Study of SPG11's splicing, in neuronal and non-neuronal cells, by long-read sequencing: implication on phenotype

Leblanc M (1,2), Gareau T (2), Bertrand M (2), Mohammad A (2), Liu E (2), Jost Mousseau C (2), Le Goff G (2), Marie Y (2), Guégan J (2), Bohl B (2), Carlile T (3), Stevanin G (1,2,4)

(1) EPHE, PSL University, Paris, France; (2) Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, Inserm, CNRS, APHP, Hôpital de la Pitié Salpêtrière, Paris, France; (3) Biogen, Cambridge (MA), USA; (4) INCIA, University of Bordeaux, CNRS, EPHE, Bordeaux, France

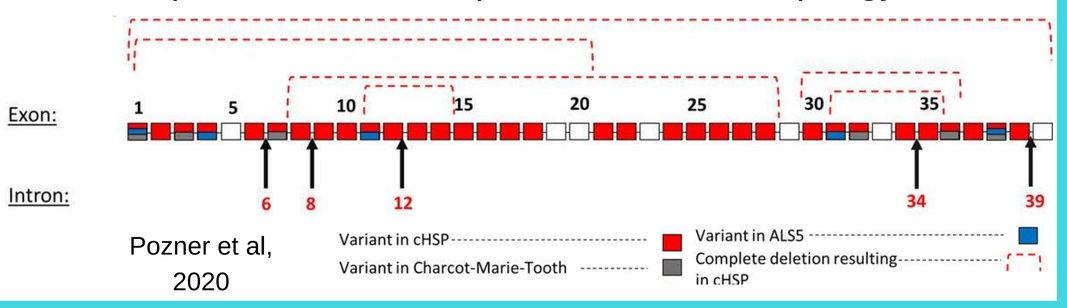
Contacts : marion.leblanc@icm-institute.org; giovanni-b.stevanin@inserm.fr

Introduction

The SPG11 gene is one of the most frequently mutated genes in the autosomal recessive forms of hereditary spastic paraplegias (HSP) and also accounts for cases with amyotrophic lateral sclerosis (ALS5) and Charcot-Marie Tooth (CMT) neuropathy. The SPG11 gene encodes a ubiquitously expressed 7.8kb full-length transcript. This transcript codes for Spatacsin, a 2 443 amino acids protein with a suspected role in autophagy. Little is known about the regulation of the expression of this gene.

Aim of the project

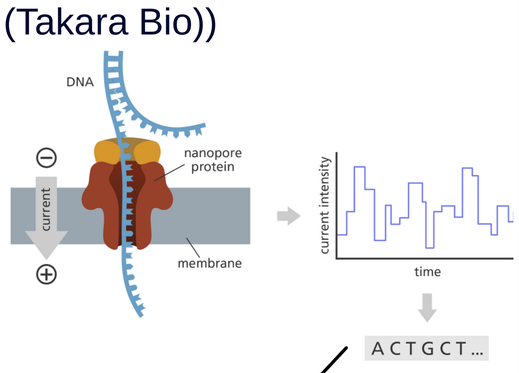
To identify all isoforms of SPG11 and establish their full expression profiles, first in non-pathological models of the disease but also in pathological models to determine if the presence of remaining isoforms could explain the variability in clinical presentation.



