(Submitted for editorial re-review to *Mitochondrion*)

Mildly deleterious mitochondrial DNA variants are less frequent in healthy older individuals, but not more prevalent in Alzheimer's patients

Ilse S. Pienaar[1,2], Neil Howell[3], Joanna L. Elson[4,5]

1. Centre for Neuroinflammation & Neurodegeneration, Division of Brain Sciences, Faculty of Medicine, Imperial College London, Hammersmith Hospital Campus, Du Cane Road, London, W12 0NN, United Kingdom
2. Department of Applied Sciences, Faculty of Health and Life Sciences, Northumbria University, Newcastle upon Tyne, NE1 8ST, United Kingdom
3. Matrilinex LLC, San Diego, CA 92122, USA
4. Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, NE1 3BZ, United Kingdom
5. Centre for Human Metabonomics, North-West University, Potchefstroom, South Africa

Running title: MutPred score variation in AD and ageing

Key words: Alzheimer's disease; Mitochondrial DNA (mtDNA); Mutational load; MutPred; mtDNA population variation

*Correspondence should be sent to: Dr. Joanna Elson, Lecturer in Genetics, Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, NE1 3BZ, United Kingdom; Email: j.l.elson@ncl.ac.uk; Tel: +44 191 241 8630

Abstract

Mitochondrial DNA (mtDNA) association studies have been conducted for over a decade using the haplogroup (lineage) association method, but this frequently produces conflicting results. Here we analyzed complete mtDNA sequence data of Alzheimer's disease (AD) patients and aged controls, from the United Kingdom (UK) and the United States (US), using a new "mutational load" method. We calculated a pathogenicity score for each of the non-synonymous substitutions of the mtDNA sequences to produce a "total mutational load" for each sequence, and compared the mutational loads of cases and controls. Using these mutational load measures, we found no evidence to support the cumulative role of mtDNA variants as a susceptibility factor in AD; that is, AD patients (UK and US cohorts) did not have higher "mutational loads" than controls. However, the US aged controls, who are significantly older than the UK ones, with many showing evidence of being healthy and having good cognition in old age, had significantly lower "mutational loads". This finding suggests that low mtDNA mutational load is more prevalent in healthy older people.

1. Introduction

The prevalence of Alzheimer's disease (AD), a neurodegenerative disorder characterized symptomatically by impaired memory, alterations to personality and decreased visual-spatial skills, is increasing as the population ages; this increase is especially pronounced in the western world, but will almost certainly become an increasingly global issue. Although there are a small number of families presenting with autosomal-dominant AD, more than 90% of AD cases are classified as sporadic-onset. Nevertheless, results from genome-wide association studies (GWAS) estimate that the heritability of sporadic AD may be as high as 80% (Harold et al. 2009; Wang et al. 2014), so this important issue is far from settled.

The neuropathology of AD is complex, and includes loss of neurons, central inflammation, as well as the formation of amyloid plaques and neurofibrillary tangles that impact on neuronal signaling (Blennow et al. 2006). Although the triggers that lead to AD remain largely unknown, there is considerable evidence for a pathological role played by abnormal mitochondrial function and oxidative stress in the etiopathogenesis of AD (Valla et al. 2006). In this regard, in the post-mortem brains of AD patients, progressive mitochondrial dysfunction in neurons was described, in addition to the more frequent occurrence of cytochrome oxidase (COX)-negative neurons, an indication of increased oxidative stress and impairment in the mitochondrial respiratory chain, compared to neurological controls (Cottrell et al. 2002). Despite such studies, whether mitochondrial dysfunction is causal in AD remains a matter of considerable debate (Howell et al. 2005), as is the question of whether inherited variants located on the mitochondrial chromosome play a role in the onset and progression of AD (Bi et al. 2015; Elson et al. 2006; Maruszak and Zekanowski 2011). "The current study investigated the role of inherited mtDNA population variants, to examine for a possible role for these variants in susceptibility to AD, by using recent computational methods, and applying a new model in an attempt to establish whether such an association exists. The work presented here does not look at somatic mtDNA mutations (Reeve et al. 2008; Reeve et al. 2009); that is, it does not consider mitochondrial dysfunction acquired on a cellular basis during an individual neurons during an

individual's lifetime (Greaves et al. 2014)."

Human mitochondrial DNA (mtDNA) encodes for 13 essential polypeptide components of the mitochondrial respiratory chain that generates adenosine triphosphate (ATP). Since mtDNA adheres to a strict maternal inheritance pattern, the evolution of mtDNA is characterised by the emergence of distinct lineages, called haplogroups (Herrnstadt et al. 2002). In light of this unique inheritance pattern, it is worth considering hypotheses by which inherited mtDNA variants might contribute to common/complex diseases such as AD. Firstly, it has been claimed that a number of complex diseases associate significantly with one mtDNA haplogroup, suggesting that one or more haplogroup-associated polymorphisms modify the risk (or outcome) of disease (Figure 1 (red star)). Whilst there are recent studies that support a role for commonly-seen inherited mtDNA variants in AD (Maruszak et al. 2009), as well as in other age-related neurodegenerative conditions, such as Parkinson's disease (PD) (Hudson et al. 2014), these associations remain to be confirmed. In this regard, a recent meta-analysis failed to find a consistent association between common mtDNA variants and AD (Hudson et al. 2012); however, another recent study did find such an association, this time in a Chinese population (Bi et al. 2015). This discrepancy highlights the issue of haplogroup studies being complicated by the occurrence of mtDNA variants that occur multiple times during human evolution, a phenomenon termed homoplasy (Herrnstadt et al. 2002) (Figure 1 (purple stars)). Another complication confounding haplogroup association studies is the high level of population stratification obtained with mtDNA analyses, due to mtDNA having an effective population size of one quarter the size of autosomal nuclear alleles. High levels of population stratification result in difficulties assembling appropriate control groups for case control studies, perhaps even requiring geographical matching of controls (Salas 2015). Traditional mtDNA haplogroup association studies have considered only the most common variants that define major haplogroups, as complete sequence data is not normally available for study cohorts.

As studies using traditional haplogroup association methods are likely to continue producing conflicting results, there is a need to develop and introduce new models, capable of determining if mtDNA variation is important in either the onset or course of complex human diseases. In this paper, we

explore a less frequently considered model in the context of mtDNA, one which also considers the impact of evolutionarily more recent variants (Elson et al. 2006; Hudson et al. 2014), or rare population variants that form the leaves of the human phylogeny (Figure 1 (yellow stars)). In this regard, rare variants that alter an amino acid have a higher likelihood of being mildly deleterious, as natural selection has had less opportunity to remove them from the population (Elson et al. 2004; Pereira et al. 2011). Importantly, the current analysis is not aimed at identifying whether a particular rare variant is more common in AD, but rather if rare variants as a group, in particular those with a higher likelihood of being mildly deleterious, are more common in the AD cases than in aged-controls.

Improved computational methods have recently been developed for determining the impact of non-synonymous changes on protein function, thereby increasing the power for finding differences between disease and control cohorts. An example of such a methodology is the analysis software program MutPred (Li et al. 2009), which has been validated in the context of mtDNA variation (Pereira et al. 2012; Pereira et al. 2011). MutPred is based upon the established SIFT (Sorting Intolerant From Tolerant) algorithm, which considers the gain or loss of 14 different structural and functional properties of proteins. While analysis tools such as MutPred (Li et al. 2009), or Ploythen2 (Adzhubei et al. 2010) alone are not sufficient in a diagnostic context (Elson et al. 2015; Yarham et al. 2011; Yarham et al. 2012). However the power of programs MutPred to differentiate between groups rich in deleterious variants and those consisting of benign polymorphic variants has been clearly demonstrated (Pereira et al. 2011).

The analysis reported in the current paper used best-available methodology and an adapted hypothesis, to re-analyse a substantial complete sequence dual location case/cohort, which was first reported in 2006 (Elson et al. 2006). Even after using these new methods and refinements of the hypothesis, the conclusion of this paper are in line with former work, namely that the frequency of rare mtDNA variation is not strikingly different in AD cases when compared to aged controls. This finding argues against a role of inherited germ-line mtDNA variants in AD.

The most interesting result of the current study is that “mutational load” was significantly lower in the aged US controls, when compared to the UK controls. This is of interest since the US controls are

significantly older than those from the UK. In addition, the US controls were clinically identified as healthy older people, whereas there was only evidence of an absence of disease in the UK controls. This difference in age and evidence of good cognition between the two control groups allowed us to suggest that low mtDNA "mutational load" might be more common in "healthy" older individuals than previously anticipated.

