The p.Ala510Val mutation in the SPG7 (paraplegin) gene is the most common mutation causing adult onset neurogenetic disease in patients of British ancestry, *Journal of Neurology*, Richard Roxburgh, Renate Marquis-Nicholson, Fern Ashton, Alice M. George, Rod A. Lea, David Eccles, Stuart Mossman, Thomas Bird, Koen L. van Gassen and Donald R. Love

Corresponding author:
  Dr Richard Roxburgh
  Department of Neurology
  Auckland City Hospital
  Private Bag 92024
  Auckland 1142, New Zealand
  Email: richardr@adhb.govt.nz
  Phone: +64 9 3074949 extension 25810

**Online Resource 3**

**Population Genomic Analysis of Variants Spanning the SPG7 Gene in Maori and European Cohorts**

**OBJECTIVE**
To estimate *in silico* whether a mutation in HSP patients from New Zealand is likely to reside on a haplotype of Maori origin given the Maori ancestry of the affected pedigree and to use this and other population genetic markers to thereby estimate whether p.Ala510Val SPG7 disease is likely to be common in the Maori population of New Zealand as a whole.

**BACKGROUND**
The purpose is to find whether the mutation (c.1529C>T; p.Ala510Val; NM_003119.2; SPG7 gene; paraplegin; chr16:88102306-88151675 hg18 build) resides on a haplotype which is specifically of Maori ancestry, rather than European. Genotype data for SNPs spanning the SPG7 locus in four affected pedigree members from an Affymetrix chip has been collected. All patients are homozygous for the p.Ala510Val mutation as well as all SNPs spanning the SPG7 gene.

**METHODS**
This was a population genomic analysis of single nucleotide polymorphisms (SNPs) spanning the SPG7 gene in a cohort of individuals of Maori ancestry (Rakaipaaka Health and Ancestry Study (RHAS), n=30) and a cohort of individuals with Caucasian/European (CEU) ancestry (CEU-HapMap, n=90) [1-2].

A) Identified all typed SNPs from a Maori (RHAS) and European (CEU) cohort that span the SPG7 gene. Assessed the relative SNP positions compared to the SPG7 gene mutation (c.1529C>T; p.Ala510Val).

B) Calculated allele frequencies and identified all ancestry informative SNPs (AIS) for the Maori cohort based comparisons with the CEU cohort.

C) Estimated haplotype frequencies and LD patterns at the SPG7 gene region in the Maori (RHAS) and European (CEU) cohorts.

D) Estimate the frequency of the (p.Ala510Val) mutation in the Maori population.
RESULTS

A. We identified 9 SNPs typed in both the Maori (RHAS) and European (CEU) groups that were also typed in the HSP patients. These SNPs span 1 MB across the SPG7 gene. The relative positions of the SNPs and the SPG7 mutation (rs61755320) is shown in Figure 1. One SNP (rs382745) resides within the SPG7 gene and is approx. 10 kb away from the mutation (rs61755320).

![Figure 1. Map showing relative positions of SNPs at the SPG7 gene region](image)

B. Table 1 shows the SNP IDs and allele frequency difference (or delta) statistics for these SNPs. A delta value of 0 means allele frequencies are equally frequent whilst a delta value of 1 means alleles are private to one ancestral group and thus completely informative of ancestry. Values >0.25 indicate moderate informativeness. There was one SNP, rs258322, that had quite different frequencies between RHAS and CEU, and three others (rs2353033, rs7188458, rs11861084) that could be considered to be ancestry-informative. None of these ancestry-informative SNPs reside within the SPG7 gene, although both rs4347628 and rs382745 (not AISs) are within the gene.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Delta</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs258322</td>
<td>0.57</td>
</tr>
<tr>
<td>rs2353033</td>
<td>0.29</td>
</tr>
<tr>
<td>rs7188458</td>
<td>0.27</td>
</tr>
<tr>
<td>rs11861084</td>
<td>0.25</td>
</tr>
<tr>
<td>rs4782449</td>
<td>0.18</td>
</tr>
<tr>
<td>rs4347628</td>
<td>0.09</td>
</tr>
<tr>
<td>rs3803681</td>
<td>0.03</td>
</tr>
<tr>
<td>rs8058269</td>
<td>0.03</td>
</tr>
<tr>
<td>rs382745</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Table 1. Allele Frequency Differences in Ancestral Cohorts
C. The haplotype block patterns across the \textit{SPG7} region are shown in Figure 2. There appears to be less recombination occurring in the European (CEU) group then in the Maori (RHAS) group. However, there is a single haplotype block including 3 SNPs and including the \textit{SPG7} gene in both cohorts: rs3803681 at pos 89544756, rs4347628 at pos 89570635, and rs382745 at pos 89603586, - a total span of about 59kb.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{haplotype.png}
\caption{The haplotype block on the left represents the individuals of European ancestry (CEU); that on the right those of Maori Ancestry (RHAS). The two populations share the central haplotype block which contains the \textit{SPG7} gene.}
\end{figure}

