Received: 15 August 2022 | Revised: 9 November 2022 | Accepted: 7 December 2022

DOI: 10.1002/jimd.12578

ORIGINAL ARTICLE

JIMD SSIEM WILEY

Transcriptomic analyses reveal neuronal specificity of Leigh syndrome associated genes

Azizia Wahedi[1,2] | **Chandika Soondram**[1,3] | **Alan E. Murphy**[4,5] | **Nathan Skene**[4,5] | **Shamima Rahman**[1,6]

[1]Mitochondrial Research Group, Genetics and Genomic Medicine Department, UCL Great Ormond Street Institute of Child Health, London, UK

[2]Medical Sciences Division, University of Oxford, Oxford, UK

[3]Department of Biochemistry, University College London, London, UK

[4]UK Dementia Research Institute at Imperial College London, London, UK

[5]Department of Brain Sciences, Imperial College London, London, UK

[6]Metabolic Unit, Great Ormond Street Hospital, London, UK

Correspondence
Shamima Rahman, Mitochondrial Research Group, UCL Great Ormond Street Institute of Child Health, London WC1N 1EH, UK.
Email: shamima.rahman@ucl.ac.uk

Funding information
Great Ormond Street Hospital Children's Charity; Lily Foundation; National Institute of Health Research (NIHR) Great Ormond Street Hospital Biomedical Research Centre

Communicating Editor: Eva Morava

Abstract

Leigh syndrome is a rare, inherited, complex neurometabolic disorder with genetic and clinical heterogeneity. Features present in affected patients range from classical stepwise developmental regression to ataxia, seizures, tremor, and occasionally psychiatric manifestations. Currently, more than 100 monogenic causes of Leigh syndrome have been identified, yet the pathophysiology remains unknown. Here, we sought to determine the cellular specificity within the brain of all genes currently associated with Leigh syndrome. Further, we aimed to investigate potential genetic commonalities between Leigh syndrome and other disorders with overlapping clinical features. Enrichment of our target genes within the brain was evaluated with co-expression (CoExp) network analyses constructed using existing UK Brain Expression Consortium data. To determine the cellular specificity of the Leigh associated genes, we employed expression weighted cell type enrichment (EWCE) analysis of single-cell RNA-Seq data. Finally, CoExp network modules demonstrating enrichment of Leigh syndrome associated genes were then utilised for synaptic gene ontology analysis and heritability analysis. CoExp network analyses revealed that Leigh syndrome associated genes exhibit the highest levels of expression in brain regions most affected on MRI in affected patients. EWCE revealed significant enrichment of target genes in hippocampal and somatosensory pyramidal neurons and interneurons of the brain. Analysis of CoExp modules enriched with our target genes revealed preferential association with pre-synaptic structures. Heritability studies suggested some common enrichment between Leigh syndrome and Parkinson disease and epilepsy. Our findings suggest a primary mitochondrial dysfunction as the underlying basis of Leigh syndrome, with associated genes primarily expressed in neuronal cells.

KEYWORDS

EWCE, Leigh syndrome spectrum, primary mitochondrial disease, transcriptomics

Azizia Wahedi and Chandika Soondram contributed equally to this work.

1 | INTRODUCTION

Leigh syndrome (subacute necrotizing encephalomyelopathy), first described in 1951, is a severe, early-onset inherited neurometabolic disorder that primarily affects the central nervous system. With a minimum birth prevalence of 1 in every 40 000 live births,[1,2] Leigh syndrome is the most frequent clinical presentation of mitochondrial disease in childhood. Onset typically varies between 3 and 12 months of age, before ultimately leading to early death.[3] Survival rates vary according to the specific gene defect, but one large multinational study reported a median age of death of approximately 2.4 years across the whole cohort.[3] Patients affected by Leigh syndrome typically experience global developmental delay and/or regression, coupled with seizures, ataxia, hypotonia, dystonia, and optic atrophy. These neurological symptoms are related to progressive bilateral, often symmetrical lesions affecting the basal ganglia, midbrain, and brainstem.[1,4] The cerebellum, white matter, and spinal cord may also be affected in some cases. In addition, Leigh syndrome was associated with multisystemic disease manifestations in some patients who presented hepatic, renal, and cardiac symptoms.

Since the advent of next-generation exome and genome sequencing, an increasing number of genetic causes of Leigh syndrome has been identified.[5] As of 2021, a total of 113 disease-causing genes have been reported (https://search.clinicalgenome.org/kb/conditions/MONDO:0009723) reflecting the phenotypic and biochemical heterogeneity of the disorder. The genetic causes encompass mutations in both mitochondrial DNA (mtDNA), which may be maternally inherited or sporadic, and nuclear DNA, which are usually inherited in an autosomal recessive pattern (but may rarely be dominant or X-linked). As far as mtDNA mutations are concerned, they can either be homoplasmic or heteroplasmic (i.e., co-existence of mutated and wild-type mtDNA), and the mutation load appears to, at least partly, determine the phenotype. Most patients with Leigh syndrome have causative mutations in nuclear DNA genes, while mtDNA mutations are responsible in approaching 25% of cases.[3]

Leigh syndrome and other mitochondrial disorders form a group of approaching 400 different rare monogenic disorders unified by impacts on mitochondrial structure and function.[6] These mitochondrial disorders are characterised by extreme clinical heterogeneity, leading to considerable diagnostic odysseys for affected patients. Moreover, there is currently a dearth of effective disease-modifying therapies. The aim of this study was to investigate the cellular specificity of Leigh syndrome associated genes within the brain, which may ultimately facilitate the identification of novel therapeutic targets.

