Progress report

Project title: Identification of the mutated genes in two novel forms of hereditary spastic paraplegia.

Applicant: Dr. med. Alfredo Ramirez

1. Linkage analysis and next generation sequencing in the Chilean family with HSP

In the Chilean family with HSP, we defined first the genomic region containing the putative pathogenic gene. For this, we analyzed in detail the linkage data searching for critical recombinations within the family. Thereby, we established an interval on the short arm of chromosome 4 within the first 13 million base pairs. In collaboration with Professor Bauer from the Institute of Human Genetics at the University of Tübingen, we searched in the Consensus Coding Sequences (CCDS) collection database to define all candidate exons within this region. To specifically sequence these segments, we sent these data to NimbleGen to design an array selectively capturing these genomic intervals. This approach identified DNA fragments with an average size of 500 base pairs allowing for an analysis of regions beyond the intron/exon boundaries. To reduce the number of potentially interesting variants upon completion of our analysis, we enriched and sequenced two affected individuals from different generations, i.e. individual 8 and 14 from the pedigree. After the enrichment, we sequenced the enriched DNA using a Roche Genome Sequencer from 454 Life Science. The first analyses of the sequencing results revealed a total of 941 variants in individual 14 and 911 in individual 8 (table 1). To search for the causative mutation among all these variants, we first excluded all known variants (single nucleotide polymorphisms, SNPs), i.e. 410 variants in individual 14 and 449 in individual 8.

Table 1: Summary of variants detected

Variants	Results Roche Software	
	Individual-14	Individual-8
Total variants	941	911
Known variants	410	449
Novel variants Non-synonym	283	213
Other novel variants	248	249

Next, we aimed to further investigate novel variants present in both individuals producing an amino acid change. Through this, we identified 6 novel variants in both persons (table 2).

Table 2. Novel variants causing aminoacid exchange found in both individuals

Gene	Exon	Variation Bases	Results Roche Software	
			Individual-14	Indiviual-8
Gene 1	Exon 7	-/T	Present	Present (doubtful)
Gene 2	Exon 1	G/A (p.A-->T)	Present (doubtful)	Present (heterozygous)
Gene 3	Exon 7	AC/-	Present (doubtful)	Present
Gene 4	Exon 1	C/A (p.P-->T)	Present	Present
Gene 5	Exon 1	T/C (p.S-->G)	Present (heterozygous)	Present (heterozygous)
Gene 6	Exon 06-07	A/C (p.S-->R)	Present (heterozygous)	Present (heterozygous)

Since next generation sequencing (NGS) is a high-throughput technique generating a large amount of data, we expected that some of these novel variants are false-positive results. Therefore, validation of NGS data is mandatory using conventional sequencing techniques (Sanger sequencing).

2. Validation of variants found with NGS

To further confirm the potentially pathogenic role of the validated variants, segregation analysis of these variants with the phenotype was performed in the family. In addition, screening of German and Chilean healthy controls was done for those variants segregating with the phenotype in the family. In this regard, a variant is classified as highly probably pathogenic, if the variant: a) can be validated with Sanger sequencing; b) segregates with the phenotype; c) is not present in healthy controls. Consequently, we initiated the validation of the variants found with NGS by designing PCR primers flanking each of these novel polymorphisms. The DNA of both Chilean HSP patients, originally sequenced with NGS, served as template to amplify the fragments containing each novel polymorphism. These PCR-fragments were directly sequenced using Sanger sequencing chemistry. This approach confirmed 4 out of the 6 variants originally identified with NGS (table 3). Interestingly, one of these confirmed polymorphisms was detected only in one of both Chilean patients excluding, thereby, this variant as potentially pathogenic.

Table 3. Novel variants confirmed by Sanger sequencing

Gene	Exon	Sanger Sequencing	
		Indiviual-14	Individual-8
Gene 1	Exon 7	Not found	Not found
Gene 2	Exon 1	Not found	Present (heterozygous)
Gene 3	Exon 7	Not found	Not found
Gene 4	Exon 1	Present (heterozygous)	Present (heterozygous)
Gene 5	Exon 1	Present (heterozygous)	Present (heterozygous)
Gene 6	Exon 06-07	Present (heterozygous)	Present (heterozygous)

In a second step, we analyzed the segregation of the three remaining variants with the phenotype within the Chilean family (table 4). This analysis showed that none of the three variants segregated perfectly in the family (i.e. at least one healthy family member carries the variant). Nevertheless, we proceed with the screening of healthy controls for these three polymorphisms because incomplete penetrance has been already described in HSP. We screened first healthy controls of German origin using the High Resolution Melting (HRM) Analysis on the LightCycler 480 (Roche Diagnostics, figure 1). Two variants were present in the German controls. Together with the segregation analysis, these findings excluded these variants as a potentially pathogenic leaving only one variant.

