and we urgently need patient-representative neuronal models.

Reprogramming of differentiated somatic cells into a pluripotent state (induced pluripotent stem cells, or iPSCs) has opened a treasure trove of opportunities for research on human diseases. It is now possible to generate iPSCs from SCA3 patients and use these cells as starting material for differentiation into disease-affected cell types [10]. Following this assumption, we generated iPSC-lines from two patients with similar CAG expansions but clearly different clinical AO and derived neurons from these iPSCs and tested whether protein aggregation and HSP expression can be related to CAG-independent variations in the AO of SCA3 pathology.
Results

Generation of iPSCs from SCA3 patients.

To understand why CAG repeats expansion in the ATXN3 gene causes the selective degeneration of human neurons, we generated iPSCs using dermal fibroblasts from two SCA3 patients (SCA3-1 and SCA3-2) and one control subject (Ctrl) who is unrelated and unaffected. Dermal fibroblasts cultured from skin punch biopsies were reprogrammed to iPSCs using episomal vectors expressing human Oct3/4, Sox2, Klf4, c-Myc, Nanog and Lin 28 [18]. We decided to use a non-viral integration-free reprogramming methods for generating iPSCs to avoid potential genomic integration of vectors or transgenes that may lead to aberrations in unrelated developmental processes [19]. Multiple clones with human embryonic stem cell (hESC) morphology were picked for all subjects. After expansion and initial characterization, one representative clone for each subject was used for further studies. As shown in Figure 1 B, all lines of iPSCs exhibited morphology indistinguishable from human embryonic stem cells (hESCs) and can be maintained indefinitely on matrigel. Pluripotency markers such as Oct4, Nanog, SOX-2, SSEA4, Tra-1- 60 and Tra-2-54 were strongly expressed in all lines (Fig. 1 B).

Notably, it has been recently shown that the process of reprogramming somatic cells to iPSCs could induce genome alterations such as copy number variation (CNV) [20,21]. Because of the fact that genome stability can have severe effects on pluripotency and differentiation of resulting iPSCs [22], we performed whole-genome SNP sequencing to investigate the possible CNV instability during somatic reprogramming in iPSCs. CNVs were detected in two of the iPSC lines, of which five CNVs were in control (Fig. 1 C-E) and two in SCA3-1 iPSC lines, respectively (Fig. 1 C). However, the low level of CNVs detected in our iPSC-derived neurons did not appear to affect the differentiation or function of these neurons (Fig. 2). Yet, in an iPSC clone derived from another SCA3 patient, we did find multiple genomic rearrangements meaning that we could not include it in our further analyses. Another iPSC clone of that same patient, however, was recently found to be normal, demonstrating that the reprogramming per se was not the cause for the genomic instability in this line.
Figure 1. Generation of induced pluripotent stem cells (iPSCs) from a control subject and two SCA3 patients. (A)
Clinical data of the two SCA3 patients included in the study. (B) Phase contrast images of control, SCA3-1 and SCA3-2 iPSC lines and staining of the lines with pluripotency markers SSEA4 (SSEA3 in SCA3-2 cells), Oct4, SOX-2, Tra-2-54, Tra-1-60 and Nanog (Tra-1-80 in SCA3-2 cells). (C) Overview of copy-number variation (CNV) analysis from DNA of control, SCA3-1 and SCA3-2 iPSC lines. (D-E) Chromosome view of genome wide genotyping data for control iPSC line. (D) Chromosomes 7 and 15 show duplications as indicated by aberrant SNP profiles (B allele frequency) and the increased log R ratios (lower panel) highlighted within the red box. (E) Chromosomes 5 and 11 exhibit deletions as indicated by the homozygous SNP profiles (B allele frequency) and reduced log R ratio highlighted within the red box.
Differentiation of iPSCs to neurons.