2. Methods and Materials

The mtDNA sequences used in the current study were reported previously by Elson and colleagues (2006) and are publically available at: http://www.phylotree.org/mtDNA_seqs.htm (van Oven and Kayser 2009). The sequences comprise two independent cohorts termed the MitoKor (MK-US) and Medical Research Council (MRC-UK) sequences. The MK-US sequence set comprises the complete mtDNA coding regions for 75 AD patients and 64 aged controls, while the MRC-UK set contains 70 AD patients and 64 controls. Only mtDNA sequences from one of the major European haplogroups (H, V, U, K, T, J, I, X, W) are used in the analysis. As reported previously, AD was confirmed in immunohistologically-stained post-mortem brain sections by using standardised neuropathological criteria (Elson et al. 2006). All the examined brain tissue sections taken from control cases were free from AD-related disease neuropathologies (Elson et al. 2006). For the MRC-UK samples, samples were included as controls if the patient, in life, had never experienced neuropsychiatric- (e.g. dementia, depression or psychosis) or cardiovascular disease (including stroke). At the neuropathological level, only appropriate age-related changes was allowed, including the presence of some small-vessel disease. The MRC-UK control brains also had to have <2 tangles/mm^2 in the cortical lobes and <5 senile plaques/mm^2 to be included in the control cohort. The 64 MK-US control sequences were derived from three sample sets: (1) 13 autopsy-confirmed brain samples that showed normal histology, with the patients from this group that had a mean age of 81 years; (2) 21 blood samples from a cohort of healthy elderly individuals (mean age = 88 years);

and (3) 30 blood samples taken from a cohort of individuals that were >80 years of age and evaluated as being cognitively normal over a period of at least 6 years (which extended up to 16 years).

Coding-region mtDNA sequences were obtained and analyzed as previously described (Herrnstadt et al. 2002). This approach incorporates several quality control measures, including a four-fold redundancy to minimize sequencing errors. Sequencing was conducted in a forward and reverse direction, with a 50% overlap of fragments (Herrnstadt et al. 2002). "Mutational load" values are given as the mean ± SD (standard deviation). The scores were compared using a one-way ANOVA, followed by a Tukey post-hoc test. All statistical analyses were conducted using GraphPad Prism software (version 6.05, GraphPad Software, Inc, La Jolla, CA, USA). The following P-value designations were applied: ***P <0.001, extremely significant; **$P \leq 0.01$, highly significant; *$P \leq 0.05$, significant and $P > 0.05$, non-significant (n/s). The present study explores a possible role for rare inherited mtDNA mutations in the susceptibility to AD, building on our previously published work (Elson et al. 2006). Only substitutions that were operationally homoplasmic were analyzed, with the sequencing having been conducted by traditional Sanger methodology, as heteroplasmies lower than ~15% could not be reliably detected. However, this would not result in missing important variants in the context of a study focused on inherited population variants. Previous studies by mitochondrial clinics and diagnostic centres have shown that mtDNA deletions and mtDNA point mutations do not create a biochemical defect resulting in a clinical defect until a "threshold level" is exceed. This threshold level varies by mutation, but is >60% and can be as high as 90% (Tuppen et al. 2010).

3. Results

In order to assess the possible phenotypic impact of the variants in the cases compared to controls, we initially compiled a list of non-synonymous variants that were present exclusively in either the AD cases

or the aged controls. We then compared the global frequency of these non-synonymous variants using a dataset consisting of 30,589 complete human mtDNA sequences; this global mtDNA dataset is available from the MitoMap database (updated 04/08/2015) (Lott et al. 2013). The rationale for this comparison being that variants that are mildly deleterious are likely to be less frequent in the population as selection acts at a number of levels to remove them (Elson et al. 2004; Stewart et al. 2008). Post-germline selection operating one mildly deleterious variants is relatively slow. Therefore, a statistically significant increase in the frequency of rare mtDNA variants in AD cases, relative to controls, is evidence that the variants play a role in the etiology of the disease. An independent samples Student's *t*-test was applied to determine whether a difference exists in the global frequency of the variants exclusive to the AD and those in control datasets. Results reveal no statistically significant difference ($P = 0.355$, n/s). However, this comparison has the drawback that it includes a number of haplogroup markers, which are relatively common in the population, but potentially different between the AD cases and controls due to sampling effects. This problem manifests even more acutely in traditional haplogroup association studies (Salas 2015), where it is normally addressed by clumping together haplogroups with frequencies of less than 5% into a single group.

In an attempt to avoid this problem, we ranked the data to consider the number of variants seen only in the AD cases or in the controls that were present at a frequency of <0.1% in the global dataset (Pereira et al. 2011). This approach allowed us to explore the possibility that very rare population variants cumulatively might be more common in AD patients than in aged controls, with mildly deleterious variants being expected to be less frequent in populations, as they are removed by natural selection. We observed 22 variants, present at a frequency of <0.1%, in the global database of 30,589 complete sequences in the AD patients, with 23 such variants being present in the control sequences, suggesting that there was no difference in the number of very rare variants across the two cohorts. This finding supports the view that a methodology based solely on the frequency of the alleles is not able to reveal differences between cohorts of a moderate size.

Next, we used the program MutPred (Li et al. 2009) to assess the likely impact of all non-synonymous changes seen in the cohort. The MutPred algorithm uses sequence conservation and protein structure to make predictions about the impact of a non-synonymous change on protein function and is therefore a more sophisticated tool for detecting mildly deleterious variants than a pure frequency count, even one using such a large global dataset (Li et al. 2009).

After calculating the MutPred scores for each variant, we summed the scores of the non-synonymous changes on each sequence to produce a "total mutational load". Mutational loads are given for the different groups as means with SD. Suppl. Table 1 (a-d) shows the haplogroup assignment (Weissensteiner et al. 2016), number of variants and "mutational loads" for each of the sequences used in the current study. The AD cases from the UK had a mean "total mutational load" of 1.798 ± 0.88, with those from the US that showed a very similar mean level (1.789 ± 0.89). The UK aged controls had a mean "total mutational load" of 2.273 ± 0.92, whilst the US controls had a notably lower mean load of 1.456 ± 0.78. Comparing all four groups with a one-way ANOVA reveled an extremely significant difference ($P < 0.0001$), with a Tukey post-hoc test indicating a significant difference between the UK AD and control cases ($P = 0.0096$). However, contrary to prediction, the mutational loads of the controls were higher than those of the AD patients. This surprising difference was not replicated in the AD cases and controls from the US ($P = 0.115$, n/s). There was also a significant difference between the UK control group and the AD cases from the US ($P = 0.0066$), but again in the opposite direction to what would have been predicted. Finally, there was a striking difference between the two control cohorts ($P < 0.0001$) (Figure 2a). It is worth noting that the number of variants, including non-synonymous changes seen in an individual, is affected by the proximity of an individual's mtDNA sequence to the reference sequence (termed the "revised Cambridge Reference Sequence" (rCRS)) on the phylogeny (Bandelt et al. 2014). An example is the large number of variants reported for many African mtDNA sequences (Bandelt et al. 2014; Herrnstadt et al. 2002). To account for this effect, an adjusted total mutational load was calculated by dividing the "total mutational load" of each sequence by the number of non-synonymous variants on that sequence, to obtain a metric that describes the average impact of mtDNA variation on an individual's

mtDNA sequence. This adjustment allowed us to test if AD sufferers have, on average, non-synonymous mtDNA variants that were more deleterious than controls. However, an adjusted mutational load will still include haplogroup markers, and thus not be entirely free from haplogroup context. The "adjusted total mutational load" for the UK AD cases was 0.41 ± 0.062 with those from the US being 0.41 ± 0.041. The UK controls had an "adjusted total mutational load" of 0.41 ± 0.048 and those from the US had 0.41 ± 0.071. Comparing the four groups using a one-way ANOVA reveled no significant differences (P = 0.395, n/s) (Figure 2b).

It is worth noting that the MutPred program designated scores of 0.5 or greater as being "actionable hypotheses" and those of 0.75 or greater as being "confident hypotheses", which indicates the likelihood that the substitution has a deleterious effect on protein function, with those below 0.3 being most likely to be variants that do not impact protein function (Li et al. 2009; Pereira et al. 2011).

A potential concern with the "total mutational load" and "adjusted total mutational load" methodology is that signal from a small number of potentially mildly deleterious mtDNA variants with high MutPred scores (>0.5) might be "averaged away" if an individual also has a large number of variants with low MutPred scores. Therefore, we decided to focus the analysis on the amino acids replacements that hold the strongest likelihood to exert a negative phenotypic effect. The UK AD cases had a mean ">0.5 total mutational load" of 0.52 ± 0.56 whilst those from the US were 0.58 ± 0.65. The UK controls that had a ">0.5 total mutational load" were 0.86 ± 0.71, with those from the US that was 0.3798 ± 0.47. Comparing all four groups using a one-way ANOVA revealed that there were significant differences (P < 0.0001), whilst application of a Tukey post-hoc test showed a significant difference between the UK AD and control cases (P = 0.0064), in the opposite direction to what was predicted. Once again this observation was not replicated in the AD cases and US controls (P = 0.576, n/s) (Figure 3a). Again, the analysis highlighted a very significant difference between the two control cohorts (P < 0.0001) (Figure 3a). When the ">0.5 total mutational load" was adjusted for the number of mtDNA variants seen in the sequences, an ANOVA revealed borderline significance (P = 0.054). A Tukey post-hoc test then revealed that the UK and US controls showed statistically significant differences (P = 0.0388), with the older, healthier

individuals showing a lower mutational load (Figure 3b). The ">0.5 total mutational load" and ">0.5 adjusted mutational load" and the number of associated variants for each sequence used in this analysis is displayed in Suppl. Table 1 (a-d).

