ELSEVIER

Alzheimer's & Dementia ■ (2017) 1-6

Alzheimer's & Dementia

Short Report

Assembly of 809 whole mitochondrial genomes with clinical, imaging, and fluid biomarker phenotyping

Perry G. Ridge[a], Mark E. Wadsworth[a], Justin B. Miller[a], Andrew J. Saykin[b], Robert C. Green[c], the Alzheimer's Disease Neuroimaging Initiative[1], John S. K. Kauwe[a,d,*]

[a]*Department of Biology, Brigham Young University, Provo, UT, USA*
[b]*Radiology and Imaging Sciences, Medical and Molecular Genetics and the Indiana Alzheimer's Disease Center, Indiana University School of Medicine, Indianapolis, IN, USA*
[c]*Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Partners HealthCare Personalized Medicine, The Broad Institute and Harvard Medical School, Boston, MA, USA*
[d]*Department of Neuroscience, Brigham Young University, Provo, UT, USA*

Abstract

Introduction: Mitochondrial genetics are an important but largely neglected area of research in Alzheimer's disease. A major impediment is the lack of data sets.
Methods: We used an innovative, rigorous approach, combining several existing tools with our own, to accurately assemble and call variants in 809 whole mitochondrial genomes.
Results: To help address this impediment, we prepared a data set that consists of 809 complete and annotated mitochondrial genomes with samples from the Alzheimer's Disease Neuroimaging Initiative. These whole mitochondrial genomes include rich phenotyping, such as clinical, fluid biomarker, and imaging data, all of which is available through the Alzheimer's Disease Neuroimaging Initiative website. Genomes are cleaned, annotated, and prepared for analysis.
Discussion: These data provide an important resource for investigating the impact of mitochondrial genetic variation on risk for Alzheimer's disease and other phenotypes that have been measured in the Alzheimer's Disease Neuroimaging Initiative samples.

Keywords: Alzheimer's disease; ADNI; Mitochondrial genetics; Whole mitochondrial genomes; Next-generation sequencing

1. Introduction

Alzheimer's disease (AD), the most common form of dementia, affects >20 million people worldwide and is the only one of the top 10 causes of death that has no effective treatments [1–3]. Full-time care is required as AD progresses, further impacting patients and their loved ones and stressing the health-care system. With incidence expected to increase to 1 in 85 people by 2050 [2], it is essential to achieve early diagnosis, effective treatments, and a better understanding of the underlying etiology.

Understanding the underlying mechanisms of risk for AD is a key for both diagnosis and treatment. Swerdlow et al. [4] proposed the Mitochondrial Cascade Hypothesis of AD. Briefly, an individual's genetics determine the baseline mitochondrial function and how mitochondria change as a person ages and declining mitochondrial function causes AD-specific pathologies.

In addition to the evidence provided by Swerdlow et al. [4], several lines of evidence support a role for mitochondria

[1]Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

*Corresponding author. Tel.: ■■■;Fax: ■■■.
E-mail address: kauwe@byu.edu

https://doi.org/10.1016/j.jalz.2017.11.013

in AD. First, mitochondria fundamentally change in a number of ways in AD and contribute to its progression and onset [5]: metabolism decreases [6], mitochondrial fusion/fission are disrupted [7], mitochondrial concentration (i.e. the ratio of mitochondrial genomes to nuclear genomes) decreases in cerebrospinal fluid [8,9], mitochondrial morphology changes [4,10], mitochondrial-encoded enzymes in the electron transport chain are altered [5,11], amyloid plaques aggregate in mitochondria [12,13], and many of these changes take place near plaques [14].

Second, individuals with a maternal family history of AD have as high as 9 times the risk of AD compared with individuals with a paternal family history of AD [15,16], or no family history. Furthermore, individuals with a maternal family history of AD also score lower on cognitive tests [17], have a lower age of onset of AD [15,18], and have more pronounced brain abnormalities consistent with AD (e.g. cerebral metabolic changes [19], higher amyloid β burden [20], reduction in gray matter volume [21,22], and increased global PiB uptake PiB-positron emission tomography [23]). Moreover, we found that some of these brain abnormalities are associated with mitochondrial haplotypes [24].