Table 2. shows that for both cohorts there were 4 identified 3-SNP \textit{<rs3803681, rs4347628, rs382745>} haplotypes of frequency > 1%. The family was homozygous for the commonest GAA haplotype. There is no substantial difference in haplotype frequencies between the two groups.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
3 SNP Haplotype & Maori (RHAS) & European (CEU) \\
\hline
GAA & 0.483 & 0.394 \\
AGG & 0.467 & 0.439 \\
AGA & 0.033 & 0.094 \\
GGA & 0.017 & 0.072 \\
\hline
\end{tabular}
\caption{Frequencies of Haplotype containing the \textit{SPG7} gene in Ancestral Cohorts}
\end{table}

Because HSP caused by \textit{SPG7} mutations, is a recessive disorder it is possible that genotype frequencies may differ between ancestral cohorts which may, therefore, provide clues as to the ancestral origin of the disease in this pedigree. To check this, genotypes of the tagging SNP (rs382742) [2] were compared for RHAS and CEU cohorts. The frequency of the homozygote (AA) genotype was 0.36 in CEU and 0.2 in the RHAS. Thus, these differences also do not indicate over representation in the Maori cohort either.
D. We identified another SNP rs2292954 (not ancestry informative so not included in the first analysis) that was only 23bp away from the mutation (which is rs61755320 [3]) This SNP is also within the single haplotype block and, given its very close proximity to the mutation, it is reasonable to assume that the frequencies of rs2292954 reflect the mutation accurately, although rs2292954 is likely to be slightly more ancient than the mutation so frequencies are a little higher.

In terms of rs2292954, frequencies of the rare allele (C) and genotype (CC) are 0.14 (CEU) and 0.05 (RHAS), and 0.06 (CEU) and 0 (RHAS), respectively. This suggests the C allele is more common in Caucasians (14%) compared to Maori (5%). In other words, the allele probably originated in Europe and decreased in Maori due to migratory effects (eg. bottlenecks). This also suggests that HSP due to homozygous mutations would be much less frequent in Maori compared to European.

In support of this, the presence of the actual mutation (rs61755320) at a heterozygous rate of 3/1000 – no homozygotes were observed - in Caucasians [4] confirms that the mutation originated in Europe. The mutation has not been typed in Maori but doesn't need to be given the proxy findings. The frequencies of the proxy SNP rs2292954 (shown above) indicate that the mutation became less frequent in Maori during their migrations. Given the tight correlation between rs2292954 and the mutation (23bp apart) the frequency of the mutation in Maori is likely to be approximately 1/3 of that seen in Caucasians.

CONCLUSIONS
The prevalence of the background haplotype (GAA) in Maori and European is high (>30%) and the frequency estimates are not sufficiently different to be ancestry-informative. Thus, there is no evidence that this inherited haplotype, which contains the SPG7 mutation (p.Ala510Val) is of Maori origin. The SPG7 mutation – also called rs61755320 present at a frequency of 0.003 in a worldwide cohort of individuals which includes mostly Europeans [4] coupled with the findings of this study suggest that the p.Ala510Val variant in the SPG7 gene in this family is an ancient mutation that occurred many generations before the Polynesian migrations. Population frequencies of an SNP in close in close linkage disequilibrium with the SPG7 p.Ala510Val mutation in a cohort of patients of Maori ancestry indicate that the frequency of this mutation is likely to be very low amongst Maori. While it is possible that the mutation travelled over time from the archaic European ancestor through this family’s Maori lineage, the results are more consistent with a recent European ancestor bringing the mutation into the family.

REFERENCES
4. The 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing Nature 2010;467:1061-73