Various protocols have been developed for the differentiation of human embryonic stem cells into a neuronal lineage by recapitulating the developmental steps from neural tube into the neuronal repertoire. We also started the neuronal differentiation of hIPSCs by inducing their differentiation into columnar epithelial cells (neural rosettes), representing neural tube cells and by priming them with the caudalizing factor retinoic acid (RA) (Fig. 2A-C). Neural rosettes were handpicked and cultured in the presence of bFGF and EGF as spheres and maintained as neural stem cells (NSC) (Fig. 2 A). Next, NSCs were dissociated and cultured in the presence of sonic hedgehog (SHH), the neurotrophic factors BDNF and GDNF, and N1 and B27 supplements essential for neuronal differentiation and their subsequent maturation (Fig. 2 A-C). Directed differentiation of neural stem cells into neurons inevitably gives rise to a mixed population of glial restricted progenitors and different neuronal subtypes. As we intended to focus our study on the aggregation sensitivity of SCA3 patient-derived neurons upon glutamate stimulation, we purified the neuronal culture by multistep FAC-sorting as previously described [23] using the cell-surface marker CD184+ and CD44- as signature for neural stem cells, CD184+/CD44+/CD24- for glia cells, and CD184-/CD44-/CD24+ for neurons. After sorting, neurons (CD184-/CD44-/CD24+) were plated and grown for 15-20 days for further axon regeneration and maturation and then used for further analyses.

Western blot analysis confirmed the expression from both alleles of normal ATXN3 (∼41kDa) in purified suspensions of control iPSC-derived neurons and normal and expanded ATXN3 (∼60kDa) in SCA3 iPSC-derived neurons (Fig. 2 D).

It has been suggested for HD iPSC-derived NSC lines that the CAG repeat length may increase by up to 10% increase after repetitive passages [24]. For this reason, we analysed CAG repeat length over the time during reprogramming and differentiation steps. For all the iPSC lines no changes in polyQ repeat length were detected, the normal and expanded CAG repeat alleles did not exhibit instability with passages or upon differentiation in NSCs (Fig. 2 E). Thus the CAG repeat length seems stable during reprogramming, long term passage, and differentiation.

Ataxin-3 aggregation

Expansion of ATXN-3 polyglutamine (polyQ) tract is believed to lead to a toxic gain of function, caused by calpain-dependent proteolysis of the mutant ataxin-3 and the consequent generation of expanded polyQ fragments, which are prone to form insoluble aggregates [9,10].

To monitor aggregation in iPSC-derived neurons, we stimulated the neuronal cultures with the excitatory neurotransmitter L-glutamate (0.1 mM) or left them
Figure 2. Directed differentiation of the iPSCs into neurons in vitro.

(A) Time line and face contrast microscopy images of iPSC-differentiation into neural rosettes, neurospheres and neurons. (B,C) iPSC-derived neurons co-stained for β-III tubulin and nuclei, in figure B, and for ataxin-3, nuclei and MAP2 in figure C. (D) Western blots of ataxin-3 expression in iPSC-derived neurons show normal (bottom arrow) and mutant (top arrow) ataxin-3. (E) CAG repeat analysis in control and SCA3 fibroblasts, iPSCs and NSCs.

<table>
<thead>
<tr>
<th>CAG repeat length</th>
<th>Fibroblast</th>
<th>iPSC</th>
<th>NSC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctrl</td>
<td>14/26</td>
<td>14/26</td>
<td>14/26</td>
</tr>
<tr>
<td>SCA3-1</td>
<td>19/71</td>
<td>19/73</td>
<td>19/72</td>
</tr>
<tr>
<td>SCA3-2</td>
<td>14/76</td>
<td>14/76</td>
<td>14/76</td>
</tr>
</tbody>
</table>

Figure 3. Ataxin-3 aggregations.

Representative western blots of fractionated ATXN-3 in control, SCA3-1 and SCA3-2 iPSC-derived neurons treated with glutamate or untreated. SDS-insoluble aggregates are seen in SCA3-1 and SCA3-2 stacking gels.
unstimulated. Immunoblotting of extracts from SCA3 iPSC-derived neurons revealed in the stacking SDS-insoluble ATXN3-containing aggregates, a hallmark of SCA3 pathology. No SDS-insoluble ATXN3-containing aggregates were detected in control neurons (Fig. 3). Unlike what was published before [10], aggregates were also present in SCA3 neurons not stimulated with L-glutamate (0.1 mM) and glutamate treatment did not lead to further increase in the amount of SDS-insoluble ATXN3 aggregates (Fig. 3). Within the limits of these initial data, there was no difference detectable in the amount of aggregation between the lines, meaning that aggregate formation was not related to the differential AO of the respective patients from which the iPSC derived neurons were generated. Obviously, however, more extensive analyses will be required with more lines to substantiate this.