4. Discussion

A wealth of evidence suggests a central role of mitochondrial dysfunction and oxidative stress in AD pathology and other neurodegenerative diseases (Coskun et al. 2006; Cottrell et al. 2002); however, it is unresolved if *inherited* mtDNA variants are involved in the etiology of such diseases (Bi et al. 2015; Coto et al. 2011; Elson et al. 2006; Hudson et al. 2012; Maruszak et al. 2009). This lack of clarity is also seen in a number of other human diseases (e.g. Diabetes (Chinnery et al. 2007; Torroni et al. 1997)) where mtDNA variation is hypothesized to play a role. The traditional haplogroup association model has been applied to obtain evidence for an etiological role but the myriad conflicting results have been obtained. This impasse argues strongly for the development of new association models to connect mtDNA variation with common complex diseases (Salas 2015). The traditional haplogroup association model aims to detect differences in the frequency of major haplogroups between cases and disease controls, with haplogroups being defined by the presence or absence of ancient (common) population variants. MtDNA has a low effective population size when compared to nuclear alleles, due to both maternal inheritance and its haploid nature. This low effective population size of mtDNA can result in stratification of these ancient variants, creating differences between populations at different geographical locations solely due to genetic drift [31]. Therefore, clines in gene frequency might be highly revealing for investigations aimed at studying population history, but they are problematic in the context of mtDNA association studies.

In this study, we applied new approaches that avoid some of the limitations and pitfalls of traditional haplogroup association studies. The "mutational load" approaches described here have more statistical power than traditional haplogroup association studies, as the mutational load analyses use fewer tests and allow the application of more powerful parametric statistical analysis. With the relatively large number of complete mtDNA sequences (n = 273) that were analyzed here, our study had 80% power to detect a 20%

difference in total mutational load between cases and controls at the 0.05 significance level. This level of power compares well to reported estimates of the power of haplogroup association studies (Samuels et al. 2006). Therefore, we propose that this method should be used – in conjunction with other approaches - in future mtDNA association studies. The advent of cheaper and more user-friendly tools for the analysis of next generation sequencing data will bring about a rise in analyses that use complete mtDNA sequence data to conduct "mutational load"-type analysis.

Our work also supports the view that subsequent mtDNA association studies should analyze only complete mtDNA sequence data and that the effects of rare variants should be analyzed as a general rule. Critically, this novel method focuses on inherited variants with high MutPred scores. Variants with high MutPred scores tend to be rare variants at the tip (or leaves) of the phylogeny and are thus not prone to stratification (Elson et al. 2004; Pereira et al. 2011). Therefore, it is also suggested that use of metrics such as mutational load and adjusted mtDNA mutational load can help avoid the false positive associations seen in the past (Salas 2015), when studying the role of the mitochondrial chromosome in the onset and development of common/complex human disease.

In terms of new models proposed for sequence data analysis, a model that considers only variants with a MutPred score >0.5 seems biologically most plausible. In addition, it can be argued that when using these variants, there is no longer a need to adjust for the position of the sequence on the phylogeny as very few common haplogroup-associated variants score above this level. As one only looks at variants of likely deleterious effect, with recent studies suggesting that natural selection has been equally efficient at removing deleterious variants in all lineages, these variants will be at the tips or leaves of the phylogeny in the vast majority of cases (Pereira et al. 2011).

In summary, the current study used new "mtDNA mutational load" methods to determine if mtDNA variation might play a role in the susceptibility to AD. Application of such new methods do not support a link between either the combined effect of all non-synonymous mtDNA variants on a person's mtDNA, nor the combined effect of variants predicted to be of deleterious effect in the susceptibility to AD. Additionally, an examination of the data does not suggest there is a sub-group of our AD patients within

the cohort that have a higher mutational load (Figures 1 and 2; Supplemental Tables (a-d)). This suggests that studies on larger cohorts, which apply the same methods, are likely to arrive at a similar conclusion, but this cannot be stated with certainty. Surprisingly, however, the results suggest a role for "low mtDNA load" in maintaining good cognition in older individuals, with the strongest difference being between the control cohorts from the two locations. This is of interest, as our previous analysis of this dataset reported that the US aged controls were significantly older than the UK aged controls, while the age of the AD cohort did not differ between the two countries. As we reported previously (Elson et al. 2006), the mean age (and SD) of the MRC control group was 77.2 ± 9.6 years, whereas the mean age of the MitoKor control group was 83.4 ± 5 years. A two-tailed *t*-test (without an assumption of equal SDs) indicates that the MitoKor control group is significantly older than the MRC controls. There was also evidence to support that a significant proportion of the US cohort had aged well. Thus US control cohort might represent a more stringently assessed collection of cases and examples of particularly healthy old age individuals (Elson et al. 2006). The significant difference in ">0.5 total mutational load" and "adjusted total mutational load" in the UK and US controls suggests that mtDNA genotypes with fewer mildly deleterious mutations might be more common in the healthy old, as compared to the old that are simply free from disease, and thus low mtDNA mutational loads might be neuroprotective. This important finding merits additional investigation and analysis.

Acknowledgements

This research did not receive any specific grant support from funding agencies in the public, commercial, or not-for-profit sectors.

Author Contributions

JLE: Thought of the concept, conducted analysis and wrote manuscript

NH: Provided sequence data and reviewed manuscript providing critical comment

ISP: Conducted analysis and reviewed manuscript providing critical comment

References

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR (2010) A method and server for predicting damaging missense mutations. Nat Methods. 7: 248-9. doi: 10.1038/nmeth0410-248.

Bandelt HJ, Kloss-Brandstätter A, Richards MB, Yao YG, Logan I (2014) The case for the continuing use of the revised Cambridge Reference Sequence (rCRS) and the standardization of notation in human mitochondrial DNA studies. J Hum Genet 59: 66-77. doi: 10.1038/jhg.2013.120. Epub 2013 Dec 5.

Bi R, Zhang W, Yu D, Li X, Wang HZ, Hu QX, Zhang C, Lu W, Ni J, Fang Y, Li T, Yao YG (2015) Mitochondrial DNA haplogroup B5 confers genetic susceptibility to Alzheimer's disease in Han Chinese. Neurobiol Aging 36: 1604.e7-16. doi: 10.1016/j.neurobiolaging.2014.10.009. Epub 2014 Oct 22.

Blennow K, de Leon MJ, Zetterberg H (2006) Alzheimer's disease. Lancet. 368: 387-403.

Chinnery PF, Mowbray C, Patel SK, Elson JL, Sampson M, Hitman GA, McCarthy MI, Hattersley AT, Walker M (2007) Mitochondrial DNA haplogroups and type 2 diabetes: a study of 897 cases and 1010 controls. J Med Genet. 44: e80.

Coskun P, Wyrembak J, Schriner SE, Chen HW, Marciniack C, Laferla F, Wallace DC (2006) Impaired platelet mitochondrial activity in Alzheimer's disease and mild cognitive impairment. Mitochondrion 6: 323-30. Epub 2006 Oct 27.

Coto E, Gómez J, Alonso B, Corao AI, Díaz M, Menéndez M, Martínez C, Calatayud MT, Morís G, Álvarez V (2011) Late-onset Alzheimer's disease is associated with mitochondrial DNA 7028C/haplogroup H and D310 poly-C tract heteroplasmy. Neurogenetics 12: 345-6. doi: 10.1007/s10048-011-0295-4. Epub 2011 Aug 7.

Cottrell DA, Borthwick GM, Johnson MA, Ince P, Turnbull DM (2002) The role of cytochrome c oxidase deficient hippocampal neurones in Alzheimer's disease. Neuropathol Appl Neurobiol 28: 390-6.

Elson JL, Herrnstadt C, Preston G, Thal L, Morris CM, Edwardson JA, Beal MF, Turnbull DM, Howell N (2006) Does the mitochondrial genome play a role in the etiology of Alzheimer's disease? Hum Genet 119: 241-54. Epub 2006 Jan 12.

Elson JL, Smith PM, Greaves LC, Lightowlers RN, Chrzanowska-Lightowlers ZM, Taylor RW, Vila-Sanjurjo A (2015) The presence of highly disruptive 16S rRNA mutations in clinical samples indicates a wider role for mutations of the mitochondrial ribosome in human disease. Mitochondrion. 25:17-27.: 10.1016/j.mito.2015.08.004. Epub 2015 Sep 5.

Elson JL, Turnbull DM, Howell N (2004) Comparative genomics and the evolution of human mitochondrial DNA: assessing the effects of selection. Am J Hum Genet 74: 229-38. Epub 2004 Jan 7.

Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, Pahwa JS, Moskvina V, Dowzell K, Williams A, Jones N, Thomas C, Stretton A, Morgan AR, Lovestone S, Powell J, Proitsi P, Lupton MK, Brayne C, Rubinsztein DC, Gill M, Lawlor B, Lynch A, Morgan K, Brown KS, Passmore PA, Craig D, McGuinness B, Todd S, Holmes C, Mann D, Smith AD, Love S, Kehoe PG, Hardy J, Mead S, Fox N, Rossor M, Collinge J, Maier W, Jessen F, Schürmann B, Heun R, van den Bussche H, Heuser I, Kornhuber J, Wiltfang J, Dichgans M, Frölich L, Hampel H, Hüll M, Rujescu D, Goate AM, Kauwe JS, Cruchaga C, Nowotny P, Morris JC, Mayo K, Sleegers K, Bettens K, Engelborghs S, De Deyn PP, Van Broeckhoven C, Livingston G, Bass NJ, Gurling H, McQuillin A, Gwilliam R, Deloukas P, Al-Chalabi A, Shaw CE, Tsolaki M, Singleton AB, Guerreiro R, Mühleisen TW, Nöthen MM, Moebus S, Möckel KH, Klopp N, Wichmann HE, Carrasquillo MM, Pankratz VS, Younkin SG, Holmans PA, O'Donovan M, Owen MJ, Williams J (2009) Genome-wide association study identifies variants at CLU and

PICALM associated with Alzheimer's disease. Nat Genet 41: 1088-93. doi: 10.1038/ng.440. Epub 2009 Sep 6.
Herrnstadt C, Elson JL, Fahy E, Preston G, Turnbull DM, Anderson C, Ghosh SS, Olefsky JM, Beal MF, Davis RE, Howell N (2002) Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. Am J Hum Genet 70: 1152-71. Epub 2002 Apr 5.
Howell N, Elson JL, Chinnery PF, Turnbull DM (2005) mtDNA mutations and common neurodegenerative disorders. Trends Genet 21: 583-6. Epub 2005 Sep 8.
Hudson G, Gomez-Duran A, Wilson IJ, Chinnery PF (2014) Recent mitochondrial DNA mutations increase the risk of developing common late-onset human diseases. PLoS Genet 10: e1004369. doi: 10.1371/journal.pgen.1004369. eCollection 2014 May.
Hudson G, Sims R, Harold D, Chapman J, Hollingworth P, Gerrish A, Russo G, Hamshere M, Moskvina V, Jones N, Thomas C, Stretton A, Holmans PA, O'Donovan MC, Owen MJ, Williams J, Chinnery PF, Consortium. G (2012) No consistent evidence for association between mtDNA variants and Alzheimer disease. Neurology 78: 1038-42. doi: 10.1212/WNL.0b013e31824e8f1d. Epub 2012 Mar 21.
Li B, Krishnan VG, Mort ME, Xin F, Kamati KK, Cooper DN, Mooney SD, Radivojac (2009) Automated inference of molecular mechanisms of disease from amino acid substitutions. Bioinformatics 25: 2744-50. doi: 10.1093/bioinformatics/btp528. Epub 2009 Sep 3.
Lott MT, Leipzig JN, Derbeneva O, Xie HM, Chalkia D, Sarmady M, Procaccio V, Wallace DC (2013) mtDNA Variation and Analysis Using Mitomap and Mitomaster. Curr Protoc Bioinformatics 44:1.23.1-26.: 10.1002/0471250953.bi0123s44.
Maruszak A, Canter JA, Styczyńska M, Zekanowski C, Barcikowska M (2009) Mitochondrial haplogroup H and Alzheimer's disease--is there a connection? Neurobiol Aging 30: 1749-55. doi: 10.1016/j.neurobiolaging.2008.01.004. Epub 2008 Mar 4.
Maruszak A, Zekanowski C (2011) Mitochondrial dysfunction and Alzheimer's disease. Prog Neuropsychopharmacol Biol Psychiatry 35: 320-30. doi: 10.1016/j.pnpbp.2010.07.004. Epub 2010 Jul 16.
Pereira L, Soares P, Máximo V, Samuels DC (2012) Somatic mitochondrial DNA mutations in cancer escape purifying selection and high pathogenicity mutations lead to the oncocytic phenotype: pathogenicity analysis of reported somatic mtDNA mutations in tumors. BMC Cancer 12:53.: 10.1186/1471-2407-12-53.
Pereira L, Soares P, Radivojac P, Li B, Samuels DC (2011) Comparing phylogeny and the predicted pathogenicity of protein variations reveals equal purifying selection across the global human mtDNA diversity. Am J Hum Genet 88: 433-9. doi: 10.1016/j.ajhg.2011.03.006. Epub 2011 Mar 31.
Salas AE, J.L. (2015) Mitochondrial DNA as a risk factor for false positives in case-control association studies. J Genet Genomics 42: 169-72. doi: 10.1016/j.jgg.2015.03.002. Epub 2015 Mar 17.
Samuels DC, Carothers AD, Horton R, Chinnery PF (2006) The power to detect disease associations with mitochondrial DNA haplogroups. Am J Hum Genet 78: 713-20. Epub 2006 Feb 17.
Stewart JB, Freyer C, Elson JL, Larsson NG (2008) Purifying selection of mtDNA and its implications for understanding evolution and mitochondrial disease. Nat Rev Genet. 9: 657-62. doi: 10.1038/nrg2396.
Torroni A, Petrozzi M, D'Urbano L, Sellitto D, Zeviani M, Carrara F, Carducci C, Leuzzi V, Carelli V, Barboni P, De Negri A, Scozzari R (1997) Haplotype and phylogenetic analyses suggest that one European-specific mtDNA background plays a role in the expression of Leber hereditary optic neuropathy by increasing the penetrance of the primary mutations 11778 and 14484. Am J Hum Genet. 60: 1107-21.

Tuppen HA, Blakely EL, Turnbull DM, Taylor RW (2010) Mitochondrial DNA mutations and human disease. Biochim Biophys Acta. 1797: 113-28. doi: 10.1016/j.bbabio.2009.09.005. Epub 2009 Sep 15.

Valla J, Schneider L, Niedzielko T, Coon KD, Caselli R, Sabbagh MN, Ahern GL, Baxter L, Alexander G, Walker DG, Reiman EM (2006) Impaired platelet mitochondrial activity in Alzheimer's disease and mild cognitive impairment. Mitochondrion 6: 323-30. Epub 2006 Oct 27.

van Oven M, Kayser M (2009) Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum Mutat. 30: E386-94. doi: 10.1002/humu.20921.

Wang HZ, Bi R, Hu QX, Xiang Q, Zhang C, Zhang DF, Zhang W, Ma X, Guo W, Deng W, Zhao L, Ni P, Li M, Fang Y, Li T, Yao YG (2014) Validating GWAS-Identified Risk Loci for Alzheimer's Disease in Han Chinese Populations. Mol Neurobiol 3: 3.

Weissensteiner H, Pacher D, Kloss-Brandstatter A, Forer L, Specht G, Bandelt HJ, Kronenberg F, Salas A, Schonherr S (2016) HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. Nucleic Acids Res. 44: W58-63. doi: 10.1093/nar/gkw233. Epub 2016 Apr 15.

Yarham JW, Al-Dosary M, Blakely EL, Alston CL, Taylor RW, Elson JL, McFarland R (2011) A comparative analysis approach to determining the pathogenicity of mitochondrial tRNA mutations. Hum Mutat. 32: 1319-25. doi: 10.1002/humu.21575. Epub 2011 Sep 19.

Yarham JW, McFarland R, Taylor RW, Elson JL (2012) A proposed consensus panel of organisms for determining evolutionary conservation of mt-tRNA point mutations. Mitochondrion. 12: 533-8. doi: 10.1016/j.mito.2012.06.009. Epub 2012 Jul 7.

Figure 1.

