Mitochondrial tRNA methylation in Alzheimer's disease and progressive supranuclear palsy

Talisa K. Silzer
University of North Texas Health Science Center
Gita A. Pathak
Yale School of Medicine
Nicole Phillips (nicole.phillips@unthsc.edu)
University of North Texas Health Science Center

Research article

Keywords: Mitochondria, methylation, neurodegenerative disease, tRNA

Posted Date: March 10th, 2020

DOI: https://doi.org/10.21203/rs.2.20546/v2

Version of Record: A version of this preprint was published at BMC Medical Genomics on May 19th, 2020.
See the published version at https://doi.org/10.1186/s12920-020-0727-9.

Abstract

Background Methylation of mitochondrial tRNAs (mt-tRNA) at the 9 th position (“p9 site”) is known to impact translational efficiency and downstream mitochondrial function; however, direct assessment of mt-RNA methylation is challenging. Recent RNA sequence-based methods have been developed to reliably identify post-transcriptional methylation. Though p9 methylation has been studied in healthy human populations and in the context of cancer, it has not yet been analyzed in neurodegenerative disease, where mitochondrial dysfunction is a prominent and early hallmark of disease progression. Results Here we show variable methylation at 11 p9 sites in post-mortem cerebellar tissue of elderly subjects who were either healthy or diagnosed with Alzheimer’s disease (AD), progressive supranuclear palsy (PSP) or pathological aging (PA). Similarities in degree of methylation were observed between AD and PSP. Expression of 5300 genes was significantly correlated with p9 methylation, with AD and PSP subjects exhibiting similar expression profiles. The top 100 transcripts were found to be enriched for mitochondrial-related toxicity. Overrepresentation testing using the top 1000 transcripts revealed enrichment for a number of molecular processes and terms including some which were mitochondrial-related. Conclusion With mitochondrial dysfunction being an established hallmark of neurodegenerative disease pathophysiology, this work sheds light on the potential molecular underpinnings of this dysfunction. Here we show overlap in cerebellar pathophysiology between common tauopathies such as Alzheimer’s disease and progressive supranuclear palsy. Whether p9 hypermethylation is a cause or consequence of pathology remains an area of focus.

Background

Mitochondrial organelles play essential roles in oxidative phosphorylation, apoptosis, calcium homeostasis and a number of cell-signalling pathways. Many lines of evidence have implicated mitochondrial dysfunction in aging and neurodegenerative disease. Though the genetic etiology of neurodegeneration is highly heterogeneous, there is an extensive literature implicating mitochondrial dysfunction as a common and early factor in the pathophysiology of a number of neurodegenerative diseases (1). Much of the mitochondrial genetic-centric studies of neurodegenerative diseases and related comorbidities have focused primarily on mitochondrial DNA (mtDNA) mutations, changes to mtDNA copy number and/or release of cell-free mtDNA (2); however, the significance of mitochondrial transcriptome epigenetics and altered mitochondrial RNA (mt-RNA) processing has not been explored. As post-transcriptional processing can alter mitochondrial protein synthesis, aberrant modifications to mitochondrial tRNAs (mt-tRNA) may contribute to the mitochondrial dysfunction underlying neurodegeneration and other age-related pathologies.

Mitochondria contain their own 16,569 base pair genome consisting of 13 mRNA genes encoding electron transport chain subunits and 2 rRNA genes, which are ‘punctuated’ by 22 tRNA genes **(3)**. The interspersed tRNA sequences serve as important sites for post-transcriptional processing of the polycistronic transcript, accomplished in part by key modifications at several sites throughout the tRNA structure (4). One of the most common modification types is methylation of adenosine or guanosine

residues (4). Of particular interest is methylation occurring at the 9th position (termed as “p9 sites” here) of mt-tRNAs, as this modification is known to be essential for proper tRNA folding, stability, and decoding capacity (5-8). Within mt-tRNAs, methyl groups are added by mitochondrial ribonuclease P protein 1 (MRPP1) to adenosine to create 1-methyladenosine (m1A). Lack of modification at this position has been associated with rare neurodegenerative disorders such as HSD10 disease (9). Further, knockdown of key processing molecules such as MRPP1 has been shown to lead to accumulation of improperly processed tRNAs resulting in decreased mitochondrial protein synthesis (10).

Though various methods exist for measuring RNA methylation including in situ hybridization, methylation-specific antibodies and methylation-sensitive enzymes, many methods lack specificity and/or have limited experimental use, due to the high concentration of input RNA required (11). Several sequence-based protocols have been established and validated demonstrating a strong correlation between variation in mt-RNA sequence data and degree of post-transcriptional methylation of mt-tRNA at specific sites (10, 12, 13). In brief, the extent of multiallelic calls (seemingly heteroplasmy) within mt-RNA sequence reads can be used to infer degree of post-transcriptional methylation based on the assumption that chemical modifications such as methyl groups, block reverse transcriptase during cDNA synthesis, often leading to incorporation of incorrect nucleotides (14). Though useful for detecting methylation, it is important to note that this method likely leads to the underestimation of methylation rate.

Using RNA sequence-based approaches, previous groups have identified variable methylation across 11 different p9 sites within mt-tRNAs of healthy Caucasian individuals; this methylation was discovered to be significantly associated with nuclear gene variants (i.e. *MRPP3*) (15). In this study, we examined methylation at these same 11 mitochondrial p9 sites in the context of different neurodegenerative pathologies and identified associated nuclear genetic variants and gene expression. The results point to altered cellular processes that may be contributing to the mitochondrial dysfunction occurring within the cerebellum that is a hallmark of both Alzheimer’s disease (AD) and progressive supranuclear palsy (PSP).

Results

Mitochondrial tRNA methylation

Frequencies of alternative allele calls measured at 11 previously identified p9 sites (585, 1610, 4271, 5520, 7526, 8303, 9999, 10413, 12146, 12274, 14734) based on the revised Cambridge reference sequence, were used to infer degree of post-transcriptional methylation. High levels of mt-RNA sequence variation were observed across all individuals (Additional Table 1). Degree of methylation, inferred from multi-allelic calls in RNA-seq data, was highly correlated across all 11 p9 sites with exception of 4271 and 14734 (Additional Figure 2). For each p9 site, comparisons were made between each of the four groups: normal controls (NC) (n= 74), Alzheimer’s disease (AD) (n= 82), progressive supranuclear palsy (PSP) (n= 83) and pathological aging (PA) (n=27) to assess differences in methylation levels. Non-parametric (Kruskal-Wallis) pairwise comparisons by disease state revealed significant differences in methylation at a number of p9 sites, with AD and PSP showing similar degrees of hypermethylation in comparison to NC

and PA; no significant difference was seen between NC and PA (Table 1; Additional File 2). Site 585 was found to be the most variable, though the reasoning for this is unclear.