This mitochondrial impact on AD risk could be influenced by several factors, including differential responses to the oxidative stress, variation in nuclear-encoded mitochondrial genes, and variation in the mitochondrial genome. In this article, we focus on an important resource for investigating mitochondrial genomic variation and others [25]. Several groups have reported a relationship between mitochondrial genetics and risk for AD (summarized in Ridge et al. [3], Table 2). Twelve different haplotypes have been implicated in mitochondrial genetic studies, but the majority of these were reported only once and not replicated [26–33], and six different studies reported no association between mitochondrial genetic variants and AD [34–39]. Among reported associations, there is no consensus, and sometimes, results appear to be contradictory. For example, Haplogroup U has been reported as both a risk and protective haplogroup [28,31,32]. However, potentially explaining the confounding nature of discoveries to date, the majority of studies used incomplete sequence data and/or had very small sample sizes [26–39], thus most were underpowered and lacked the resolution to identify correlations for all but the most common haplogroups. Only a single study used whole mitochondrial data [30], whereas most genotyped only a handful of SNPs. Furthermore, only one study used a large data set, but in this particular data set, the authors only genotyped 138 SNPs [39]. In summary, there is strong evidence to suggest a relationship between the mitochondrial genome and AD, yet the relationship remains undefined.

The Alzheimer's Disease Neuroimaging Initiative (ADNI) recently sequenced the whole genomes, including mitochondrial genomes, of 809 individuals. Each of the genomes was analyzed using tools and pipelines developed for diploid genomes. However, these analysis pipelines, particularly variant identification that relies on a likelihood model expecting diploid sequences, are inaccurate for use on the mitochondrial genome, which is haploid. Here, we report not only an AD data set of 809 annotated whole mitochondrial genomes with extensive phenotypes (Table 1) but also an appropriate pipeline to analyze mitochondrial genomes. We hope to facilitate research in this important area by providing a data set and analysis pipeline for future researchers to augment this initial data set.

2. Methods

2.1. Alzheimer's Disease Neuroimaging Initiative

Data used in the preparation of this article were obtained from the ADNI database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public–private partnership and is an ongoing, longitudinal, highly collaborative study. The primary goal of ADNI has been to test whether serial magnetic resonance imaging, positron emission tomography, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment and used for the early diagnosis of AD. ADNI has undergone several phases (ADNI1, ADNI GO, and ADNI 2), with each phase adding additional samples. In 2012, 818 ADNI samples were selected for whole genome sequencing to further the goals of ADNI. DNA sequence data were collected from DNA derived from whole blood. All subjects in our analyses had self-reported ancestry of non-Hispanic European American. All the data (whole genome sequence, phenotype, and newly assembled and annotated whole mitochondrial genomes) are publically available through ADNI (http://adni.loni.usc.edu/data-samples/).

2.2. Genome sequencing, assembly, and variant detection

ADNI genomes were sequenced on an Illumina HiSeq. Reads were paired-end, 100 base-pair reads. Before read mapping, adapters were removed. ADNI mapped the whole genome sequences and called variants using default settings in the Burrows-Wheeler Aligner [40] for mapping and

Table 1
Demographics

	Count	Sex (male/female)	Average age	APOE status (22/23/33/34/44/24)
Cases	191	126/65	74.42	0/8/74/80/25/4
Controls	279	135/144	74.51	0/35/167/68/7/2
MCI	333	183/149*	71.57*	1/26/162/110/25/9
Total	803	444/358*	73.17*	1/69/403/258/57/15

APOE, apolipoprotein E; MCI, mild cognitive impairment.
Demographic and phenotype information is available for 803 of the 809 mitochondrial genomes in the data set. APOE status refers to APOE genotype.
*Missing data for one sample.