Whole genome sequencing a tool to unravel rare variants associated with ALS survival

Philip Van Damme & Matthieu Moisse
Clinical presentation of ALS

- Third most frequent neurodegenerative syndrome
- Epidemiology
  - Incidence 2/100,000 a year
  - Prevalence 4-8/100,000
  - Life time risk 1/400
  - Variability in clinical presentation
Worldwide Project MinE
More and more countries join this groundbreaking genetic ALS research!

38%
8,335.98 / 22,500.00 DNA profiles collected
Learn more

https://www.projectmine.com/
Survival analysis
ALS and survival

Single variants associated with survival

- 1552 cases from BE and NL origin had the required clinical info and used in the analysis
- Cox-regression with the following cofactors:
  - Age
  - Gender
  - Sequencing technology
  - Population through pca1-10
  - C9orf72
  - Site of onset
Findings

- 138 loci have a p-value < 5e-8
- Smallest p-value 3.7e-13
- HR between 1.7 and 44
- MAF from 0.07% up to 4.5%
Findings

• 138 loci have a p-value < 5e-8
• Smallest p-value 3.7e-13
• HR between 1.7 and 44
• MAF from 0.07% up to 4.5%
Additive effect

Nbr loci per patient

Survival probability

Survival (Years)

Survival probability

Survival (years)

\( p < 0.0001 \)
QQ plots reveal deviation from expected
QQ deviation due to rare variants
Issues

- **Issue:** Hard from distinguish real an association from a random
- **Solution:** permutations
- **Currently at 1,789,090**

Table 1 Adaptive permutation recommendations

<table>
<thead>
<tr>
<th>m</th>
<th>a_p</th>
<th>b</th>
<th>r</th>
<th>b</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.05</td>
<td>1,900</td>
<td>115</td>
<td>475</td>
<td>34</td>
</tr>
<tr>
<td>5</td>
<td>0.01</td>
<td>9,900</td>
<td>120</td>
<td>2,475</td>
<td>36</td>
</tr>
<tr>
<td>50</td>
<td>1e-3</td>
<td>99,900</td>
<td>121</td>
<td>24,975</td>
<td>36</td>
</tr>
<tr>
<td>500</td>
<td>1e-4</td>
<td>999,900</td>
<td>121</td>
<td>249,975</td>
<td>36</td>
</tr>
<tr>
<td>1,000</td>
<td>5e-5</td>
<td>1,999,900</td>
<td>121</td>
<td>499,975</td>
<td>36</td>
</tr>
<tr>
<td>10,000</td>
<td>5e-6</td>
<td>19,999,900</td>
<td>121</td>
<td>4,999,975</td>
<td>36</td>
</tr>
<tr>
<td>100,000</td>
<td>5e-7</td>
<td>199,999,900</td>
<td>121</td>
<td>49,999,975</td>
<td>36</td>
</tr>
<tr>
<td>1,000,000</td>
<td>5e-8</td>
<td>1,999,999,900</td>
<td>121</td>
<td>499,999,975</td>
<td>36</td>
</tr>
</tbody>
</table>

Recommendation of the number of permutations (b) and cut-off value (r) varying the number of SNPs (m), PWER (a_p) and precision level (c), with a fixed EWER (a_c = 0.05).

Che et al., BioData mining 2014
Permutations
Validation cohort

- Cohort of 616 ALS patients
- Could decrease the p-value of 6 with a p-value < 5e-8
ALS and survival extremes

Single variants associated with extreme survival

- 613 cases from BE and NL origin with survival > 4 or < 1.5 years
- Linear-regression with the following cofactors:
  - Age
  - Gender
  - Sequencing technology
  - Population through pca1-10
  - c9orf72
  - Site of onset
Extremes analysis
• 0 loci have a p-value < 5e-8
• 59 variants have a p-value < 5e-6
• HR between 0.34 and 2.6
• MAF from 14% up to 50%
Onset analysis
ALS and onset

Single variants associated with onset

• 1552 cases and 881 controls from BE and NL origin had the required clinical info and used in the analysis
• Cox-regression with the following cofactors:
  • Gender
  • Seq technology
  • Population through pca1-10
Findings

- 337 loci have a p-value < 5e-8
- Smallest p-value 4.6e-18
- HR between 2.3 to 93
- MAF between 0.05% and 1.0%
Findings

- 337 loci have a p-value < 5e-8
- Smallest p-value 4.6e-18
- HR between 2.3 to 93
- MAF between 0.05% and 1.0%
- 51 loci validated
Pathway analysis

-log(BH p-value)

0.00 0.25 0.50 0.75 1.00 1.25 1.50 1.75

Synaptic Long Term Depression eNOS Signaling Nitric Oxide Signaling in the Cardiovascular System Neuropathic Pain Signaling in Dorsal Horn Neurons Synaptic Long Term Potentiation Sperm Motility Calcium-induced T Lymphocyte Apoptosis ErbB4 Signaling Dopamine-DARPP32 Feedback in cAMP Signaling Gap Junction Signaling GPCR-Mediated Nutrient Sensing in Enteroendocrine Cells α-Adrenergic Signaling Neuregulin Signaling CREB Signaling in Neurons ErbB Signaling

Threshold

Ratio

0.00 0.025 0.050 0.075 0.100 0.125 0.150
Path Designer Synaptic Long Term Potentiation
QQ plots reveal deviation from expected
QQ deviation due to rare variants
Additive effect

(p < 0.0001)
Previously reported associated loci

UNC13A
Exonic loci in HENMT1
5’UTR IL5RA
CDK14 - gene with the most hits.
... followed by RNU6-830P
Previously reported associated loci

CAMTA1

IDE
Previously reported associated loci

KIFAP3

UNC13A

chr1:169873354

chr19:17644214
Previously reported associated loci

DOA

EPHA4