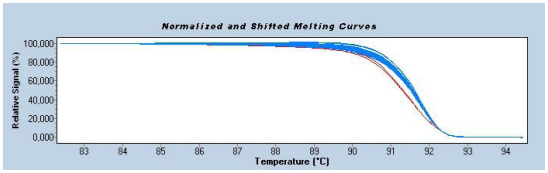
Table 4. Segregation analysis of validated variants in the Chilean family

	Gene 4	Gene 5	Gene 6
Individual_ID	Exon01 (C>A; p.P>T)	Exon01 (T>C; p.S>G)	Exon06-07 (A>C; p.S>R)
1			
2			
3			
5			
6			
7			
8			
10			
11			
12			
13			
14			
15			

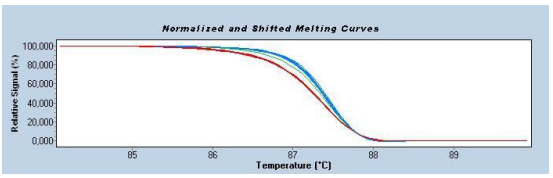
In red are shown affected individuals and also individuals heterozygous for each variant

The remaining polymorphism is a heterozygous cytosine to adenine transversion which is located in the first coding exon of a novel gene (gene 4). At the protein level, this variant predicted a non-conserved amino acid exchange of a proline for a threonine. The affected gene codes for a protein of unknown function with no homology to any known gene family. For this reason, the putative functional consequences of this polymorphism cannot be readily predicted. However, its absence in 167 German controls prompted us to screen Chilean healthy controls. This analysis detected the polymorphism only in one Chilean control out of 285 Chilean controls.

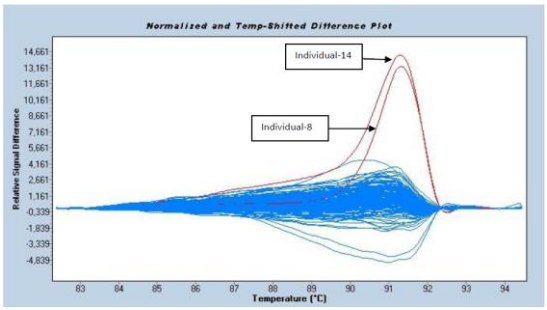
(A) **Gene 4**: Exon 1; ; C>A; p.P>T



(C) **Gene 6**: Exon 6-7; A>C; p.S>R



(B)



(D)

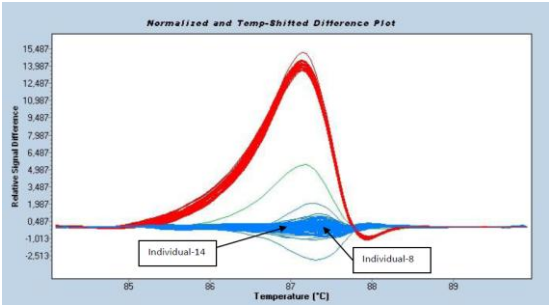


Figure 1. Screening of German controls using HRM analysis. (A) and (C) show the temperature-shifted melting curves for two different variants analyzed in controls. Each melting pattern (i.e. genotype) detected in each experiment is shown with a different color (red and blue). In (B) and (D) are depicted the normalized melting curves for (A) and (C), respectively. This approach improves the ability to recognized different melting patterns, each representing a particular genotype. In each HRM experiment, both Chilean patients were included as positive control. Their melting curves are indicated with arrows in (B) and (D). Thus, in German controls, the variant analyzed on the left panels is only present in both patients. Conversely, the variant on the right panels is present also in healthy controls. In conclusion, the first variants may represent a pathogenic mutation in the Chilean family and the second variant only a common polymorphism.

At present, our results might be in keeping with a possible pathogenic role of the variant located on gene 4. On the other hand, some findings argue against this variant being causative for the disease. First, we did not find a perfect segregation of the variant with the phenotype in our Chilean family. Second, the variant is also present in one healthy Chilean control. Furthermore, the polymorphism is located in an orphan gene with unknown function, hampering, thereby, the prediction of functional consequences for this amino acid exchange. For this reason, we revisited the NGS data initiating the analysis of variants with a putative effect on splicing. Additionally, we aim to screen a sample of sporadic HSP patients for mutations in this gene.

Thank you very much for your support of our project.

With kind regards

Dr. med. Alfredo Ramirez