Table 1. Summary table detailing p9 methylation data distributions and significant group differences per site. Mean ± standard deviation of degree of methylation (inferred from alternative allele frequency) is shown across the top for each site. Non-parametric Kruskal-Wallis p-values with Bonferroni correction (to adjust for multiple comparisons) is shown on the far right. P-values from pair-wise Kruskal-Wallis comparisons between groups are shown within the matrix; p-values<0.05 (*), p-values<0.001 (**).

p9 site		NC	AD	PSP	PA	
585	Mean±SD	0.0140±0.0057	0.0210±0.01067	0.0201±0.0092	0.0142±0.0053	p-value (Bonferroni adj.)
	NC		<0.001	<0.001		<0.001
	AD	**			0.01	
	PSP	**			0.02	
	PA		*	*		
1610		0.1276±0.0456	0.1601±0.0579	0.1546±0.0599	0.1239±0.0339	0.001
	NC		0.003	0.045		
	AD	*			0.018	
	PSP	*			0.1	
	PA		*			
4271		0.1308±0.0247	0.1343±0.0234	0.1304±0.0229	0.1379±0.0239	0.602
	No significant group differences					
5520		0.0610±0.0393	0.0810±0.0409	0.0790±0.0496	0.0576±0.0250	0.002
	NC		0.004	0.072		
	AD	*			0.069	
	PSP				0.357	
	PA					
7526		0.0954±0.0468	0.1217±0.0492	0.1247±0.0685	0.0888±0.0353	0.001
	NC		0.007	0.059		
	AD	*			0.019	
	PSP				0.83	
	PA		*			
8303		0.0687±0.0419	0.0830±0.0455	0.0966±0.0572	0.0758±0.0346	0.01
	NC		0.17	0.006		
	AD					
	PSP	*				
	PA					
9999		0.0929±0.0527	0.1130±0.0546	0.1370±0.0808	0.0921±0.0416	0.001
	NC		0.072	0.001		
	AD				0.518	
	PSP	**			0.055	
	PA					
10413		0.0754±0.04986	0.1000±0.0514	0.1209±0.0857	0.0728±0.0389	<0.001
	NC		0.007	<0.001		
	AD	*			0.131	
	PSP	**			0.031	
	PA			*		
12146		0.0445±0.0278	0.0523±0.0289	0.0658±0.0512	0.0492±0.0232	0.063
	No significant group differences					
12274		0.0103±0.0069	0.0140±0.0086	0.0167±0.0151	0.0099±0.0059	0.004
	NC		0.038	0.024		
	AD	*			0.22	
	PSP	*			0.17	
	PA					
14734		0.0480±0.0211	0.0484±0.0480	0.0474±0.0226	0.0450±0.0178	0.933
	No significant group differences					

Nuclear gene-based associations

Gene-based GWAS mapped typed SNPs to 17079 protein coding genes. Two SNPs (rs12343928; rs4877837) located on chromosome 9 were mapped to *SLC28A3* (p-value = $1.935x10^{-6}$, CADD score = 24.7) and were determined to be significantly associated with methylation at site 585; the regional plot reveals that several SNPs in LD with the top lead SNP are also associated in the local signal (Figure 1; Additional File 3). A single SNP (rs2034879) mapping to *SENP8* (CADD score = 13.22), *MYO9A* (CADD score = 18.35), *GRAMD2* (CADD score = 16.49) on chromosome 15 had suggestive associations with methylation at p9 site 585 (Figure 1; Additional File 3). Another SNP (rs9872864) mapping to *TRAIP* (p-value = $8.32x10^{-7}$, CADD score = 7.558) and nearby gene *IP6K1* (p-value = $6.67x10^{-7}$, CADD score = 20.7) located on chromosome 3 was also found to have significant associations with methylation at six out of the 11 p9 sites (5520, 7526, 8303, 9999, 10413, 12146) (Figure 2; Additional File 3).

Mitochondrial-related gene expression

Expression of 5300 transcripts was found to significantly correlate with degree of methylation at site 585; the top 1000 transcripts (565 unique genes) are visualized in Figure 3. Based on the primary bifurcation, left and right clades appear to be enriched for AD and PSP, and NC and PA respectively; though enrichment was not formally tested (Figure 3).

The top 1000 transcripts correlated with p9 methylation at site 585 were analyzed using PANTHER to test for significantly overrepresented cellular components, molecular functions and biological processes. We identified a number of GO terms to be overrepresented in our gene-set including several that were mitochondrially-associated (Figure 4; Additional File 4).

Further, using IPA software, molecular toxicity analysis of the top 100 transcripts correlated with methylation at site 585 yielded oxidative stress (*p*-value= 0.0115) and gene regulation by peroxisome proliferators via PPARa (*p*-value= 0.03) to be the most significantly enriched processes (Additional Figure 3).

Discussion

Our analysis of RNA sequence has validated that methylation occurs consistently at functionally important sites across the mitochondrial transcriptome, and that this methylation is highly correlated across the 11 multiallelic sites. Here we show significant hypermethylation at specific sites in the mitochondrial transcriptome (i.e. mt-tRNA p9 sites) in the cerebellar tissue of individuals with Alzheimer's disease and progressive supranuclear palsy. Similarities in methylation between these two pathologies may suggest a comparable genetic etiology underlying the cerebellar mitochondrial dysfunction that is a hallmark of both diseases (16),(17). Interestingly, no significant difference was observed between NC and PA, which may be due to the ambiguity of differentiating healthy from pathological aging. Individuals diagnosed with pathological aging often appear to be cognitively intact despite the presence of neuritic plaques. Similar presentation has also been observed in healthy aged individuals, making it difficult to distinguish between the two diagnoses (18). Moreover, a recent study by Idaghdour and Hodgkinson

(2017) revealed significant hypermethylation at mitochondrial p9 sites in cancer. This substantiates the notion that mt-tRNA hypermethylation may be more closely associated with pathology rather than normal aging processes. Whether this hypermethylation is a cause or consequence of pathology is still unclear.

Nuclear genes are associated with p9 methylation

The gene-based genome-wide association analysis revealed that variation in several nuclear encoded genes is significantly associated with methylation at different p9 sites throughout the mitochondrial transcriptome. Two SNPs (rs12343928; rs4877837) mapping to solute carrier family 28 member 3 (*SLC28A3*) located on chromosome 9 were found to be significantly associated with methylation at p9 site 585. The strength of this result was further substantiated by the clustering of neighbouring SNPs in linkage disequilibrium within the gene region (Figure 1). The protein encoded by this gene is involved in regulating neurotransmission, vascular tone and metabolism of nucleoside drugs, which are processes that are physiologically relevant to our tissue and diseases of interest. Nucleoside drugs are often used to treat age-related diseases such as type 2 diabetes and hypertension, as well as mitochondrial dysfunction in Parkinson's disease (PD) (19). This is interesting, as PSP pathology also features mitochondrial dysfunction and bears phenotypic similarity to PD, while T2D and hypertension often serve as comorbidities for AD.

A single SNP (rs2034879) that mapped to several genes (*SENP8, MYO9A, GRAMD2*) on chromosome 15 (likely due to the restrictive window size of ±10kb that was used) reached the suggestive threshold for association with p9 methylation at position 585. SENP8 is known to catalyze pathways associated with neddylation, a post-transcriptional modification analogous to ubiquitination; these pathways are also responsible for the endocytic degradation of APP. Dysfunction of these pathways has been observed in AD leading to characteristic accumulation of APP and Ab (20). MYO9A is an unconventional myosin responsible for the regulation of Rho GTPase activity within neurons, which in turn regulates neuronal morphology and function; for these reasons, the Rho family has been a recent therapeutic target for neurodegenerative diseases such as AD (21). In turn, various Rho GTPases have also been implicated in maintenance of mitochondrial homeostasis and apoptotic signaling (22). GRAMD2 is involved in the organization of endoplasmic reticulum-plasma membrane contact sites (EPCS), which are key modulators of calcium homeostasis (23). Calcium serves as an important regulator of mitochondrial bioenergetics (e.g. activation of different Kreb's cycle enzymes (24)), dynamics and apoptotic signaling. Subsequent contact sites between the ER and mitochondria, termed mitochondrial associated membranes (MAMs), facilitate uptake of calcium from the ER to the mitochondria (25), although there is not a clear role for GRAMD2 in this process.

