|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **No** | **Study** | **Country** | **Ethnicity** | **Disease group** | **No of cases** | **Repeat size in cases** | **No of controls** | **Repeat size in controls** | **Predominant haplotype** | **Repeat sizing method** | **Cut-off for expansion** | **Remarks** |
| 1 | Rollinson et al, 2012 | UK | British | AD | 568  | ≤21 | 314 | 1-23 | Not studied | rp-PCR | Not mentioned | 80% of AD patients from a specialist young-onset dementia clinic (mean age 62.9 years, range 37-90).  |
| 2 | Cacace et al, 2013 | Northern European | Caucasian |  AD | 1217 | 2-25 | 1119 | 2-24 | Intermediate repeats associated with rs2814707 | rp-PCR, and STR genotyping assay (STR-PCR) for samples with <30 repeats | Short 2-6 repeats; intermediate 7-25 repeats | No difference in repeat length distribution between MCI and AD or MCI and controls. MCI with intermediate repeats not at greater risk for AD. No association of intermediate alleles with CSF t-tau  |
|  MCI | 200 | 2-24 |
| 3 | Harms et al, 2013 | USA | Not specified | AD | 872 | Mean (sd) 6.5 (4.1) | 888 | Mean (sd) 4.48 (3.7) | Founder haplotype in expansion carriers | rp-PCR; Fluorescence-based fragment size analysis. Southern blot in carriers. | >30 | No difference in mean length of longest allele between AD & controls; p = .10). Distribution of longest allele lengths was similar. Within normal ranges, higher repeat length not a risk factor for AD. No association between longest non-expanded allele length and age at onset (p = .52) or interaction with APOE genotype. |
| 4 | Kohli et al, 2013 | USA | Caucasian (European)  | AD | 1184 | ≤4-23 | 1039 | ≤4-23 | Risk-haplotype A-allele in allexpanded samples | rp-PCR and fragment length analysis | >30 | Mean repeat length slightly greater in European versus African American controls and in the combined case and control samples (6.97 vs.6.72 repeat copies, t test p=0.023). |
| African-American | AD | 291 | 620 |
| 5 | Xi et al, 2012 | UK, Italy, Spain, North America | European and North American | AD | 424 | 2-30  | 602 | 2-30 | Not mentioned | 2-step PCR – fragment length analysis followed by rp-PCR | >30 | Intermediate allele frequency similar between AD and controls (1%). Alleles with 21 and 23 repeats did not segregate with AD in 2 families. 15 autopsy AD cases with TDP-43 pathology had 2-12 repeats. Trend toward association between 10-repeat allele and risk for all 4 disorders (OR 1.72-2.14) seen. |
| 6 | Luo et al, 2014 | China | Han Chinese | Genetically unidentified (HSP) | 8329 | Mean (sd) 5.44 ± 2.87(2-14) | 308 | 2-14 | Not mentioned | 2-step PCR with rp-PCR then FAM-fluorescent labelled PCR assay  | >30 | No case with pathogenic expansion.Longest repeat allele per sample converted into a binary categorical variable: short alleles (<7 repeats); intermediate alleles (7–27 repeats). |
| Spastic paraplegia 4 (SPG4) | 68 | Mean (sd) 6.67 ± 2.55(2-11) |
| 7 | Fahey et al, 2014 | Ireland | Irish | All psychosis  | 1243 | <30; Only 2 cases had >20 repeats (27 & 28) both with schizophrenia | 1234 | <30; Only 5 controls had >20 repeats (23-26)  | Founder haplotype and anear-identical haplotype rs10511816  | rp-PCR | >30 | All samples with ≥17 repeats carried a version of founder haplotype with very strong correlation with repeat length (Spearman’s Rho =0.714, p < 0.001). 78% of samples with >6 repeats carried the founder haplotype vs only in 3% of samples carrying 2-6 repeats. |
| 8 | Galimberti et al, 2014 | Italy and Germany | Italian and German | Schizophrenia or schizoaffective disorder | 298 | Estimated 2-30 | Number not specified | Estimated 2-15 | Not studied | rp-PCR | >40 | Pathogenic expansion detected in 2 cases, both with acoustic hallucinations and alcohol abuse with positive family history for dementia and psychiatric conditions. |
| 9 | Solje et al, 2016 | Finland | Finnish | Psychosis, schizophrenia | 130 | Mean 5 (1-26) | None | - | Not mentioned | rp-PCR and fragment length analysis | >30-40 | 4 samples with intermediate repeats 17-26, no detailed clinical data reported. |
| 10 | Gami et al, 2015 | UK Queen Square Brain Bank | British | Cognitively normal aged controls  | - | - | 86 | Not mentioned |  | PCR  | - | 1 control with ~30 repeats had no TDP43-positive lesions and sparse p62-positive inclusions containing all five DPRs. DPR inclusions and RNA foci absent in 2 controls with 20 repeats. |
| 11 | Alias et al, 2014 | Spain | Spanish | Gene-confirmed SMA | 149 | ≤17 | None | - |  | rp-PCR and Southern blot |  | Longest allele in unrelated series was 17. Larger allele detected contained 10 repeats; 80% of patients carried <5 repeats. |

Supplementary Table 5. AD = Alzheimer’s disease; rp-PCR = repeat-primed polymerase chain reaction; MCI = mild cognitive impairment; STR = short tandem repeat; HSP = hereditary spastic paraplegia; TDP43 = TAR DNA-binding protein 43; DPR = dipeptide repeat proteins; RNA = ribonucleic acid; SMA = spinal muscular atrophy