This project involved using a systems biology approach to explore the cellular specificity of 113 genes linked to Leigh syndrome, which have recently been curated by a global ClinGen expert panel (https://search.clinicalgenome.org/kb/conditions/MONDO:0009723). We hypothesised that genes with mutations known to be causative of Leigh syndrome might have increased or decreased expression in brain regions most affected in Leigh syndrome and aimed to determine which of these possibilities was the case. In summary, a transcriptomic data set was utilised used to (i) compare gene expression levels across 10 UKBEC brain regions available in transcriptomic data from the United Kingdom Brain Expression Consortium (UKBEC) network, (ii) investigate the cell type(s) which are most relevant to the disease, (iii) identify whether the genes enriched in selected modules are associated with presynaptic or postsynaptic structures, and (iv) test for genetic commonalities, if any, between Leigh syndrome and other neurological disorders.

2 | MATERIALS AND METHODS

2.1 | Co-expression network analysis

Co-expression (CoExp) is a resource developed to exploit gene co-expression networks (GCNs) to build gene groupings based on transcriptomic profiling.[7] The CoExp tool is an online web application consisting of a collection of 109 networks powered by the CoExpNets suite of R packages. The enrichment of the 113 genes of interest (Table 1) was tested within the CoExp networks generated from different brain regions, available in the CoExp Web Application (https://rytenlab.com/coexp). The 10 human brain regions within the UKBEC network which were investigated were as follows: cerebellar cortex (CRBL), frontal cortex (FCTX), hippocampus (HIPP), medulla oblongata (MEDU), occipital cortex (OCTX), putamen (PUTM), substantia nigra (SNIG), temporal cortex (TCTX), thalamus (THAL), and white matter (WHMT). The UKBEC data set comprises 134 individuals confirmed to be neuropathologically normal and of European descent from 16 to 102 years of age and was generated using Affymetrix exon 1.0 ST arrays.[8]

The gene annotation functionality of the CoExp website returns a table with detailed information regarding the genes present and the modules in which they are found within a specific network. Table S5 provides details of the colour codes for modules assigned by the CoExp software, which groups genes based on ontology terms and similar functions. The FDR-adjusted p-value was used for analysis, as described by Mencacci et al.[9] where

TABLE 1 List of Leigh syndrome associated genes used in the EWCE analysis study

Function	Biochemical defect	Gene defect(s)	Inheritance mode
Complex I	Enzyme subunits	*MT-ND1, MT-ND2, MT-ND3, MT-ND4, MT-ND5, MT-ND6*	Maternal/sporadic
		NDUFA2, NDUFA9, NDUFA10, NDUFA12, NDUFA13, NDUFB8, NDUFC2, NDUFS1, NDUFS2, NDUFS3, NDUFS4, NDUFS7, NDUFS8, NDUFV1, NDUFV2	
		NDUFA1	X-linked
	Assembly factors	*NDUFAF2, NDUFAF4, NDUFAF5, NDUFAF6, NDUFAF8, FOXRED1, NUBPL, TIMMDC1*	AR
Complexes II, III, V	Enzyme subunits	*SDHA, UQCRQ, MT-ATP*	AR
		ATP5MK	Maternal/sporadic
	Assembly factors	*SDHAF1, BCS1L, TTC19*	AR
Complex IV	Enzyme subunits	*MT-CO1, MT-CO2, MT-CO3*	Maternal/sporadic
		COX4I1, COX8A, NDUFA4	AR
	Assembly factors	*COX10, COX15, SCO2, SURF1, PET100, PET117*	AR
Disorders of pyruvate metabolism	Pyruvate dehydrogenase	*PDHA1*	X-linked
		DLAT, DLD, PDHB, PDHX	AR
Disorders of vitamin transport and metabolism and cofactor biosynthesis	Biotin	*BTD*	AR
	Thiamine	*SLC19A3, SLC25A19, TPK1*	AR
	Coenzyme Q10	*COQ9, PDSS2*	AR
	Lipoic acid	*LIAS, LIPT1, MECR*	AR
mtDNA maintenance	MtDNA depletion and/or multiple mtDNA deletions	*POLG, SUCLA2, SUCLG1, RNASEH1*	AR
		SLC25A4, SSBP1	De novo
Disorders of mitochondrial gene expression	Impaired mitochondrial protein synthesis	*MT-TI, MT-TK, MT-TL1, MT-TL2, MT-TV, MT-TW*	Maternal/sporadic
		EARS2, FARS2, GFM1, GFM2, GTPBP3, IARS2, LRPPRC, MRPS34, MTFMT, MTRFR, NARS2, PNPT1, PTCD3, TACO1, TARS2, TRMU, TSFM	AR
Mitochondrial protein quality control, mitochondrial membrane, and mitochondrial dynamics	Proteostasis	*CLPB, LONP1*	AR
	Lipid remodelling	*SERAC1*	AR
	Fission	*MFF, SLC25A46, VPS13D, DNM1L*	AR
	Fission	*DNM1L*	AD
	Fusion	*OPA1*	AR
Mitochondrial toxicity	Sulphide metabolism	*ETHE1, SQOR*	AR
	Valine degradation	*ECHS1, HIBCH*	AR
	Detoxification	*NAXE, SLC39A8*	AR
Miscellaneous	Unknown	*AIFM1*	X-linked
		HPDL, FBXL4, ADAR, RANBP2, MORC2, NUP62	AR