Pozner et al, 2020

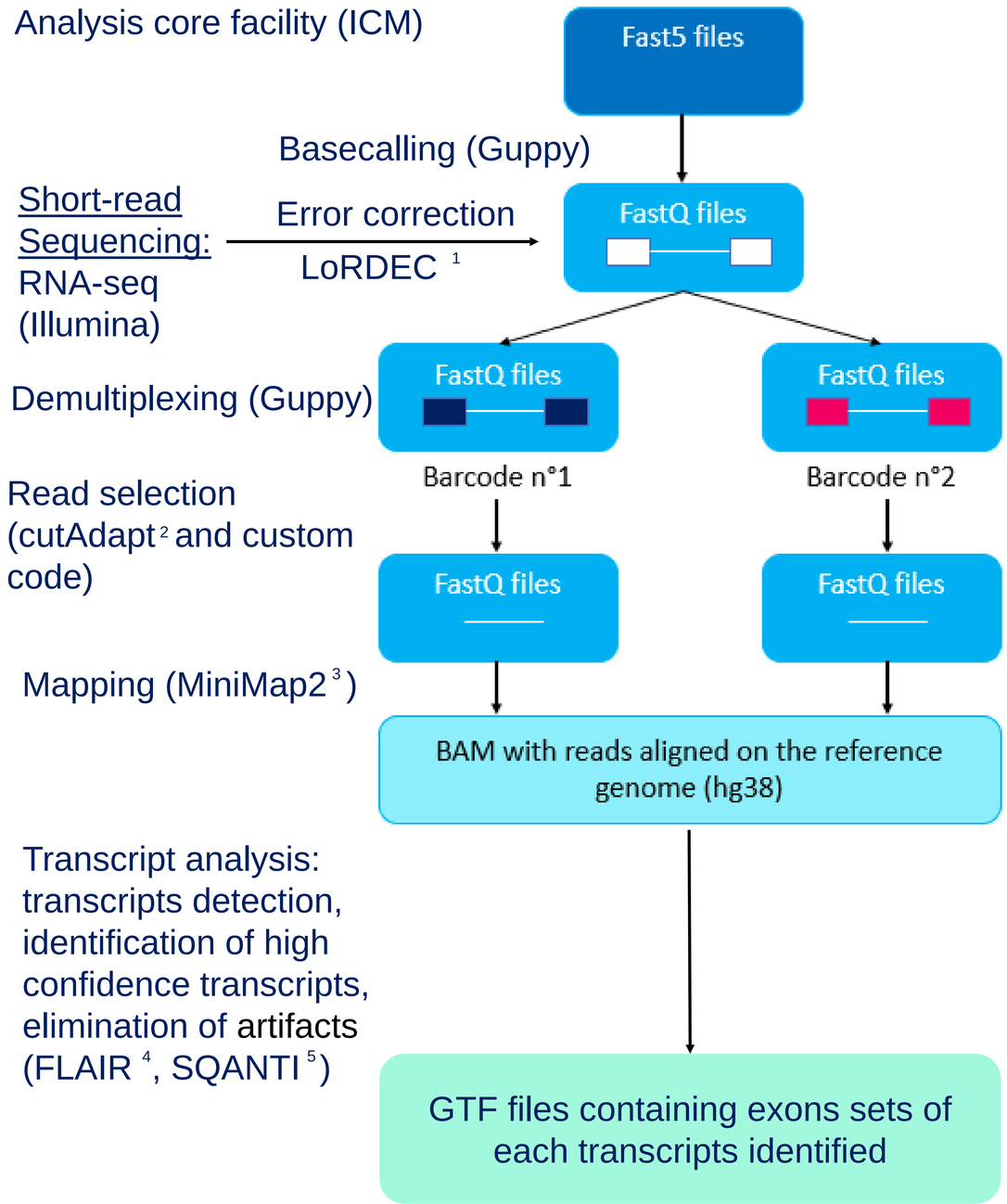
Protocol

- Retrotranscription of RNA into double-stranded cDNA (SmartSeq v4 ultra low input kit (Takara Bio))
- Library preparation (ONT PBK-004 kit, adaptation of the protocol to favor larger fragments)
- Sequencing on the MinION for 48 to 72 hours



Data Analysis

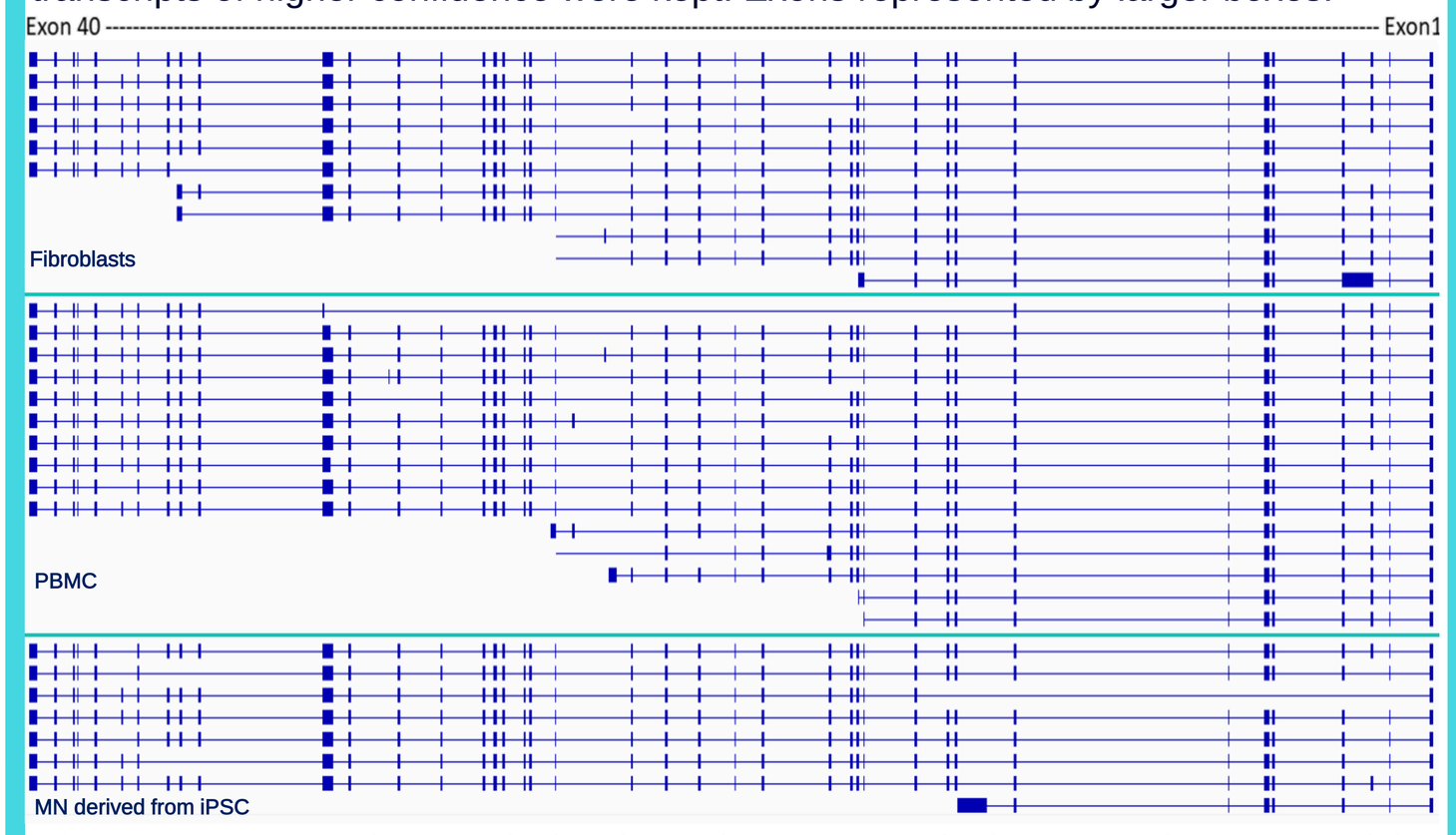
In collaboration with Data and Analysis core facility (ICM)