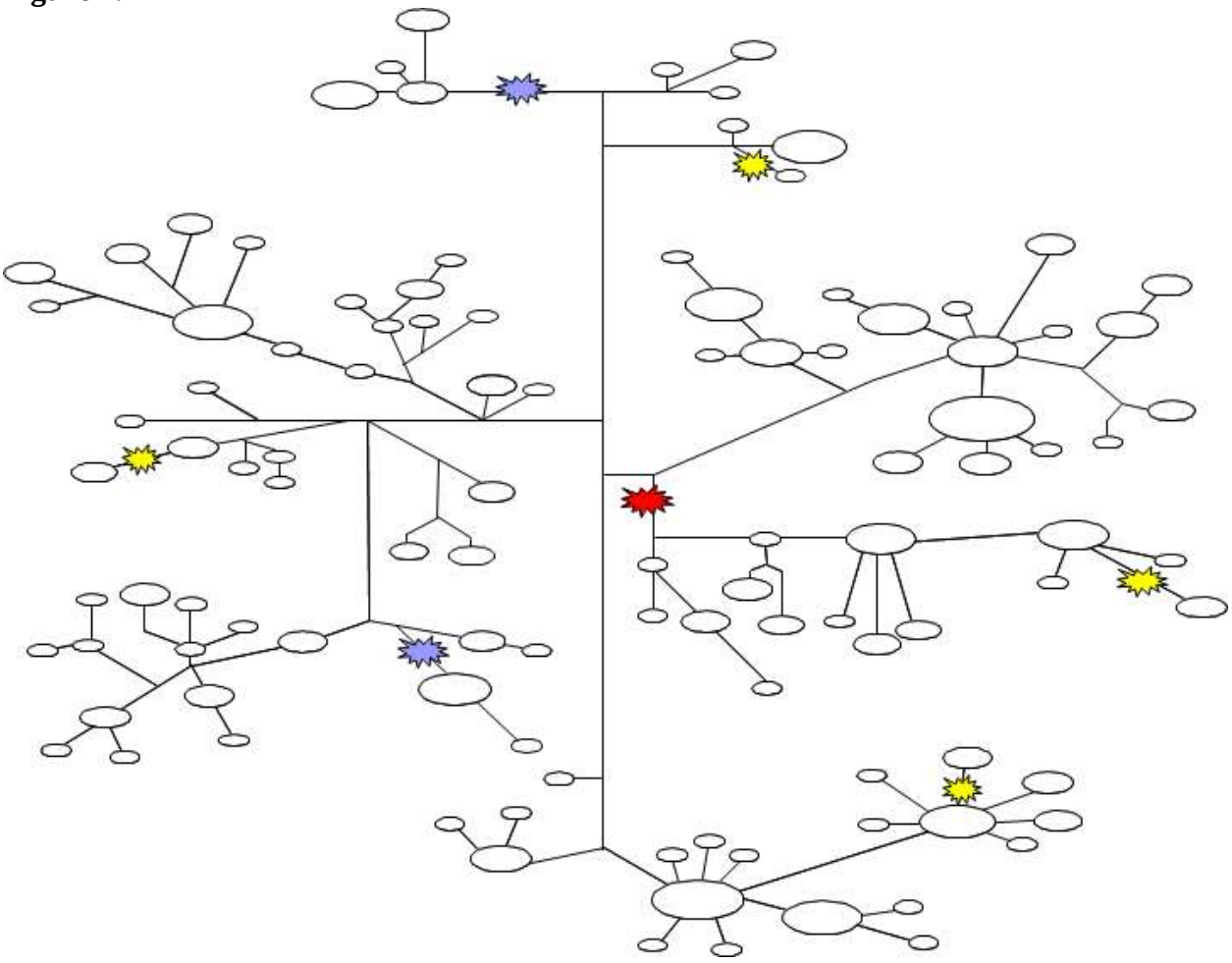
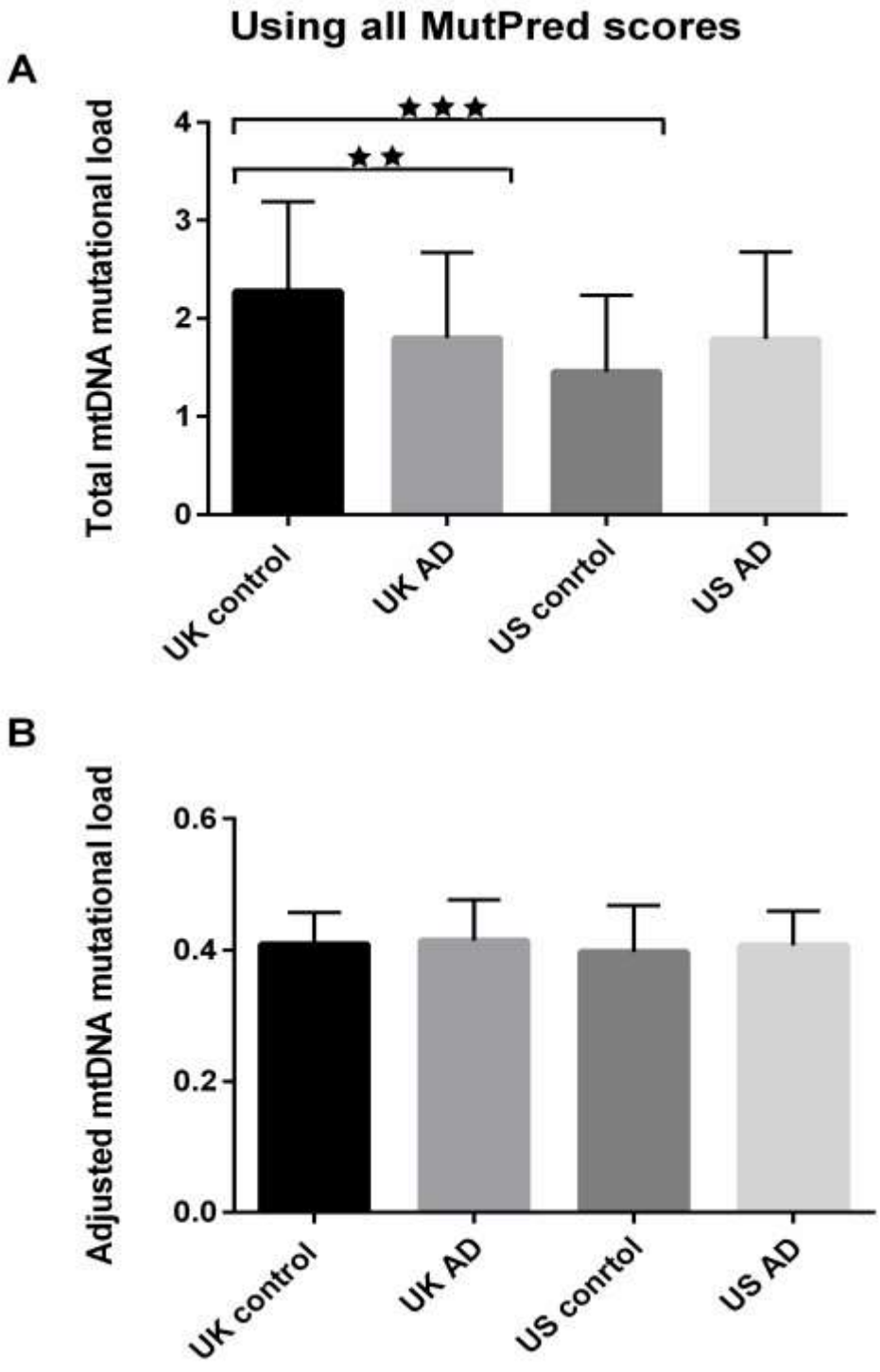


Figure 2 [original and modified versions – for decision]



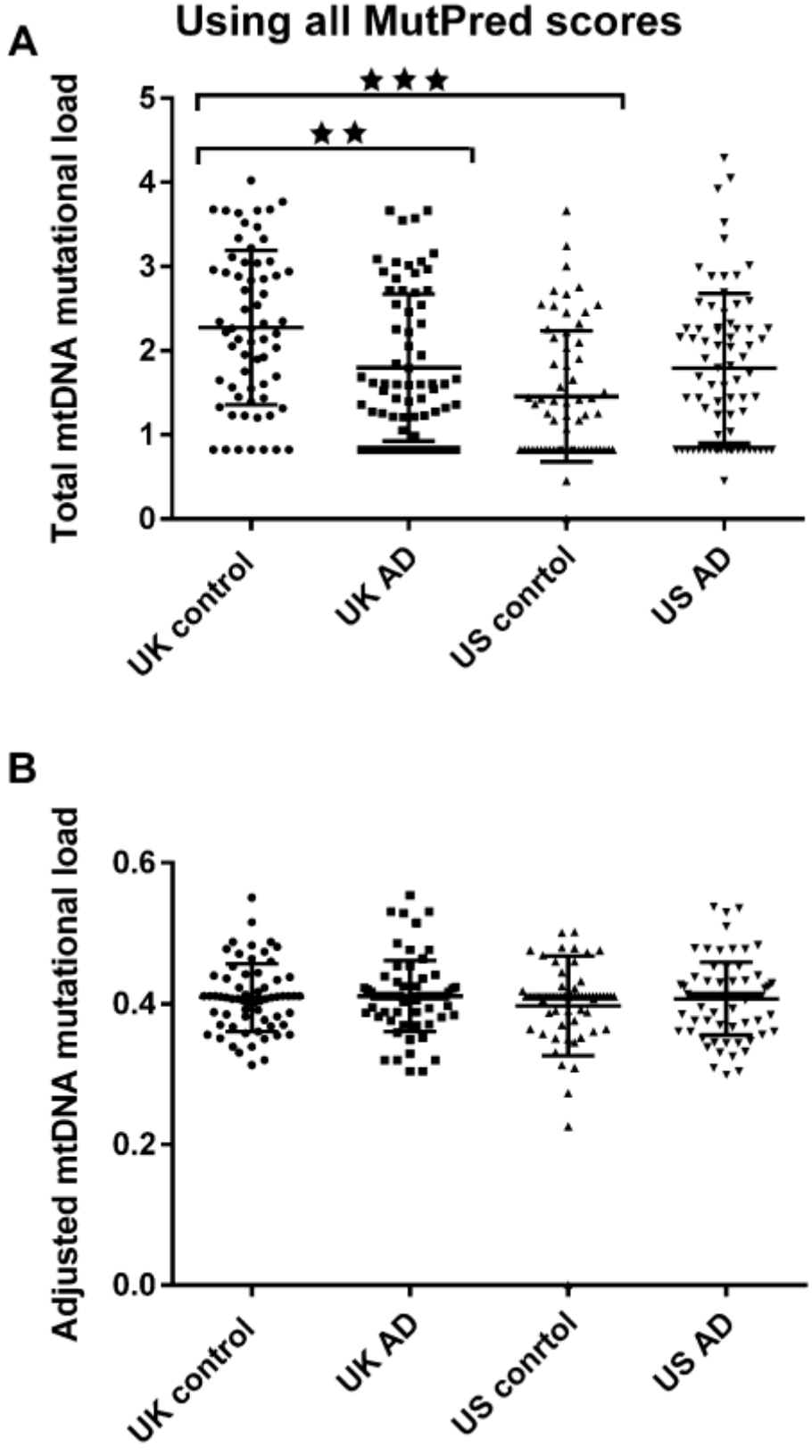
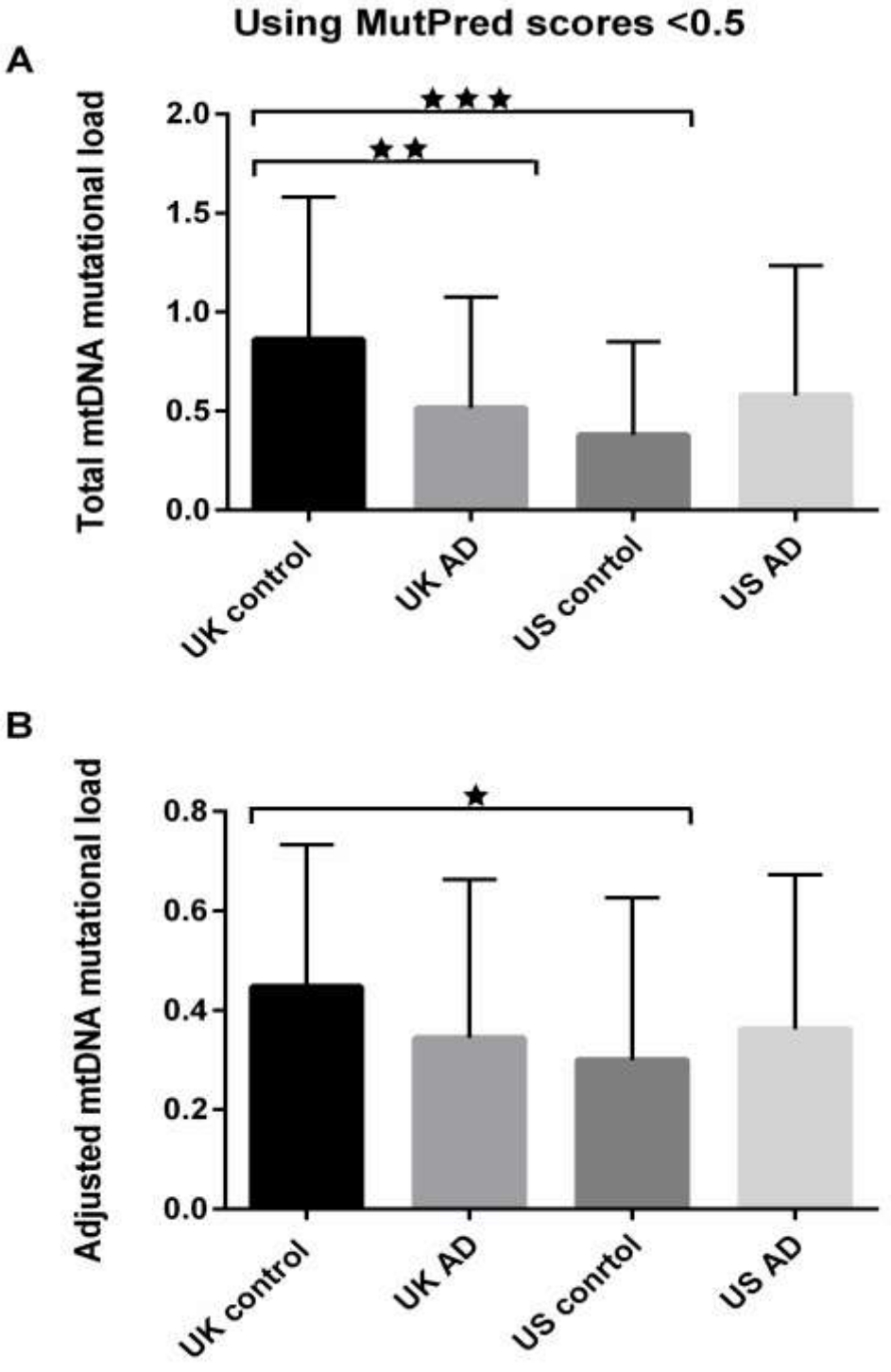