Expression levels of HSPs in iPSC-derived SCA3 neurons

The family of heat shock proteins (HSPs) was initially characterized as a highly conserved battery of genes whose expression can be induced by proteotoxic stress and that could be related to handling protein aggregates.

The human HSP family is remarkably large, with over 13 HSPA, 41 DNAJ and 11 HSPB members in the genome [13]. Various members within the class of HSP families have been shown to associate with poly-Q inclusions in vivo, and their overexpression can affect aggregation rates, toxicity and disease progression [16,17,25–27]. Therefore, we hypothesized that differences in HSPs expression might be a modifying factor of pathogenesis in patients with SCA3 and expression could be related to AO. We measured expression levels of different HSP family members in control and SCA3-1 neurons. Yet a limited set of HSPs were analysed due to limited availability of cell material and specific antibodies. No significant changes in HSPA1A/Hsp70, HSPC/Hsp90 and DNAJ/Hsp40 levels were detected (Fig. 3 A) indicating that expression of the mutant SCA3 protein per se did not induce a stress response, consistent with earlier data in patient fibroblasts (Zijlstra et al) and in human SCA3 brains (Seidel et al). In fact, the levels of two members of the HSPB family, HSPB1 and HSPB8, were actually found to be lower in SCA3-1 patient neurons compared to the control cells (Fig 3 A). We next compared HSPs expression levels and insolubilization between the two SCA3 patients. SCA3-2 neurons showed lower HSPA1A protein levels compared to the SCA3-1 line and Hsp70 became associated with the SDS-insoluble fraction only in the SCA3-1 line (Fig. 4 B). In addition, while HSPB1 was significantly upregulated in the SCA3-1 line, it was not detectable/absent in the SCA3-2 line. It is interesting to note that the SCA3-2 cells showing low/undetectable HSPB1 and HSPA1A levels were from the patient that showed an earlier AO of the disease (Fig. 1 A), which may suggest a correlation between the levels of these chaperone and AO (although our limited analysis does not yet allow firm conclusions).
Discussion

Recently, human induced pluripotent stem cells (iPSCs) have emerged as a useful tool for research on the pathogenesis of many neurodegenerative diseases [22,28]. An increasing number of studies have employed iPSCs derived from patients with neurological diseases, with many focusing on their potential use to understand the cellular mechanisms behind the pathogenesis of these disorders [10,24,29–31]. Previous data regarding the molecular pathogenesis of neurodegenerative diseases like SCA3 have come from investigations utilizing cells and animal models and post-mortem brain tissues, that all have their limitations, e.g., especially when trying to understand why different patient with comparable CAG repeat sizes in their affected ATXN3 genes show such large difference in age at onset (AO).

In line with what has been previously reported [10,24], we confirm that it is possible to generate disease-specific iPSCs from SCA3 patients and differentiate them into neurons (Fig. 1, 2) that show aggregation related to the expression of the expanded allele (Fig. 3, 4). However, unlike what was found by Koch and co-workers [10],

Figure 4. HSP expression levels in control and SCA3-derived neurons.
(A) Representative western blots of ATXN-3, β-III tubulin, GFAP, GAPDH and selective members of HSP family in control and SCA3-1 treated with glutamate or untreated. (B) Representative western blots of ATXN-3, β-III tubulin, HSPA1A and HSPB1 SCA3-1 and SCA3-2 treated with glutamate or untreated.
we find that glutamate-induced excitation is not critical under our culture conditions to induce aggregation of ATXN-3 (Fig. 3). The reasons for this can be manifold. First, the protocols used for generating neurons are longer than previously reported. Duration of differentiation towards neuron was prolonged to increase the amount of glutamate receptors and other neuronal characteristics in order to elucidate maximum response (ataxin aggregation and HSP response) upon glutamate treatment. NSC were derived independently for each experiment from each SCA and control iPSc line to avoid long-term culturing effect, such as chromosomal aberrations and biased neuronal differentiation towards specific subtype. This could imply that more mature and similar population of neurons derived from the neural progenitors are used here, excitotoxic triggers may not be required to induce ATXN3 aggregation. Moreover, it is important to note that we sorted our neuronal population after 80-90 days in culture and so eliminated any possible buffering capacity of astrocytes. Also, we omitted growth factors during the last 10 days before the aggregation assay, which may have stressed neurons in a similar manner as glutamate stimulation inducing ATXN3-cleavage and its consequent aggregation. So, the requirement for an external trigger to induce aggregation still needs further investigation.