Methylation at six of the 11 p9 sites (5520, 7526, 8303, 9999, 10413, 12146) was significantly associated with a single SNP (rs9872864) mapping to two genes *TRAIP* and *IP6K1* located on chromosome 3. *TRAIP* encodes TRAF interacting protein, an E3 ubiquitin ligase that plays a key role in cell survival and apoptosis. TRAIP serves important roles in cell cycle checkpoints by regulating spindle assembly, appropriate chromosome distribution during cell division, and DNA damage responses (26). Cell cycle

checkpoints have been proposed to serve neuroprotective roles; checkpoint dysregulation and cell cycle re-entry of post-mitotic neurons have been observed in a number of tauopathies (27). *IP6K1* encodes an inositol phosphokinase responsible for synthesis of 5-diphosphoinositol phentakisphosphate (5-IP_7) from hexakisphosphate (IP_6). IP6K1 serves as an important upstream regulator for various metabolic processes (e.g. glucose homeostasis, lipolysis), apoptosis and global transcription (28). It has been investigated as a therapeutic target for obesity and type 2 diabetes (29) and variants within the *IP6K1* gene region have been previously identified, using tag SNP analysis, as being associated with AD (30). IP6K1 also impacts mitochondrial function by regulating ATP concentration via alteration of the ratio of glycolytic to oxidative phosphorylation (31). Knockout (*IP6K1*-/-) yeast models have shown decreased mitochondrial respiration, yet increases in ATP, a paradox that is thought to be due to enhanced glycolysis and depletion of metabolic processes requiring ATP (31). Further, its role in lipolysis is pathologically relevant as findings from Wan et al. (32) demonstrate the capacity of Ab accumulation to promote lipolysis, increasing lipid toxicity and lipid peroxidation, leading to subsequent downstream mitochondrial dysfunction (33), a phenotype common to many neurodegenerative diseases.

While several of the genes associated here are loosely connected to mitochondrial function, we did not identify any loci that may directly affect mt-tRNA processing. It is interesting to note that 5 out of the 6 genes identified in the gene-based GWAS had CADD scores >10. In brief, Combined Annotation Dependent Depletion (CADD) scoring is a predictive machine learning tool that allows for estimation of the deleteriousness and/or pathogenicity of causal genetic variants (34). In general, raw CADD scores exceeding 10 are considered to be in the top 10% of deleterious variants in the human genome (34). Our findings here suggest that the genes identified as having significant associations with p9 methylation, may play a role in pathogenesis; though their exact role remains ambiguous at this time.

Hodkinson et al. (15), reported p9 methylation associations with several genes—one of which is *MRPP3*, a key player in tRNA processing in the mitochondria. We did not replicate the *MRPP3* association which is not entirely surprising since their study was conducted in normal adults (40-69 years) without any particular pathological conditions versus our study of aged subjects with neurodegenerative disease. It may follow that the variability in *MRPP3* associated with p9 methylation levels described in Hodgkinson et al., is presumably normal, whereas p9 methylation in our cohort is presumably a result of pathological process(es). The resulting genotype associations may be pointing to novel gene variants that are upstream effectors of mitochondrial function.

Mitochondrially-related gene expression is correlated with p9 methylation

We identified 5300 genes significantly correlated with p9 methylation. Using principle component analysis, we identified clustering of gene expression profiles by disease state with AD and PSP, and NC and PA grouping together respectively. Of the top gene expression hits, we identified over-represented GO terms related to a number of cellular components and processes using PANTHER. Interestingly, there was an enrichment for mitochondrial components. In addition, some of the GO terms such as *DNA repair* and *cellular response to DNA damage stimulus*, appeared to correlate with one of the top genes (i.e. *TRAIP*)

that was identified in the gene-based GWAS. Though some of the GO terms are rather ambiguous, it does imply that p9 methylation is associated with altered transcription of gene-sets important for mitochondrial function.

Further, of the top 100 transcripts correlated with p9 methylation we identified significant enrichment for toxicity processes including oxidative stress and gene regulation by peroxisome proliferators via PPARa. This is logical given that both oxidative stress and PPARa function have been implicated in aging and age-related pathologies such as diabetes, cardiovascular disease and neurodegeneration (35-37), and have obvious ties to mitochondrial function. PPARa is a transcription factor best recognized for its ability to regulate fatty acid uptake and oxidation and has been implicated in inflammatory processes (38, 39). It is capable of mediating expression of PGC1a, a key regulator of mitochondrial biogenesis and antioxidant defense. Because of this, PPARa is capable of functioning as a neuroprotective factor against oxidative stress through regulation of several mitochondrial proteins, including those involved in processes of mitochondrial dynamics (i.e. fission, fusion) (40) and has also been identified as a therapeutic target for neuroinflammation (41).

Conclusion

This is the first work analyzing mitochondrial p9 methylation in the context of neurodegeneration. Here we report an association between p9 methylation and nuclear-encoded gene expression, as well as mitochondrial-related regulation; although, the implications of these associations are unclear. We observed comparable post-transcriptional mt-tRNA hypermethylation and nuclear gene expression profiles in the cerebellum of Alzheimer's disease and progressive supranuclear palsy patients. This points to molecular similarities underlying the mitochondrial dysfunction already known to occur within the cerebellum of individuals diagnosed with either of these tauopathies.

A basal level of p9 methylation is required for proper tRNA folding and stability (7) and the degree of methylation at p9 sites has been shown by other groups to have low inter-individual variability (42). In addition, p9 hypermethylation has been observed in other age-related diseases such as cancer (43). We propose that p9 hypermethylation in AD and PSP may be more closely associated with pathology as opposed to normal aging. As shown in Figure 5, we observed that specific nuclear encoded variants increase risk for hypermethylation at mitochondrial p9 sites within cerebellar neurons, but importantly, causality is still not clear. For example, hypermethylation of p9 may be the primary event that impacts the function of the electron transport chain (ETC) and results in mitochondrial dysfunction, due to the influence that p9 methylation has on downstream mitochondrial protein translation. Alternatively, it is possible that ETC dysfunction is the primary event that causes altered p9 methylation, in which case, the p9 hypermethylation is serving as an endophenotype of genetic risk for mitochondrial dysfunction. Regardless, through retrograde signaling, mitochondrial dysfunction may impact nuclear gene expression resulting in alterations in cellular and molecular processes to further exacerbate neurodegenerative pathology.

Though at present, it is unclear whether mt-tRNA methylation is a cause or consequence of pathology, these results point to a potential role for mitochondrial post-transcriptional methylation in the pathophysiology of common tauopathies such as Alzheimer's disease and progressive supranuclear palsy. Given this, post-transcriptional mt-tRNA methylation may serve as an exciting area of investigation for the development of future treatment strategies for neurodegenerative disease.

Methods

Samples used

Samples used for this study were post-mortem cerebellar tissue of 266 elderly Caucasian individuals who were either healthy or diagnosed with Alzheimer's disease (AD), progressive supranuclear palsy (PSP) or pathological aging (PA) (Table). Samples were originally collected by the Mayo RNAseq study led by Dr. Nilüfer Ertekin-Taner, Mayo Clinic, Jacksonville, FL as part of the multi-PI U01 AG046139 (MPIs Golde, Ertekin-Taner, Younkin, Price). All individuals with AD "…had definite diagnosis according to the NINCDS-ADRDA criteria and had Braak NFT stage of IV or greater. Control subjects had Braak NFT stage of III or less, CERAD neuritic and cortical plaque densities of 0 (none) or 1 (sparse) and lacked any of the following pathologic diagnoses: AD, Parkinson's disease (PD), DLB, VaD, PSP, motor neuron disease (MND), CBD, Pick's disease (PiD), Huntington's disease (HD), FTLD, hippocampal sclerosis (HipScl) or dementia lacking distinctive histology (DLDH)" (doi:10.7303/syn5550404).

Table 2. Demographics of subjects used.