A
Using all MutPred scores
Total mtDNA mutational load
★★★
★★
5
4
3
2
1
0
UK control
UK AD
US conrtol
US AD
B
Adjusted mtDNA mutational load
0.6
0.4
0.2
0.0
UK control
UK AD
US conrtol
US AD

Figure 3 [original and modified versions – for decision]



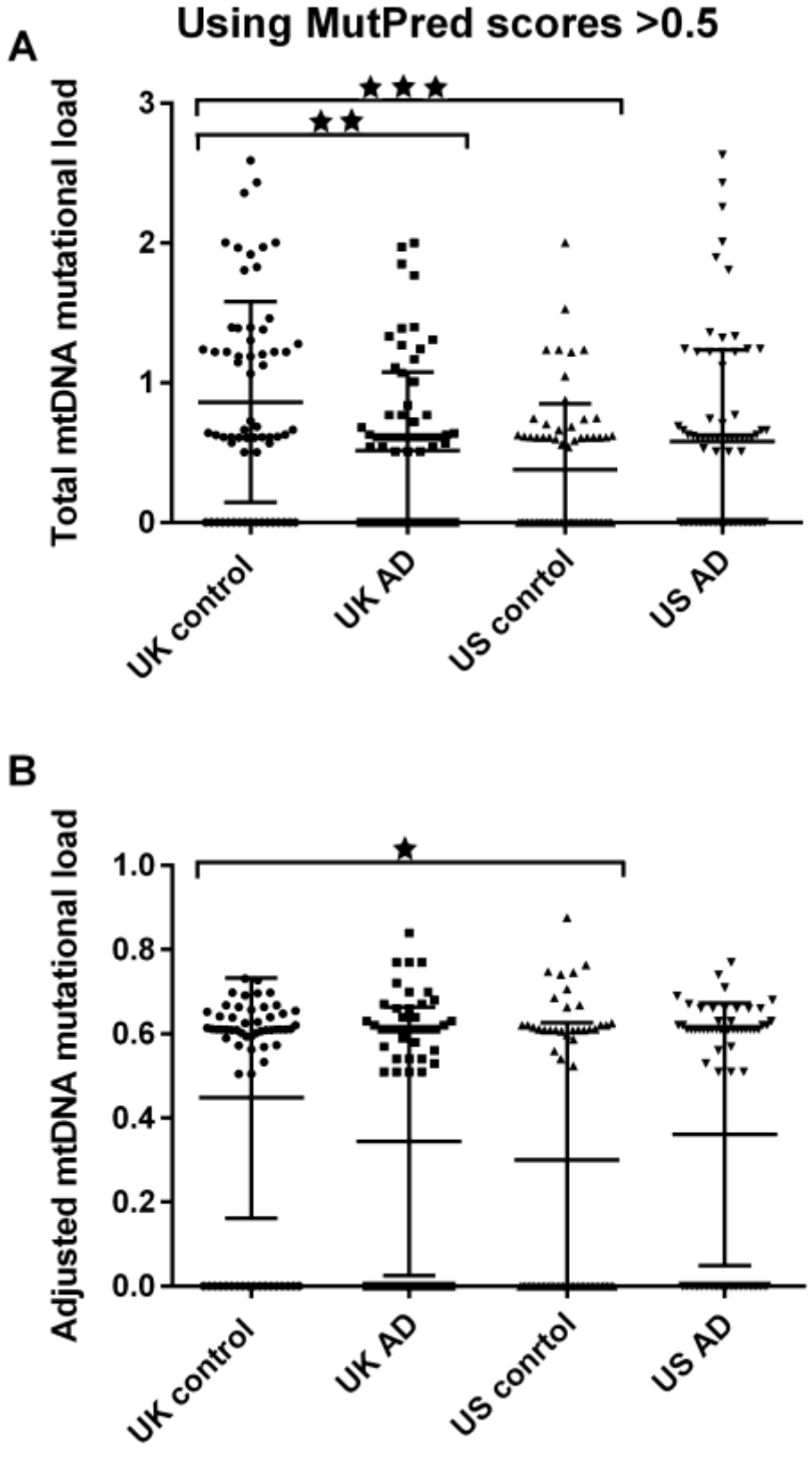

A
Using MutPred scores >0.5
Total mtDNA mutational load
3
2
1
0
UK control
UK AD
US conrtol
US AD
B
Adjusted mtDNA mutational load
1.0
0.8
0.6
0.4
0.2
0.0
UK control
UK AD
US conrtol
US AD

Figure 1

Schematic representation of the European phylogeny, showing the nine major haplogroups. Stars illustrate the hypothesis that might link mtDNA variation to disease. (i) Red star – the traditional haplogroup association hypothesis, a single common variant affects susceptibility or is protective; (ii) purple stars – how the traditional haplogroup association hypothesis might be complicated by the occurrence of such a variant more than once on the phylogeny, a homoplasy; and (iii) yellow stars – how there might be a cumulative effect of multiple rare variants.

Figure 2

(a) Total mutational loads using all non-synonymous substitutions for the AD cases and controls in the UK and US cohorts. Comparing all four groups with a one-way ANOVA revealed an extremely significant differences (***$P < 0.0001$), with a Tukey post-hoc test showing a highly significant difference between the UK AD and control cases (**$P = 0.0096$), which was not replicated when the AD and control cohort from the US was analysed ($P = 0.115$, n/s). Additionally, a difference was seen between the two control cohorts from the two continents (***$P < 0.0001$). **(b) Adjusted mtDNA mutational loads using all non-synonymous substitution (total load per sequence/number of variants per sequence).** No significant differences were observed.