In how far the iPSC-derived neurons are a suitable model to study the genetic co-factors that modulates the AO in SCA3 patients requires further investigation. Our preliminary data on components of the PQC system do not allow any conclusion in this respect yet. However, in line with data in fibroblast from the same patients (Zijlstra et al) and in line with cell model data (Ref) and data from post mortem human brains (Seidel), we find no evidence of an up-regulated stress-response in neurons expressing the expanded polyQ protein (Fig. 4). In fact, in one line, HSPB1 actually seems to be expressed at a lower level. Furthermore, the neurons from the SCA3 patient with the early onset (SCA3-2) express lower levels of HSPA1A and HSPB1 than the neurons from the SCA3 patient with the average age at onset (SCA3-1), which could be taken as an indication that these HSP levels may play a modulating role in (CAG-size independent) SCA3 pathogenesis. However, more lines and more extensive analyses are needed to substantiate these suggestions.

In all, however, our feasibility study is encouraging and suggests that patient-derived iPSC lines may become a valuable tool in SCA3 research.

Materials and methods

Human subjects. The experiments were undertaken with the understanding and written consent of each subject, and were been carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki) for experiments involving humans. Fibroblast samples were obtained from one healthy
individuals and two clinically affected and genetically confirmed Dutch patients with SCA3 (Fig. 1A). Subjects were randomly approached for participation in this research. A single experienced neurologist determined the AO in all patients as the age at which the first clinical manifestations of unsteadiness of gait and stance were unmistakably present. The patient group contained heterozygotes only.

**Generation of iPSCs with episomal vectors.** Human dermal fibroblasts (HDFs) were cultured in Dulbecco’s modified Eagle media (DMEM, Gibco) containing 10% fetal bovine serum (FBS), 1 mM non-essential amino acids (NEAAs), 1× GlutaMAX, and 100 unit/ml penicillin with 100 μg/ml streptomycin. The episomal iPSC reprogramming plasmids, pCXLE-hOCT3/4, pCXLE-hSK and pCXLE-hMLN were purchased from Addgene. The plasmids used in our experiments were mixed in a ratio of 1:1:1 for efficient reprogramming. Three micrograms of expression plasmid mixtures were electroporated into 5× 105 HDFs with Amaxa® Nucleofector Kit according to the manufacturer’s instructions. After nucleofection, cells were plated in DMEM containing 10% FCS and 1% penicillin/streptomycin until it reaches 70-80% confluence. The culture medium was replaced the next day by human embryonic stem cell medium (HESM) containing knock-out (KO) DMEM, 20% KO serum replacement (SR), 1 mM NEAAs, 1× GlutaMAX, 0.1 mM β-mercaptoethanol, 1% penicillin/streptomycin, and 10ng/ml bFGF (Invitrogen). Between 26-32 days after plating colonies developed and colonies with a phenotype similar to human ESCs were selected for further cultivation and evaluation. Selected iPSC colonies were mechanically passaged on matrigel (BD, hES qualified matrigel) coated plates containing mTeSR™1 (defined, feeder-free maintenance medium for human ESCs and iPSCs).

**Generation of iPSC-derived neural stem cells.** iPSCs were dissociated manually and plated on a non-coated dish in human embryonic stem cell medium (HESM). After 4 days, embryoid bodies (EBs) were formed and transferred to neural differentiation medium containing DMEM/ F12, 1 mM NEAAs, 1× GlutaMAX, 1% penicillin/streptomycin, and 1× N1 supplement (100X) for another 4 days. EBs were plated on matrigel-coated plates for neural rosette formation for 8-10 days with 0.01mM retinoic acid. Neural rosettes were handpicked and cultured in neural stem cell medium containing DMEM/ F12, 1 mM NEAAs, 1× GlutaMAX, 1% penicillin/streptomycin, 1× N1 supplement (100X), 20 ng/mL FGF2 (peprotech), 20 ng/mL EGF (peprotech), and 2μl/ml B27 supplement (Invitrogen).