Diagnosis	Sex	Age at Death	APOE		
			e2	e3	e4
NC (n = 74)	Male (n = 40)	80.93±8.15	7.43%	86.40%	6.08%
	Female (n = 34)	85.15±6.76			
AD (n = 82)	Male (n = 34)	81.68±8.17	2.44%	68.30%	29.30%
	Female (n = 48)	83.21±7.19			
PSP (n = 83)	Male (n = 51)	74.22±6.94	9.04%	82.53%	8.43%
	Female (n = 32)	73.65±5.99			
PA (n = 27)	Male (n = 11)	85.18±3.74	7.41%	75.92%	16.67%
	Female (n = 16)	84.06±4.65			

Age is reported as an average ± standard deviation for each respective group. Normal control (NC), Alzheimer's disease (AD), progressive supranuclear palsy (PSP), pathological aging (PA).

Data obtainment and QC

SNP, RNA-seq data and gene expression data were obtained from the Synapse Data Repository (doi:10.7303/syn5550404). Genomic and RNA-seq data was originally generated and QC'd by the Golde, Ertekin-Taner, Younkin and Price laboratories. Briefly, genotypes were obtained using the Illumina Human Omni 2.5Exome array. The protocol used for QC of genomic data included the following procedures: (1)

duplicate or relatedness, (2) discrepant sex information, (3) low genotyping call rate (98%) and (4) outlying heterozygosity (± 3 standard deviations from the mean) were removed using PLINK. EIGENSTRAT was used to identify outliers (>6 standard deviations from the mean after 5 iterations) (44). Based on this QC protocol, a total of 10 individuals were removed from the original dataset of 278 individuals.

Briefly, RNA was extracted from brain tissue using TRIzol® and Qiagen RNeasy columns treated with DNase. Sequencing libraries were prepared using the Illumina Truseq kit. An Illumina Hiseq 2000 was used to obtain 101bp paired-end reads (performed in triplicate; 3 lanes/sample). RNA-seq data of individuals with low mapped reads (<85%) or sex discrepant gene counts (Y-chromosome gene expression) were removed from further analysis (20 individuals total). Using SNAPR, raw RNA-seq reads were aligned using the human genome and transcriptome builds GRCh38 and GRCh38.77 respectively, while filtering on Phred scores <= 20 (45). Subsequent BAM files were utilized for determining alternative allele frequency at the 11 p9 sites. Gene expression was derived from transcript counts generated using SNAPR (45); counts were normalized using the trimmed mean of m-values (TMM) method in edgeR taking into account differences in library size (calcNormFactors).

Mitochondrial RNA heteroplasmy

Mitochondrial specific reads were extracted from total RNA-seq BAM files using samtools v.1.34 *view.* An mpileup file was created from each BAM file using samtools *mpileup*. Heteroplasmy was measured at 11 different p9 sites in the mitochondrial transcriptome (585, 1610, 4271, 5520, 7526, 8303, 9999, 10413, 12146, 12274, 14734) from the mpileup files using *readcounts* in VarScan v.2.4 with default parameters. Of note, none of the aforementioned p9 sites have previously been shown to overlap with mitochondrial pseudogene (NUMTs) sequences (15). Frequency of alternative allele calls (non-reference calls based on Cambridge Reference Sequence (CRS)) was measured at each of the 11 sites. Degree of post-transcriptional methylation was inferred from the frequency of alternative allele calls based on previous published methods (15). A Kolmogorov-Smirnov test determined the data to have a non-normal distribution. Correlation coefficients for p9 methylation across all 11 sites were determined using Spearman's rho. Degree of p9 methylation was analyzed between disease states using pairwise Kruskal-Wallis tests with Bonferroni correction for multiple testing at each individual site (not accounting for the 11 site tests).

Gene-based genome-wide association

Principle component analysis using EIGENSOFT was conducted to determine if a considerable amount of heterogeneity was present in the SNP data (46). Linkage disequilibrium-based pruning of the SNPs was conducted in PLINK (–indep-pairwise 50 5 0.2) yielding a final dataset of 58,174 SNPs in linkage equilibrium (47). Linear association adjusting for age, sex and eigenvectors 1-10 was applied to identify genetic variants associated with degree of methylation (log10 transformed) at each respective p9 site using the SNP2GENE function of the Functional Mapping and Annotation (FUMA) platform (48); this

utilizes MAGMA to map SNPs to protein-coding genes (49). Position map window size was conservatively set to 10 kb. The genome-wide significance threshold was set to a Bonferroni adjusted p-value of $2.93x10^{-6}$ based on the mapping of input SNPs to 17079 genes; Bonferroni correction was applied within each linear association test. Manhattan plots and local regional plots as well as Q-Q plots to identify genomic inflation in the SNP-set used for the gene-based GWAS, were generated in FUMA.

Gene expression analysis

Gene expression significantly correlated with methylation at site 585 was determined using Spearman's rho. The top 1000 transcripts (565 unique genes) and the first two principle components were visualized using the ClustVis online tool (50). Rows and columns of the heatmap were clustered based on correlation distance and average linkage with unit variance scaling. The top genes (based on top 1000 transcripts) were input into the PANTHER classification system to identify overrepresented cellular components, molecular functions and/or biological processes using Fisher's exact test with FDR correction (46). Enrichment for particular toxicity processes based on the top 100 transcripts was also performed using the Ingenuity Pathway Analysis (IPA) (Qiagen Inc., https://www.qiagenbioinformatics.com/products/ingenuitypathway-analysis) Core Analysis Toxicity function.

Abbreviations

mtDNA

mt-RNA

mt-tRNA

NC

AD

PSP

PA

GWAS

SNP

Declarations

Ethics approval and consent to participate

Access to download the datasets used in this study was approved under the North Texas Regional IRB exempt protocol #2016-090.

Consent for publication

Not applicable.

Availability of data and materials

Data used in this study was obtained from the Synapse data repository (syn5550404). Subsequent analyses not included in the additional materials are available by the author upon request.

Competing interests

The authors declare that they have no competing interests

Funding

Not applicable.

Authors' contributions

NP and TS conceptualized and designed the study. Analysis and data interpretation were carried out by all authors. TS drafted the manuscript. All authors have read and approved the final submitted version of the manuscript.

Acknowledgements

Study data were provided by the following sources: The Mayo Clinic Alzheimer's Disease Genetic Studies, led by Dr. Nilufer Taner and Dr. Steven G. Younkin, Mayo Clinic, Jacksonville, FL using samples from the Mayo Clinic Study of Aging, the Mayo Clinic Alzheimer's Disease Research Center, and the Mayo Clinic Brain Bank. Data collection was supported through funding by NIA grants P50 AG016574, R01 AG032990, U01 AG046139, R01 AG018023, U01 AG006576, U01 AG006786, R01 AG025711, R01 AG017216, R01 AG003949, NINDS grant R01 NS080820, CurePSP Foundation, and support from Mayo Foundation. Study data includes samples collected through the Sun Health Research Institute Brain and Body Donation Program of Sun City, Arizona. The Brain and Body Donation Program is supported by the National Institute of Neurological Disorders and Stroke (U24 NS072026 National Brain and Tissue Resource for Parkinson's Disease and Related Disorders), the National Institute on Aging (P30 AG19610 Arizona Alzheimer's Disease Core Center), the Arizona Department of Health Services (contract 211002, Arizona Alzheimer's Research Center), the Arizona Biomedical Research Commission (contracts 4001, 0011, 05-901 and 1001 to the Arizona Parkinson's Disease Consortium) and the Michael J. Fox Foundation for Parkinson's Research.