1 Salmela and Rivals, 2014 2 Marcel, 2011 3 Li, 2018 4 Tang et al, 2020 5 Tardaguila, 2018

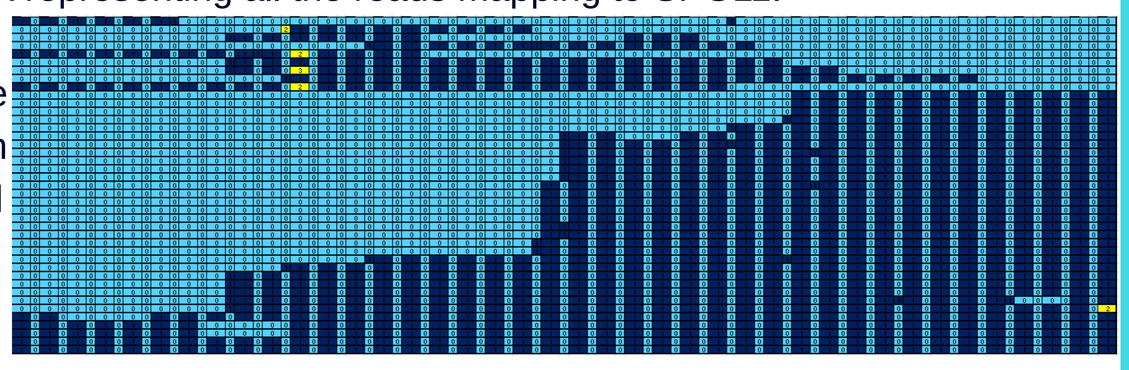
Results

1/ Controls: Transcripts in 4 human samples from 3 different control individuals and 3 different cell types (fibroblasts, peripheral blood mononuclear blood cells (PBMC) and motor neurons (MN) derived from induced pluripotent stem cells (iPSC). Only transcripts of higher confidence were kept. Exons represented by larger boxes.



2/ Patients: Transcripts analysis of one human sample from a patient mutated for SPG11 (homozygous stop codon in exon 32) and presenting a complex form of HSP. Splice matrix representing all the reads mapping to SPG11.

Presence in dark blue and absence in light blue. Exon 1 on the right and exon 40 on the left.



Conclusions

- >15 different transcripts detected, the full length one is the most expressed
- New SPG11 transcripts identified in addition to the full-length transcript
- Cell-type specificities -> transcripts but also exon usage
- Events observed: exon skipping, alternative polyadenylation, intron retention/ new exons
- SPG11 mutated patients: high decrease of the major canonical transcript but other isoforms remaining that may retain partial function

Perspectives

- Sequencing of RNA extracts from fibroblasts, iPSC-derived MN and PBMC of other control individuals, and from SPG11 and SPG11-mutated ALS5 patients are under processing to increase our dataset
- Study of transcripts stability in cellular context
- Addition of a protein prediction step in the analysis pipeline
- Study of proteins expression in vitro -> localization and function
- Phenotype-genotype correlation (ALS5 versus SPG11 patients)

Acknowledgements

The research leading to these results has received funding in the context of the PSL-Biogen doctoral partnership (CIFRE convention n°2019/1067).
We gratefully acknowledge the iGenSeq core and Data and Analysis core facilities at ICM.