Figure 3

(a) Mutational loads, including only those using all non-synonymous substitutions with MutPred scores >0.5 in AD cases and controls, in both the UK and US cohorts. A highly significant difference was seen between UK AD and control cases (**$P = 0.0064$), but not in the replication cohort from the US. Again there was an extremely significant difference seen between the two control cohorts (***$P < 0.0001$). **(b)** Adjusted mtDNA mutational loads (including only those variants with MutPred scores >0.5), in AD cases and controls from the UK and US cohorts showed the only significant difference after correcting for multiple comparisons between the controls from the US and the UK (*$P = 0.0388$).

Supplemental tables (a-d): Raw values by sequence for all the analysis conducted in this manuscript.

UK Contol	Haplogroup	Total load	No. varinats	Adjusted load	> 0.5 Total load	No. varinats	> 0.5Adjusted load
mrc71	X2b	2.349	6	0.392	0.569	1	0.569
mrc72	I1a1b	2.344	6	0.391	1.188	2	0.594
mrc73	I2a	2.261	6	0.377	0.606	1	0.606
mrc74	W3a1	2.719	7	0.388	0.505	1	0.505
mrc75	J2a1a1	3.635	9	0.404	1.972	3	0.657
mrc76	J2a1a1a2	3.059	8	0.382	1.396	2	0.698
mrc77	J1c	2.883	8	0.360	1.220	2	0.610
mrc78	J1c1b	3.666	9	0.407	2.003	3	0.668
mrc79	J1c1a	3.666	9	0.407	2.003	3	0.668
mrc80	H6a1b2b	1.393	3	0.464	0.572	1	0.572
mrc81	J1c4	3.218	9	0.358	1.220	2	0.610
mrc82	J1c5c	2.959	7	0.423	1.918	3	0.639
mrc83	J1c2g	3.469	9	0.385	1.806	3	0.602
mrc84	J1c	4.021	10	0.402	2.358	4	0.590
mrc85	J1c3b	3.677	10	0.368	1.220	2	0.610
mrc86	T2f1a1	3.518	8	0.440	2.434	4	0.609
mrc87	T2b3	3.05	7	0.436	1.966	3	0.655
mrc88	T2b3	3.676	8	0.460	2.592	4	0.648
mrc89	H1j	1.449	3	0.483	0.628	1	0.628
mrc90	H1e	1.326	3	0.442	0.505	1	0.505
mrc91	H1c14	2.928	6	0.488	1.188	2	0.594
mrc92	H1i1	1.897	4	0.474	0.641	1	0.641
mrc93	H1be	1.548	3	0.516	0.727	1	0.727
mrc94	H1y	2.202	4	0.551	1.381	2	0.691
mrc95	H1g1	1.435	3	0.478	0.614	1	0.614
mrc96	H2a1b	0.821	2	0.411			
mrc97	H3g1b	1.924	4	0.481	0.664	1	0.664
mrc98	H6a1a3	1.227	3	0.409			
mrc99	H31	0.821	2	0.411			
mrc100	H31	1.203	3	0.401			
mrc101	H5a1c1a	1.313	3	0.438			
mrc102	H	0.821	2	0.411			
mrc103	H7	1.23	3	0.410			
mrc104	H; H27a	1.952	4	0.488	0.686	1	
mrc105	U8a1a1	1.751	5	0.350			
mrc106	K2b1	2.539	6	0.423	0.609	1	0.609
mrc107	K2a6	3.767	8	0.471	1.828	3	0.609
mrc108	K1a4a1d	3.04	7	0.434	1.392	2	0.696
mrc109	K1a4a1	2.852	7	0.407	1.204	2	0.602
mrc110	K1	3.336	8	0.417	1.279	2	0.640
mrc111	K1c1	3.109	7	0.444	1.461	2	0.731
mrc112	K1c2	2.488	7	0.355	0.609	1	0.609
mrc113	U3a1	1.565	5	0.313			
mrc114	U5a1a2a	2.939	8	0.367	1.126	2	0.563
mrc115	U5a1a1	2.036	6	0.339			
mrc116	U5b2a6	2.83	7	0.404	1.303	2	0.652
mrc117	U5b2a2b	2.137	6	0.356	0.610	1	0.610
mrc118	U5b2a3	2.673	7	0.382	1.146	2	0.573
mrc119	J1c3a2	3.329	9	0.370	1.220	2	0.610
mrc120	T2b	1.695	5	0.339	0.611	1	0.611
mrc121	H1e2	0.821	2				
mrc122	H1	0.821	2	0.411			
mrc123	H1	0.821	2	0.411			
mrc124	H1h1	0.821	2	0.411			
mrc125	H6a1a2b1	1.227	3	0.409			
mrc126	U4b1b1	2.22	6	0.370	0.625	1	0.625
mrc127	U5b2a1b	2.137	6	0.356	0.610	1	0.610
mrc128	U2e2a1	2.052	5	0.410	1.066	2	0.533
mrc129	J2a1a1a2	2.889	7	0.413	1.396	2	0.698
mrc130	J1c5	2.104	6	0.351	0.611	1	0.611
mrc131	T2b	2.323	6	0.387	1.239	2	0.620
mrc132	H4a1a1	2.263	5	0.453	0.664	1	0.664
mrc133	K1	1.648	5	0.330			
mrc134	U5a1a1	1.601	5	0.320			

US Control	Haplogroup	Total load	No. varinats	Adjusted load	> 0.5 Total load	No. varinats	> 0.5Adjusted load
mk76	X2c1	2.099	6	0.350	0.741	1	0.741
mk77	I	1.82	5	0.364	0.606	1	0.606
mk78	J1b1a	2.758	8	0.345	0.611	1	0.611
mk79	J1c7a	3.243	9	0.360	1.22	2	0.61
mk80	T2b	2.545	7	0.364	1.239	2	0.6195
mk81	H1c10	1.38	3	0.460	0.559	1	0.559
mk82	H1c3	1.506	3	0.502	0.685	1	0.685
mk83	H1c2a	1.429	3	0.476	0.608	1	0.608
mk84	H1	1.436	3	0.479	0.615	1	0.615
mk85	H1a5	1.902	4	0.476	0.706	1	0.706
mk86	H1b1	1.25	4	0.313	0.54	1	0.54
mk87	H1a1	0.821	2	0.411			
mk88	H1n	0.821	2	0.411			
mk89	H1	0.821	2	0.411			
mk90	H1az	0.821	2	0.411			
mk91	V2c	2.03	5	0.406	0.746	1	0.746
mk92	H40b	0.821	2	0.411			
mk93	H5k	0.821	2	0.411			
mk94	HV5	1.171	3	0.390			
mk95	H2a1	0.821	2	0.411			
mk96	H2a2b1a	0.821	2	0.411			
mk97	H2a2a1						
mk98	H2a	0.821	2	0.411			
mk99	H3h2	1.408	3	0.469	0.587	1	0.587
mk100	H3a1	1.441	3	0.480	0.62	1	0.62
mk101	H4a1a1	2.672	6	0.445	0.664	1	0.664
mk102	H6a1	1.569	4	0.392	0.748	1	0.748
mk103	H53	1.838	4	0.460	0.611	1	0.611
mk104	H7d3a	1.418	3	0.473	0.597	1	0.597
mk105	H5a6	1.363	5	0.273	1.048	2	0.524
mk106	H5a1a	0.821	2	0.411			
mk107	H5a1	0.821	2	0.411			
mk108	H5a1	1.07	3	0.357			
mk109	H	1.253	3	0.418			
mk110	U4a1b	2.531	6	0.422	0.876	1	0.876
mk111	U4a1a	1.655	5	0.331			
mk112	K	3.006	6	0.501	1.528	2	0.764
mk113	K1a4a1a2b	2.461	7	0.352	0.609	1	0.609
mk114	K1b2b	2.454	7	0.351	0.609	1	0.609
mk115	K1c2	2.716	7	0.388	0.609	1	0.609
mk116	T1a1	2.323	6	0.387	1.239	2	0.6195
mk117	T2b	2.158	5	0.432	1.239	2	0.6195
mk118	H1c7	0.821	2	0.411			
mk119	H1c3	0.821	2	0.411			
mk120	H1c3	0.821	2	0.411			
mk121	H1c2	0.821	2	0.411			
mk122	H3c2b	0.821	2	0.411			
mk123	H3a1a	0.821	2	0.411			
mk124	H3ap	0.821	2	0.411			
mk125	H3	0.821	2	0.411			
mk126	H	0.821	2	0.411			
mk127	H6a1a5	1.227	3	0.409			
mk128	U4b1b1	2.22	6	0.370	0.625	1	0.625
mk129	K2a2a1	2.548	6	0.425	0.609	1	0.609
mk130	U5b1	1.235	4	0.309			
mk131	J1c1a	3.666	9	0.407	2.003	3	0.667666667
mk132	HV	0.821	2	0.411			
mk133	H2a2	0.821	2	0.411			
mk134	H31	0.452	2	0.226			
mk135	H1	0.821	2	0.411			
mk136	H5a1	1.171	3	0.390			
mk137	H	0.821	2	0.411			
mk138	X2b	0.821	2	0.411			
mk139	J1c1a	2.257	6	0.376	0.609	1	0.609