References

1. Johri A, Beal MF. Mitochondrial Dysfunction in Neurodegenerative Diseases. Journal of Pharmacology and Experimental Therapeutics. 2012;342(3):619-30.
2. Swerdlow RH, Koppel S, Weidling I, Hayley C, Ji Y, Wilkins HM. Mitochondria, Cybrids, Aging, and Alzheimer's Disease. Prog Mol Biol Transl Sci. 2017;146:259-302.
3. Ojala D, Montoya J, Attardi G. tRNA punctuation model of RNA processing in human mitochondria. Nature. 1981;290(5806):470-4.
4. Powell CA, Nicholls TJ, Minczuk M. Nuclear-encoded factors involved in post-transcriptional processing and modification of mitochondrial tRNAs in human disease. Frontiers in Genetics. 2015;6(79).
5. Duechler M, Leszczyńska G, Sochacka E, Nawrot B. Nucleoside modifications in the regulation of gene expression: focus on tRNA. Cell Mol Life Sci. 2016;73(16):3075-95.
6. Voigts-Hoffmann F, Hengesbach M, Kobitski AY, van Aerschot A, Herdewijn P, Nienhaus GU, et al. A Methyl Group Controls Conformational Equilibrium in Human Mitochondrial tRNALys. Journal of the American Chemical Society. 2007;129(44):13382-3.
7. Helm M, Giegé R, Florentz C. A Watson–Crick Base-Pair-Disrupting Methyl Group (m1A9) Is Sufficient for Cloverleaf Folding of Human Mitochondrial tRNALys Biochemistry. 1999;38(40):13338-46.
8. Zhang C, Jia G. Reversible RNA Modification N(1)-methyladenosine (m(1)A) in mRNA and tRNA. Genomics Proteomics Bioinformatics. 2018;16(3):155-61.
9. Bohnsack MT, Sloan KE. The mitochondrial epitranscriptome: the roles of RNA modifications in mitochondrial translation and human disease. Cellular and Molecular Life Sciences. 2018;75(2):241-60.
10. Lopez Sanchez MIG, Mercer TR, Davies SMK, Shearwood A-MJ, Nygård KKA, Richman TR, et al. RNA processing in human mitochondria. Cell Cycle. 2011;10(17):2904-16.
11. Ovcharenko A, Rentmeister A. Emerging approaches for detection of methylation sites in RNA. Open Biology. 2018;8(9):180121.
12. Schwartz S, Motorin Y. Next-generation sequencing technologies for detection of modified nucleotides in RNAs. RNA biology. 2016;14(9):1124-37.
13. Mercer TR, Neph S, Dinger ME, Crawford J, Smith MA, Shearwood A-MJ, et al. The human mitochondrial transcriptome. Cell. 2011;146(4):645-58.
14. Hauenschild R, Tserovski L, Schmid K, Thüring K, Winz M-L, Sharma S, et al. The reverse transcription signature of N-1-methyladenosine in RNA-Seq is sequence dependent. Nucleic acids research. 2015;43(20):9950-64.
15. Hodgkinson A, Idaghdour Y, Gbeha E, Grenier J-C, Hip-Ki E, Bruat V, et al. High-Resolution Genomic Analysis of Human Mitochondrial RNA Sequence Variation. Science. 2014;344(6182):413-5.
16. Swerdlow RH. Mitochondria and Mitochondrial Cascades in Alzheimer's Disease. Journal of Alzheimer's disease : JAD. 2018;62(3):1403-16.

17. Albers DS, Swerdlow RH, Manfredi G, Gajewski C, Yang L, Parker WD, et al. Further Evidence for Mitochondrial Dysfunction in Progressive Supranuclear Palsy. Experimental Neurology. 2001;168(1):196-8.
18. Reas ET. Amyloid and Tau Pathology in Normal Cognitive Aging. The Journal of Neuroscience. 2017;37(32):7561.
19. Tufi R, Gandhi S, de Castro IP, Lehmann S, Angelova PR, Dinsdale D, et al. Enhancing nucleotide metabolism protects against mitochondrial dysfunction and neurodegeneration in a PINK1 model of Parkinson's disease. Nat Cell Biol. 2014;16(2):157-66.
20. Chen Y, Neve RL, Liu H. Neddylation dysfunction in Alzheimer's disease. J Cell Mol Med. 2012;16(11):2583-91.
21. Aguilar BJ, Zhu Y, Lu Q. Rho GTPases as therapeutic targets in Alzheimer's disease. Alzheimer's research & therapy. 2017;9(1):97-.
22. Fransson A, Ruusala, A, Aspenström, P. Atypical Rho GTPases have roles in mitochondrial homeostasis and apoptosis. J Biol Chem. 2003;278(8).
23. Saheki Y, De Camilli P. Endoplasmic Reticulum–Plasma Membrane Contact Sites. Annual Review of Biochemistry. 2017;86(1):659-84.
24. Traaseth N, Elfering S, Solien J, Haynes V, Giulivi C. Role of calcium signaling in the activation of mitochondrial nitric oxide synthase and citric acid cycle. Biochimica et Biophysica Acta (BBA) - Bioenergetics. 2004;1658(1):64-71.
25. Burgoyne T, Patel S, Eden ER. Calcium signaling at ER membrane contact sites. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research. 2015;1853(9):2012-7.
26. Chapard C, Meraldi P, Gleich T, Bachmann D, Hohl D, Huber M. TRAIP is a regulator of the spindle assembly checkpoint. Journal of Cell Science. 2014;127(24):5149.
27. Khurana V, Merlo P, DuBoff B, Fulga TA, Sharp KA, Campbell SD, et al. A neuroprotective role for the DNA damage checkpoint in tauopathy. Aging Cell. 2012;11(2):360-2.
28. Ghoshal S, Tyagi R, Zhu Q, Chakraborty A. Inositol hexakisphosphate kinase-1 interacts with perilipin1 to modulate lipolysis. Int J Biochem Cell Biol. 2016;78:149-55.
29. Zhu Q, Ghoshal S, Rodrigues A, Gao S, Asterian A, Kamenecka TM, et al. Adipocyte-specific deletion of Ip6k1 reduces diet-induced obesity by enhancing AMPK-mediated thermogenesis. The Journal of clinical investigation. 2016;126(11):4273-88.
30. Crocco P, Saiardi A, Wilson MS, Maletta R, Bruni AC, Passarino G, et al. Contribution of polymorphic variation of inositol hexakisphosphate kinase 3 (IP6K3) gene promoter to the susceptibility to late onset Alzheimer's disease. Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease. 2016;1862(9):1766-73.
31. Szijgyarto Z, Garedew A, Azevedo C, Saiardi A. Influence of Inositol Pyrophosphates on Cellular Energy Dynamics. Science. 2011;334(6057):802.

32. Wan Z, Mah D, Simtchouk S, Kluftinger A, Little JP. Role of amyloid β in the induction of lipolysis and secretion of adipokines from human adipose tissue. Adipocyte. 2014;4(3):212-6.
33. Schrauwen P, Schrauwen-Hinderling V, Hoeks J, Hesselink MKC. Mitochondrial dysfunction and lipotoxicity. Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids. 2010;1801(3):266-71.
34. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Research. 2018;47(D1):D886-D94.
35. Liguori I, Russo G, Curcio F, Bulli G, Aran L, Della-Morte D, et al. Oxidative stress, aging, and diseases. Clinical interventions in aging. 2018;13:757-72.
36. Fruchart J-C. Peroxisome proliferator-activated receptor-alpha (PPARα): At the crossroads of obesity, diabetes and cardiovascular disease. Atherosclerosis. 2009;205(1):1-8.
37. Agarwal S, Yadav A, Chaturvedi RK. Peroxisome proliferator-activated receptors (PPARs) as therapeutic target in neurodegenerative disorders. Biochemical and Biophysical Research Communications. 2017;483(4):1166-77.
38. Fan W, Evans R. PPARs and ERRs: molecular mediators of mitochondrial metabolism. Current Opinion in Cell Biology. 2015;33:49-54.
39. Corona JC, Duchen MR. PPARγ as a therapeutic target to rescue mitochondrial function in neurological disease. Free Radic Biol Med. 2016;100:153-63.
40. Zolezzi JM, Silva-Alvarez C, Ordenes D, Godoy JA, Carvajal FJ, Santos MJ, et al. Peroxisome proliferator-activated receptor (PPAR) γ and PPARα agonists modulate mitochondrial fusion-fission dynamics: relevance to reactive oxygen species (ROS)-related neurodegenerative disorders? PloS one. 2013;8(5):e64019-e.
41. Esmaeili MA, Yadav S, Gupta RK, Waggoner GR, Deloach A, Calingasan NY, et al. Preferential PPAR-α activation reduces neuroinflammation, and blocks neurodegeneration in vivo. Human molecular genetics. 2016;25(2):317-27.
42. Ali AT, Idaghdour Y, Hodgkinson A. Nuclear genetic regulation of human mitochondrial RNA modification. bioRxiv. 2019:666339.
43. Idaghdour Y, Hodgkinson A. Integrated genomic analysis of mitochondrial RNA processing in human cancers. Genome medicine. 2017;9(1):36-.
44. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nature Genetics. 2006;38:904.
45. Magis AT, Funk CC, Price ND. SNAPR: a bioinformatics pipeline for efficient and accurate RNA-seq alignment and analysis. IEEE Life Sci Lett. 2015;1(2):22-5.
46. Patterson N, Price AL, Reich D. Population Structure and Eigenanalysis. PLOS Genetics. 2006;2(12):e190.

47. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira Manuel A R, Bender D, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. American Journal of Human Genetics. 2007;81(3):559-75.
48. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nature Communications. 2017;8(1):1826.
49. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. PLOS Computational Biology. 2015;11(4):e1004219.
50. Metsalu T, Vilo J. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. Nucleic acids research. 2015;43(W1):W566-W70.

Additional File Legends

Additional File 1. Additional figures.

Additional File 2. Pairwise comparisons of p9 methylation across disease state.

Additional File 3. Gene-based GWAS results.

Additional File 4. PANTHER over-representation test results.

Figures

Figure 1

Schematic diagram displaying the 11 p9 sites analyzed and significant group differences in methylation. The outer circle represents the mitochondrial genome, with protruding tRNA illustrations signifying the 11 p9 sites analyzed. The inner coloured circles represent the different phenotypes (green = NC, purple = AD, blue = PSP, orange = PA). Significant differences in p9 methylation between the different groups are indicated by solid and dotted lines ($p<0.05$ and $p<0.001$) respectively, based on pairwise KW tests with Bonferroni correction.

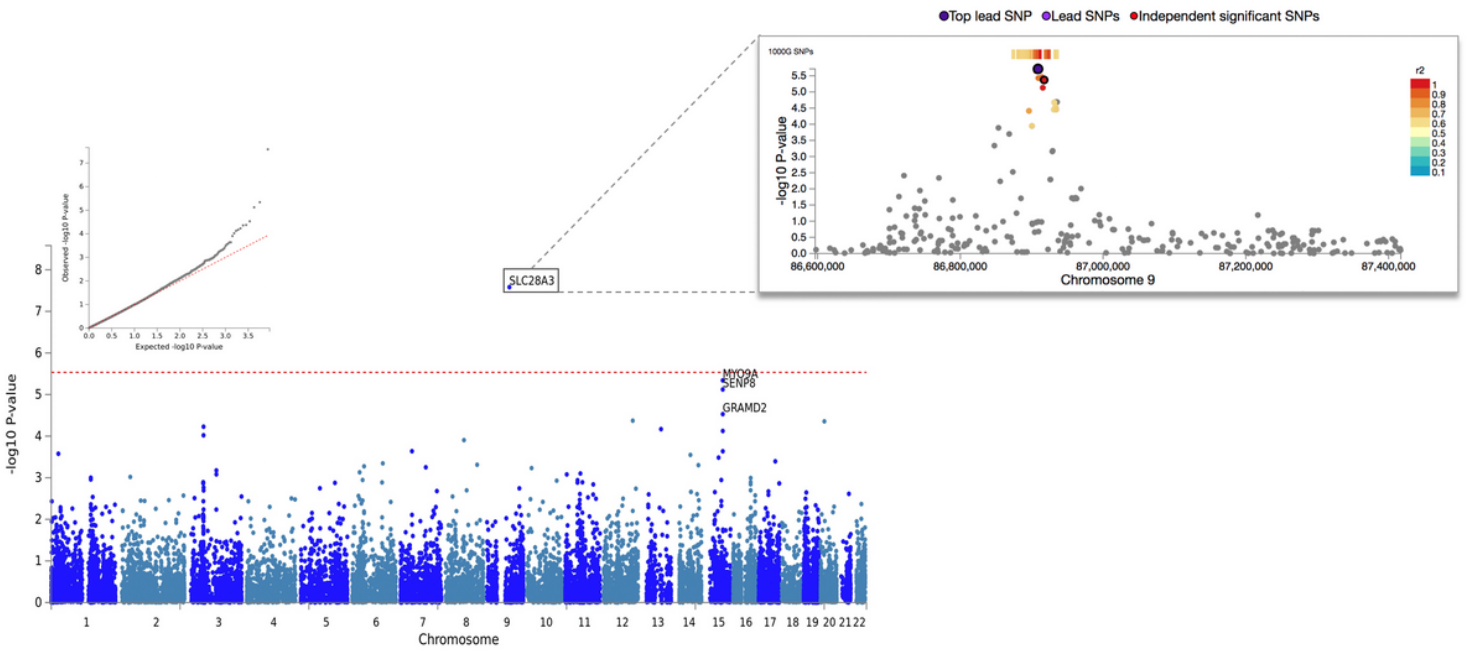

Figure 2

Association results for gene-based GWAS with methylation at site 585. The Q-Q plot (top left) conforms to expectations and shows no sign of genomic inflation. The Manhattan plot (bottom left) illustrates a primary signal in SLC28A3 on chromosome 9 (p-value = 1.935x10-6); as well as multiple suggestive associations on chromosome 15. The regional plot (top right) displays clustering of individual SNPs within an 800kb window and shows multiple SNPs in LD with the top lead SNP in the gene-based signal.

