



From left to right: Guus Korver, Sanne Golembiewski, Marijn Verhaeg, Romy Kuipers, Luuk Hesselink.

Understanding the highly variable deterioration rates in adult Huntington's disease

Sanne Golembiewski, Luuk Hesselink, Guus Korver, Romy Kuipers, Marijn Verhaeg.

Supervisor: dr. H. Zhou

Abstract: Huntington's disease (HD) is a neurodegenerative disease which is caused by mutations in the Huntingtin (HTT) gene. These mutations cause abnormally long repeats of CAG codons which codes for glutamine. Although the genetic factor is known for the cause of HD, the mechanisms of the disease progression are still unknown. Depending on the disease progression, patients will suffer from choreatic movements, intellectual deterioration and a variety of physiological symptoms. The adult-onset form of Huntington's disease shows variability in disease progression causing the remaining life-time of one patient to possibly differ 10 years compared to other patients.

This research aims to give insight into the mechanism of HD disease by finding correlations between cellular phenotypes, gene expression levels and the patient deterioration on a clinical level. To identify these correlations, induced pluripotent stem cells (iPSCs) will be used as the main model. These iPSC derived cells can be differentiated into different types of neurons and data will be obtained at cellular and molecular levels. By analysing this data using machine learning algorithms, we expect to gain more knowledge about the underlying mechanism of HD, which can eventually lead to better prognosis and improved personalized treatments.

Samenvatting: De ziekte van Huntington is een neurodegeneratieve ziekte waarbij neuronen in de hersenen afsterven. Dit leidt tot verschillende mentale en lichamelijke symptomen en uiteindelijk tot het overlijden van de patiënt. De ziekte ontstaat door een mutatie in het gen dat het eiwit Huntingtin maakt. Er is bij de ziekte van Huntington namelijk sprake van een abnormaal lange CAG-herhaling in het DNA (>40) dat codeert voor het aminozuur glutamine. Echter is er een groot verschil in snelheid van de progressie van de ziekte tussen patiënten. Zo overlijdt de ene patiënt 10 jaar na diagnose en de andere patiënt na 20 jaar, ongeacht op welke leeftijd de ziekte intrad. Het mechanisme achter de progressie van de ziekte is nog niet bekend binnen het onderzoeksveld. In dit project willen we de mechanismen achter het progressie proces beter leren begrijpen met behulp van een grootschalige data-analyse. Meer kennis over deze mechanismen kan uiteindelijk leiden tot een verbetering in de behandelingen gericht op het vertragen van het ziekteproces en zal in het meest succesvolle geval kunnen leiden tot persoonsgerichte medicatie.

Introduction: Huntington's disease (HD) is an autosomal dominant neurodegenerative disease in which patients are suffering from physical symptoms such as chorea, rigidity, abnormal posturing and cognitive as well as behavioral symptoms like aggression, anxiety and depression. In contrast to most other neurodegenerative diseases, the cause of HD is largely known. Namely, the disease is caused by an expansion of the CAG repeats within exon 1 of the Huntingtin (HTT) gene that encodes for the Huntingtin protein. Because of this expansion of the CAG repeat, the protein is produced in an abnormal form and is called mutant Huntingtin (mHTT). The elongated CAG repeats form patches where aggregation of mHTT occurs, inducing apoptosis and leading to the degeneration of cells in the brain. The mutation is fully penetrant if the expansion goes beyond 40 CAG repeats, which means that these individuals will eventually develop HD at some point in their lives [1]. It is striking that there is a big variability in disease progression rate between individuals with HD [2]. Some patients die 10 years after diagnosis, while other patients die after 20 years. The precise mechanism

behind differences in disease progression between individuals is not understood so far in the research field of HD.

Research Proposal: In our research proposal, we describe a method to gain a better understanding of the underlying disease mechanisms within Huntington’s disease. This method consists of three phases (I, II and III). In Phase I we acquire data from 30 adult Huntington patients and 10 controls. From the 30 Huntington patients, we assess the deterioration rate by using retrospective data from clinical databases, using the Unified Huntington’s Disease Rating Scale (UHDRS) [3]. Furthermore, we take blood samples from all subjects and obtain iPSCs by reprogramming CD34+ enriched hematopoietic cells (HSCs) present in the blood [4]. We further differentiate the iPSCs into five different types of neurons important to our research: neural stem cells, striatal medium spiny neurons, striatal interneurons, pyramidal neurons of the cerebral cortex and cortical interneurons. To measure the cellular phenotypes, we stress the neurons by downregulating the DNAJB6 gene and exposing them to 30-minute glutamate pulses [5]. The two cellular phenotypes we measure are mHTT-aggregation and caspase activity (cell death) using fluorescence-based imaging. At the molecular level, we determine the CAG-repeat length and measure the gene expression patterns in the iPSC-derived neurons and also in the peripheral blood by RNA-sequencing. After we have gathered all this information, we will begin by analysing the data in Phase II. In Phase II, we will look for phenotype-correlations and subgroups within the patients by making use of the machine learning algorithm Support Vector Machine (SVM) [6]. Based on the subgroups we find in Phase II, we will further analyse the data in Phase III. In Phase III, we use the machine learning algorithm Random Forest (RF) for each subgroup to find the importance of each phenotype and relations between phenotypes [7-9]. This will tell us how subgroups may differ in their Huntington disease mechanisms.