**Neural differentiation of neural stem cells.** Terminal neural differentiation was induced by dissociating the neural stem cells (NSCs) using accutase (Sigma) for 20 min at
37ºC and plating them on a matrigel-coated plated for attachment. The next day, the medium of these cell cultures were changed to neuronal differentiation medium containing DMEM/F12, 1 mM NEAAs, 1× GlutaMAX, 1% penicillin/streptomycin, 1× N1 supplement (100X), 20 ng/mL BDNF (Peprotech), 20 ng/mL GDNF (Peprotech), 50 ng/mL SHH (Peprotech), 1 mM dibutyryl-cAMP (Sigma) and 2 μl/ml B27 supplement (Invitrogen) for 80-90 days. Differentiated neurons were purified by multistep FAC-sorting as previously described [23] using cell-surface marker signatures for the specific isolation of neural stem cells, glia and neurons, respectively CD44, CD184 and CD24.

Immunocytochemistry. IPSCs derived from SCA patients, control iPSCs and in vitro differentiated neurons were fixated with 4% paraformaldehyde for 20 min. Cells were blocked in 5% normal goat serum and 2% Fetal calf serum; subsequently, samples were probed with primary antibodies: SSEA-4 (1:500), TRA-1-60 (1:500), TRA-2-54 (1:500), OCT-4 (1:500), Sox-2 (1:500), NANOG (1:500), Ataxin-3 (1:300), MAP-2 (1:500), βIII tublin (1:500). Alexa 488 and Cy3-conjugated secondary antibodies were used in combination with Hoechst nuclear staining. Confocal imaging was performed with Zeiss LSM confocal laser scanning microscope.

Excitatory stimulation of neurons. SCA3 neurons or control neurons cultured in 3.5-cm dishes were washed three times with 2 ml BSS (balanced salt solution) containing 25 mM Tris, 120 mM NaCl, 15 mM glucose, 5.4 mM KCl, 1.8 mM CaCl2, 0.8 mM MgCl2, pH 7.4. After treatment with L-glutamate 0.1 mM (Sigma, no. G8415) in BSS for 30 min cells were washed again three times and left them to recover for 30 min in differentiation media followed by a second 30 min L-glutamate treatment in BSS, and subsequently cultured in differentiation media for 24 hr until analysed. For analysis of fragmentation and aggregation of ATXN3 by western blotting, extracts were analysed either immediately after lysis or after fractionation.

Western blotting. Neuronal cells were washed in PBS and scraped them. Cells were immediately frozen in liquid N2 followed by lysis in RIPA buffer (50 mM Tris, 150 mM NaCl, 0.2% Triton X-100) containing 25 mM EDTA. For fractionation, lysates containing 1–2 mg/ml total protein dissolved in 50 mM Tris, 150 mM NaCl, 0.2% Triton X-100, 25 mM EDTA (RIPA buffer) were centrifuged at 22,000 g for 30 min at 4 C. The pellet fractions were separated from supernatants (Triton X-100-soluble fraction) and homogenized by sonication in RIPA buffer containing 2% SDS (SDS fraction). β-mercaptoethanol (5%) was added in all the samples and subsequently incubated at 99 C for 5 min. Gels were loaded with 40 mg of the Triton X-100 fraction and 40 μl of the SDS fraction. Proteins were resolved by SDS-PAGE, transferred to
nitrocellulose membrane and then processed for western blotting. Membranes were subsequently incubated with HRP-conjugated secondary antibodies (Amersham) at 1:7000 dilution. Visualization was performed with enhanced chemiluminescence and Hyperfilm (ECL, Amersham). Mouse monoclonal anti-ATXN3 (1H9) from Acris Antibodies, while mouse monoclonal anti-HSPA1A/Hsp70, mouse monoclonal anti-HSPB1, rabbit polyclonal anti-DNAJBs/HSP40 were from Enzo life science. Mouse monoclonal anti-GAPDH was from RDI Research Diagnostics, while rabbit polyclonal anti-GFAP, mouse monoclonal anti-HSPB8 (M04) and anti-HSPCs/HSP90 were from DAKO, Abnova and Stressmarq, respectively.

**Genome-wide SNP genotyping and Genomic CAG repeat length analysis.** Genomic DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen). The Genomic CAG repeat length analysis of fibroblast, iPSC and NSC samples was performed as previously described [33]. Genome wide SNP genotyping was performed using 320k cyto Illumina arrays as per the manufacturer’s protocol (Illumina). Data were collected using the Illumina Bead Station scanner and data software. Genotypes were produced using the genotyping module of Genome Studio and copy number variation (CNV) analysis was performed. In addition, the B-allele frequencies and log R ratios were visualized using the genome viewer tool within this package.
References