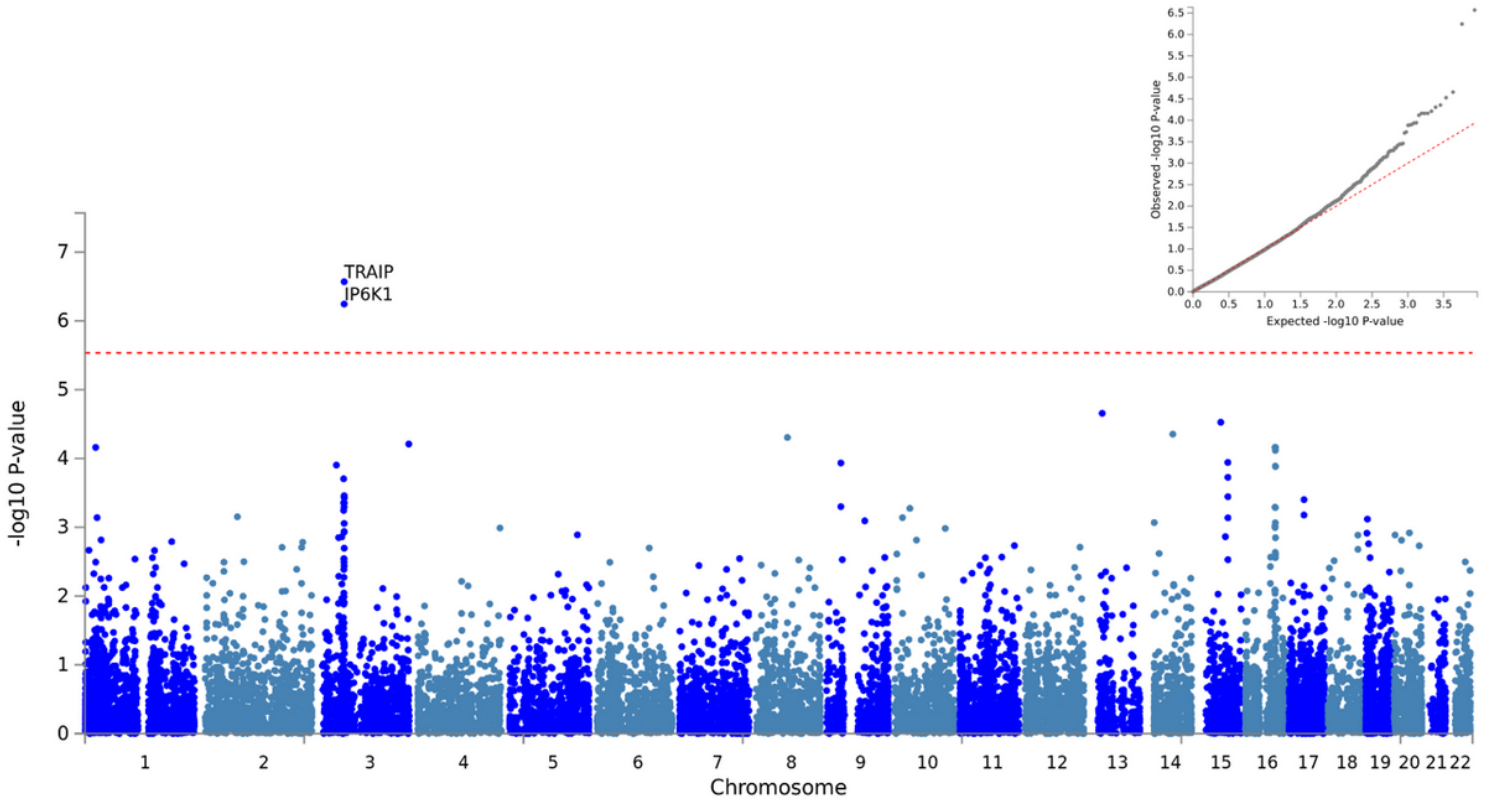

Figure 3

Association results for gene-based GWAS with methylation across 6 out of 11 sites. The Q-Q plot (top right) conforms to expectations and shows no sign of genomic inflation. The Manhattan plot (bottom left) illustrates a primary signal in TRAIP and IP6K1 on chromosome 3 (8.32x10-7, 6.67x10-7). This result was replicated in 6 out of 11 p9 sites (results for site 12146 shown here).

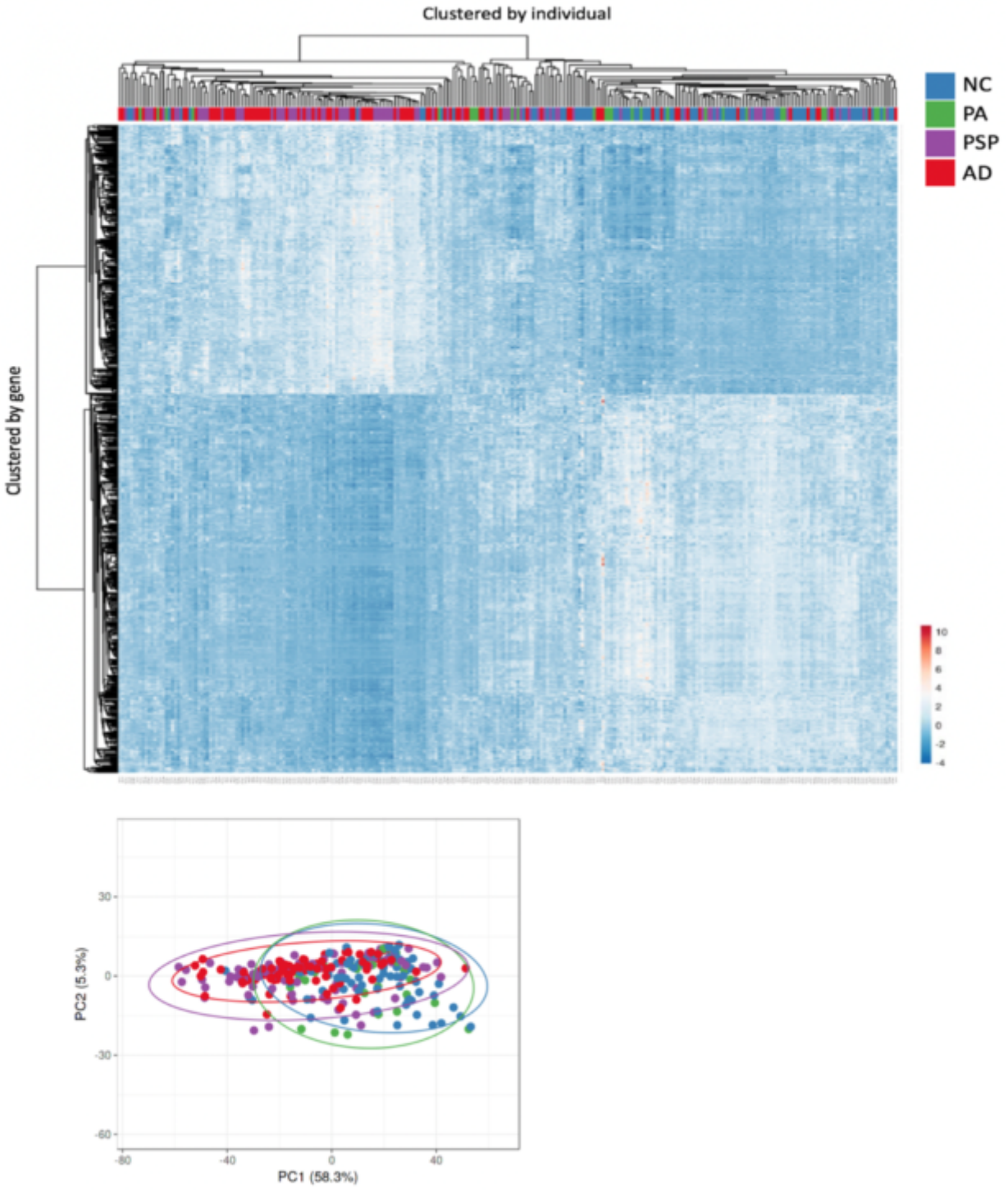

Figure 4

Heatmap displaying the top 1000 transcripts significantly associated with degree of p9 methylation at site 585. Expression profiles are clustered by gene (y-axis) and individual (x-axis) (top). Principle

component analysis of gene expression displays overlap between case groups (bottom left). Colors represent disease status; NC (blue), PA (green), AD (red) and PSP (purple) (bottom). Gradient bar (bottom right) represents gene expression level; high expression (red), low expression (blue).