UK AD	Haplogroup	Total load	No. varinats	Adjusted load	> 0.5 Total load	No. varinats	> 0.5Adjusted load
mrc1	X2b	3.09	7	0.441	1.310	2	0.655
mrc2	J1c1a	3.666	9	0.407	1.392	3	0.696
mrc3	J1c1	3.666	9	0.407	2.003	3	0.668
mrc4	W1e	3.571	9	0.397	0.505	1	0.505
mrc5	W5a1a1	2.719	7	0.388	0.505	1	0.505
mrc6	W5a	2.687	7	0.384	1.011	2	0.506
mrc7	J2a1a1a	3.059	8	0.382	1.396	2	0.698
mrc8	T2; T2f	2.938	7	0.420	1.854	3	0.618
mrc9	T2b17	2.323	6	0.387	1.239	2	0.620
mrc10	T2b3b	3.05	7	0.436	1.966	3	0.655
mrc11	R0a2	1.525	4	0.381	0.539	1	0.539
mrc12	R0a	0.986	3	0.329			
mrc13	HV16	1.252	3	0.417			
mrc14	HV	0.821	2	0.411			
mrc15	HV	0.821	2	0.411			
mrc16	V	0.821	2	0.411			
mrc17	V10	0.821	2	0.411			
mrc18	V3a1	1.209	3	0.403			
mrc19	HV17	1.544	3	0.515	0.723	1	0.723
mrc20	HV9	0.821	2	0.411			
mrc21	H1am	0.821	2	0.411			
mrc22	H1ag1	1.661	3	0.554	0.840	1	0.840
mrc23	H1ag1	1.593	3	0.531	0.772	1	0.772
mrc24	H1	1.593	3	0.531	0.772	1	0.772
mrc25	H1b1	1.687	4	0.422	0.642	1	0.642
mrc26	H1	1.361	3	0.454	0.540	1	0.540
mrc27	H1c1	1.587	3	0.529	0.766	1	0.766
mrc28	H1c1	1.272	3	0.424			
mrc29	H1c1	0.821	2	0.411			
mrc30	H1ad	0.821	2	0.411			
mrc31	K1c2	2.716	7	0.388	0.609	1	0.609
mrc32	H8c2	1.318	3	0.439			
mrc33	H7	1.943	4	0.486	0.625	1	0.625
mrc34	H6a1a3a	1.227	3	0.409			
mrc35	H6a1a	1.62	4	0.405			
mrc36	H6a1b2b	1.393	3	0.464	0.572	1	0.572
mrc37	H	1.431	3	0.477	0.610	1	0.610
mrc38	H3	1.275	3	0.425	0.682	1	0.682
mrc39	H17	0.821	2	0.411			
mrc40	H2a2	1.43	3	0.477	0.609	1	
mrc41	H3q1	0.821	2	0.411			
mrc42	H2a1d	0.821	2	0.411			
mrc43	U4b1b1	2.22	6	0.370	0.625	1	0.625
mrc44	K1c2	2.716	7	0.388	0.609	1	0.609
mrc45	K1a4a1a2b	2.966	8	0.371	1.114	2	0.557
mrc46	K1	2.921	7	0.417	1.273	2	0.637
mrc47	K2a5a1	3.014	7	0.431	0.609	1	0.609
mrc48	K2b1	2.539	6	0.423	0.609	1	0.609
mrc49	U3a1a	2.254	6	0.376			
mrc50	U5b2b	3.157	8	0.395	1.769	3	0.590
mrc51	U5b2b	3.548	9	0.394	1.169	2	0.585
mrc52	U5b2c2	1.845	5	0.369	0.610	1	0.610
mrc53	U5b2a4a	2.86	7	0.409	1.333	2	0.667
mrc54	U5a1a1	1.601	5	0.320			
mrc55	U5a1a1	1.601	5	0.320			
mrc56	U5a1a1	1.601	5	0.320			
mrc57	U1a1	1.046	3	0.349	0.512	1	0.512
mrc58	H47	0.821	2	0.411			
mrc59	H26a	0.821	2	0.411			
mrc60	H	0.821	2	0.411			
mrc61	U4b	1.215	4	0.304			
mrc62	U4a	1.215	4	0.304			
mrc63	K1a4a1a	2.461	7	0.352	0.609	1	0.609
mrc64	K2a4	2.548	6	0.425	0.609	1	0.609
mrc65	U2e2a3	2.052	5	0.410	1.066	2	0.533
mrc66	U3a1	1.793	5	0.359			
mrc67	H1b1	1.361	3	0.454	0.540	1	0.540
mrc68	H5a	0.821	2	0.411			
mrc69	H	0.821	2	0.411			

US AD	Haplogroup	Total load	No. varinats	Adjusted load	> 0.5 Total load	No. varinats	> 0.5Adjusted load
mk1	N1a1a1a2	2.529	7	0.361	1.122	2	0.561
mk2	I2a1	3.012	7	0.430	1.357	2	0.679
mk3	X2k	2.112	5	0.422	0.741	1	0.741
mk4	W1c	2.236	6	0.373	0.505	1	0.505
mk5	J1c2r	3.925	10	0.393	2.262	4	0.566
mk6	J1c9	2.883	8	0.360	1.220	2	0.610
mk7	J1; J1c12	4.047	10	0.405	2.014	3	0.671
mk8	J1c	2.883	8	0.360	1.220	2	0.610
mk9	J1c3g	3.329	9	0.370	1.220	2	0.610
mk10	J1c1b1a	4.288	10	0.429	2.625	4	0.656
mk11	J1d1a	2.274	7	0.325	0.611	1	0.611
mk12	T1a1	2.323	6	0.387	1.239	2	0.620
mk13	T1a2	2.572	7	0.367	1.239	2	0.620
mk14	T2b	2.987	7	0.427	1.903	3	0.634
mk15	T2c1d	2.895	7	0.414	1.811	3	0.604
mk16	T2f1a1	3.518	8	0.440	2.434	4	0.609
mk17	H1ag1	0.821	2	0.411			
mk18	H1e	2.041	4	0.510	1.220	2	0.610
mk19	H1ay	2.15	4	0.538	1.329	2	0.665
mk20	H	2.145	4	0.536	1.324	2	0.662
mk21	H1a5	1.902	4	0.476	0.706	1	0.706
mk22	H1a3a	0.996	3	0.332			
mk23	H1g1	1.435	3	0.478	0.614	1	0.614
mk24	H1g1	1.435	3	0.478	0.614	1	0.614
mk25	H1g1	1.435	3	0.478	0.614	1	0.614
mk26	H1	0.821	2	0.411			
mk27	H1c6	1.59	3	0.530	0.769	1	0.769
mk28	H1	0.821	2	0.411			
mk29	H5a	1.234	3	0.411			
mk30	H5a1c1a	1.313	3	0.438			
mk31	H5a1	1.448	3	0.483	0.627	1	0.627
mk32	H7a1b	1.436	3	0.479	0.615	1	0.615
mk33	H1c1	0.821	2	0.411			
mk34	H1ae1	0.821	2	0.411			
mk35	J1c	1.822	6	0.304	0.611	1	0.611
mk36	H1a6	0.821	2	0.411			
mk37	H7b2a	1.036	3	0.345			
mk38	H6a1a1a	1.732	4	0.433	0.505	1	0.505
mk39	H6a1b2	0.821	2	0.411			
mk40	H4a1a1	2.263	5	0.453	0.664	1	0.664
mk41	H4a1a1	2.263	5	0.453	0.664	1	0.664
mk42	H4a1a3a	2.591	6	0.432	0.664	1	0.664
mk43	H3k1a	0.821	2	0.411			
mk44	H11a1	1.581	4	0.395	0.509	1	0.509
mk45	H13a1	0.821	2	0.411			
mk46	U4b1b	1.388	4	0.347	0.625	1	0.625
mk47	K1a1b1b	1.648	5	0.330			
mk48	K1	2.257	6	0.376	0.609	1	0.609
mk49	K1a16	2.689	7	0.384	0.609	1	0.609
mk50	H6a1b	0.821	2	0.411			
mk51	U5b2a	2.137	6	0.356	0.610	1	0.610
mk52	U5a1b	2.062	6	0.344			
mk53	U5a2b3	1.925	5	0.385	0.690	1	0.690
mk54	U6a3	1.688	5	0.338	0.527	1	0.527
mk55	H49a	0.821	2	0.411			
mk56	H7b	0.821	2	0.411			
mk57	H16b	0.821	2	0.411			
mk58	H3aa	0.452	1	0.452			
mk59	H7b	0.821	2	0.411			
mk60	H24a	0.821	2	0.411			
mk61	K1a1b1a	2.257	6	0.376	0.609	1	0.609
mk62	K2a2a1	2.548	6	0.425	0.609	1	0.609
mk63	K1b2b	2.454	7	0.351	0.609	1	0.609
mk64	U5b1c2b	1.235	4	0.309			
mk65	U5a1b	2.062	6	0.344			
mk66	HV	0.821	2	0.411			
mk67	K1	2.257	6	0.376	0.609	1	0.609
mk69	U3a1	1.792	6	0.299			
mk70	H1a1	1.276	3	0.425			
mk71	T2b	2.158	5	0.432	1.239	2	0.620
mk72	H1e6	0.821	2	0.411			
mk73	H2	0.821	2	0.411			
mk74	H1	0.821	2	0.411			
mk75	H1	0.821	2	0.411			