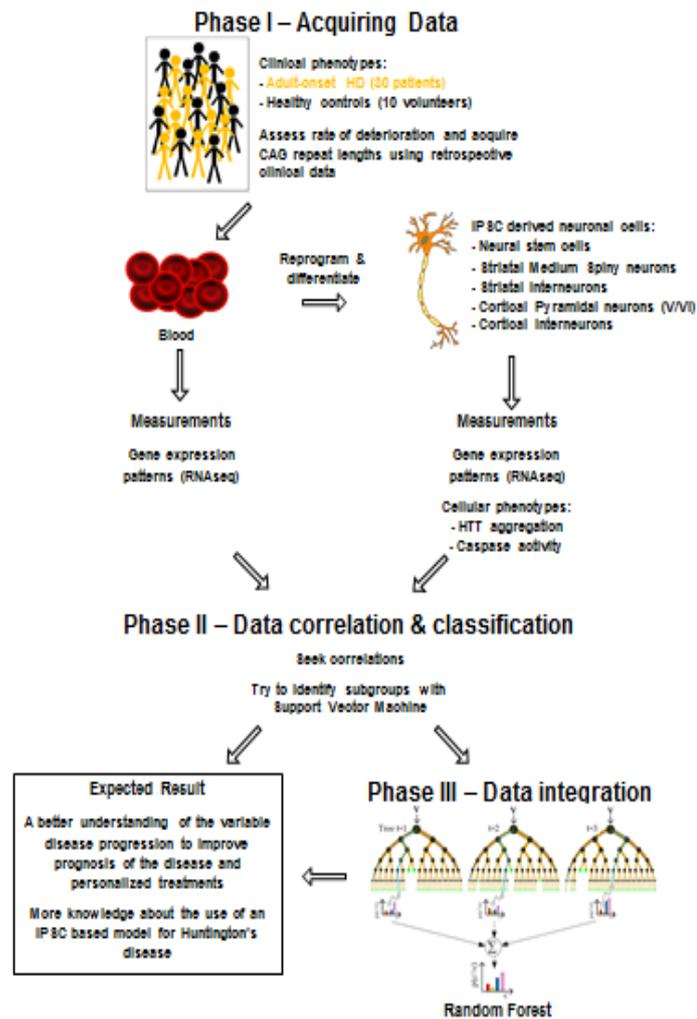


Figure 1: Schematic overview of our research methodology

Outlook: The results of our research could be very important to the field of Huntington’s disease since the iPSC-model is a novel model in the field and could be a great improvement to the rodent-model and human fibroblast-model that are currently used to model the disease. Also, if subgroups were to be found by our research, indicating that there could be different disease mechanisms, this could give rise to better prognosis and personalized treatments for patients suffering from the disease.

References

[1] Lee, J. et al. CAG repeat expansion in Huntington disease determines age at onset in a fully dominant fashion; *Neurology* 78(10).; 2012, pp 690–695.
 [2] Zhang, N. et al. iPSC-based drug screening for Huntington’s disease; *Brain research*, 1638.; 2016, pp 42–56

- [3] HSG Unified Huntington's Disease Rating Scale: Reliability and-Consistency; *Movement Disorders*, 2(2).; 1996, p 136142.
- [4] Merling, R. et al. Transgene-free iPSCs generated from small volume peripheral blood nonmobilized CD34+ cells; *Blood*, 121(14).; 2013, pp 98–107.
- [5] Thiruvalluvan, A. et al. Induced pluripotent stem cells: cell therapy and disease modeling; University of Groningen.; 2018.
- [6] Meyer, D. *Support Vector Machines*; 2017.
- [7] Donges, N. *The Random Forest Algorithm.*; 2014
- [8] Opitz, D. *Classifier Ensembles: Bagging Classifiers.*; 1999.
- [9] liaw A.; Wiener, M. Classification and regression by randomForest; *Rnews*, 2(3).; 2002, pp 18–22.