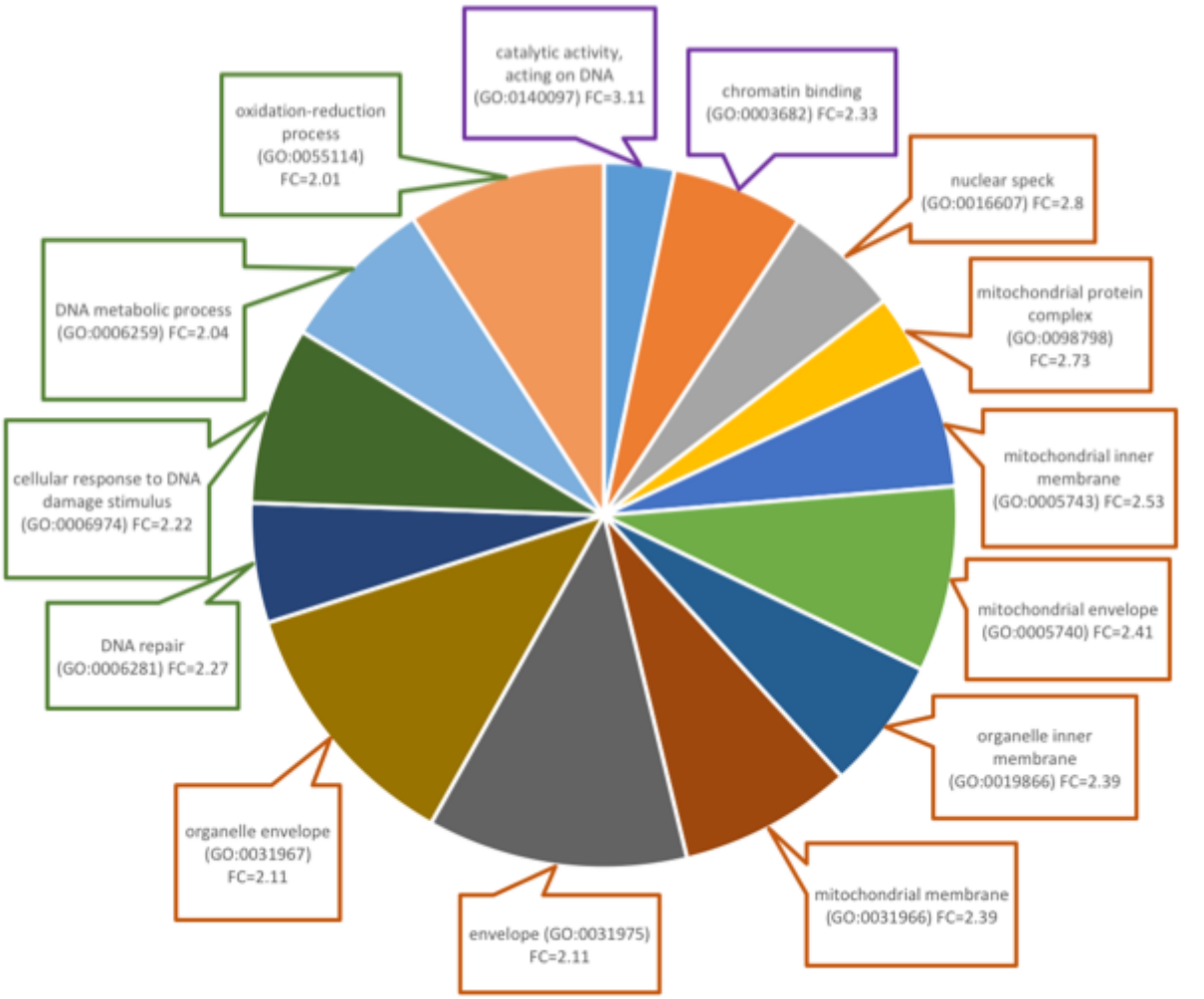

Figure 5

Overrepresented GO terms based on the top 1000 transcripts significantly correlated with p9 methylation. GO terms are color coded by category; Molecular function (purple), cellular component (orange), biological process (green). Terms were selected using a threshold of >2 fold change. Fold change was determined based on number of enriched genes from our 1000 transcripts list divided by the number of expected enriched genes.

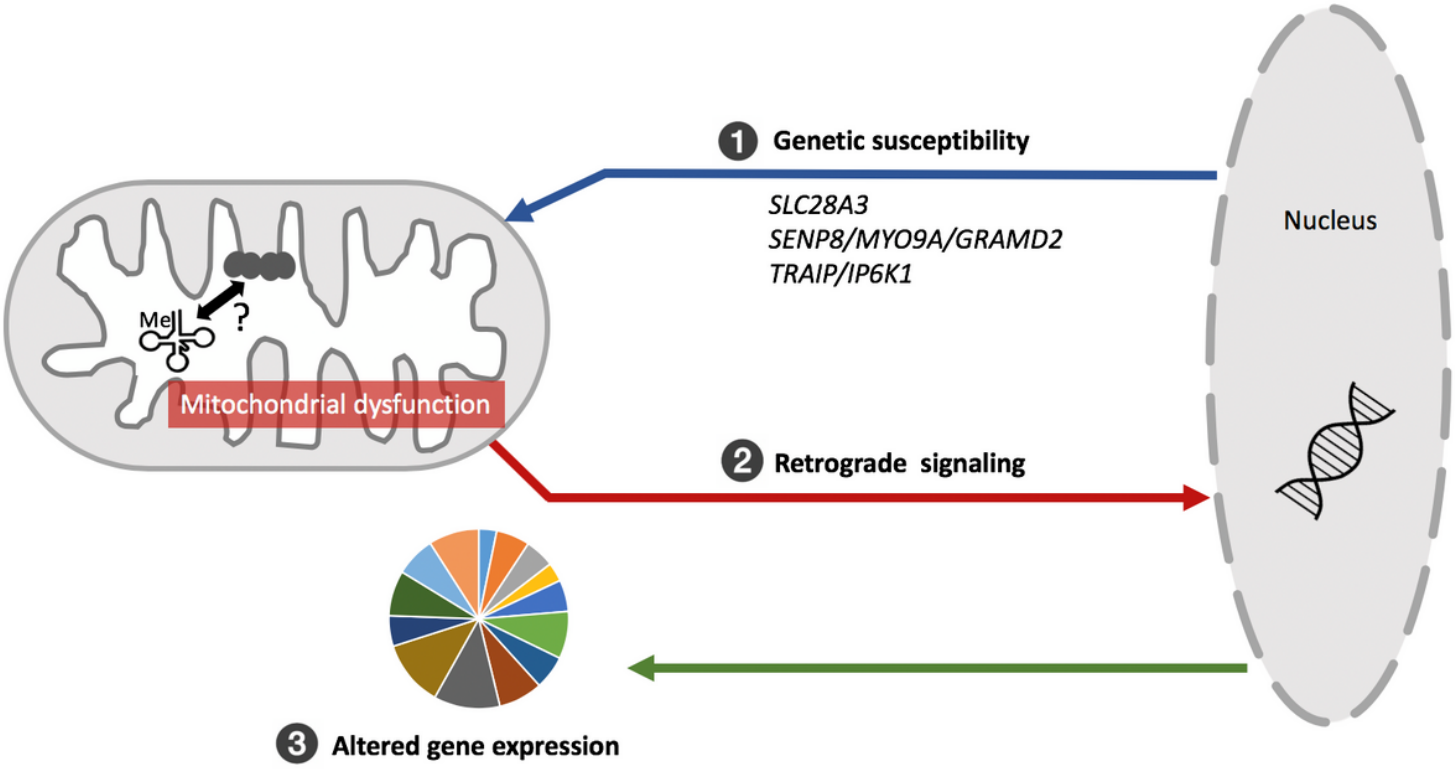

Figure 6

Hypothetical schematic of potential causes and effects of p9 methylation in neurodegeneration. (1) Nuclear encoded genetic variants may be causal for p9 hypermethylation within neuronal mitochondria, resulting in onset or exacerbation of electron transport chain dysfunction. (2) Subsequent mitochondrial dysfunction may lead to altered downstream nuclear gene expression via retrograde signaling resulting in (3) changes to various molecular and cellular processes that further contribute to neurodegenerative pathology.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile2pairwisecomparisons.xlsx
- AdditionalFile4PANTHEROverrep.xlsx
- AdditionalFile3genebasedGWAS.xlsx
- AdditionalFileboxplotbysitebydiagnosis.pdf
- Additionalfile1Figures.docx