Note: Genes have been categorised based on the disease mechanism.

Abbreviations: AD, autosomal dominant; AR, autosomal recessive; mtDNA, mitochondrial DNA; OXPHOS, oxidative phosphorylation.

Brain regions	Modules			
CRBL	Tan	Magenta		
FCTX	Yellow	Red		
HIPP	Tan	Cyan	Steel blue	White
MEDU	Yellow	Light cyan	Midnight blue	
OCTX	Green yellow	Black	Dark magenta	
PUTM	Grey60	Red	Brown	
SNIG	Yellow	Salmon	Brown	
TCTX	Green	Salmon	Magenta	
THAL	Dark green	Light green	Magenta	
WHMT	Grey60	Pink		

TABLE 2 Leigh syndrome-enriched modules, generated from UKBEC gene co-expression network analysis

Abbreviations: CRBL, cerebellum; FCTX, frontal cortex; HIPP, hippocampus; MEDU, medulla; OCTX, occipital cortex; PUTM, putamen; SNIG, substantia nigra; TCTX, temporal cortex; THAL, thalamus; WHMT, white matter.

values <0.05 satisfied the argument that those genes were significantly enriched within the specific modules. For 79 genes of interest that were available in the UKBEC network, specific modules in which Leigh syndrome associated genes were significantly enriched were identified (Table 2).

2.2 | Venny analysis

A Venny diagram (https://bioinfogp.cnb.csic.es/tools/venny/) was used to test for overlapping Leigh syndrome genes amongst enriched modules and to summarise the CoExp findings from Table 3. Only the genes within modules with enrichment of the largest number of Leigh associated genes (substantia nigra, putamen, medulla, and thalamus) were used as input for the diagram.

2.3 | Expression-weighted cell-type enrichment analysis

Expression-weighted cell-type enrichment (EWCE) analysis was performed to investigate whether certain cell types express the Leigh syndrome associated genes more than can be attributed to random chance.[10] In this study, we used the EWCE method with the target gene list (Table 1) and background gene set (defined as all other genes within the data set) as the two arguments to determine the probability that the 113 genes exhibit higher specificity in a given cell type. The data set within the EWCE package includes the Karolinska Institute single-cell RNA-Seq data from mouse (juvenile P22-P32 CD1 *Mus musculus*) brain,[11] thus genes in our target list were converted to their mouse homologous, through a function within EWCE that converts the HGNC gene name into its corresponding MGI mouse orthologue gene. Based on the standard deviation from the mean, the enrichment of Leigh syndrome genes in level 1 and level 2 cell types was determined, with the levels representing higher (cell types) and lower level (cell subtypes) groupings, respectively. To investigate whether subgrouping the genes based on product function would reveal differential enrichment, we reran the EWCE analysis on functional groupings of the Leigh syndrome associated genes (Table 1). We were unable to perform the same analysis with the genes subdivided based on whether the mode of inheritance is mitochondrial or nuclear because no mitochondrial encoded genes are present within the EWCE dataset. Instead, we utilised the online STRING database platform to identify predicted direct (physical) and indirect (functional) associations of the products of the protein coding mitochondrial genes linked to Leigh syndrome. The corresponding genes for proteins interacting with our target gene list products were compiled and utilised as an extended 'mitochondrial' gene set and then run through EWCE. EWCE was run on 10 000 iterations for the complete list of genes and was increased to 100 000 bootstrap replicates to compensate for the multiple testing. We also controlled for transcript length and GC content of the genes. Data are displayed as standard deviations from the mean, and *p*-values were obtained using the Benjamini–Hochberg FDR method to account for multiple testing.

Next, we investigated three cell populations in which Leigh syndrome associated genes were found to be significantly enriched, namely, pyramidal CA1, pyramidal SS, and interneurons, to determine the specificity values of the genes. Specificity is defined as the average expression of a gene in a particular cell type by the average

TABLE 3 List of Leigh syndrome disease-causing genes enriched in UKBEC modules

Modules	Leigh syndrome disease genes
CRBL tan	*NDUFA9, NDUFA12, NDUFC2, NDUFS3, COX15, PDHA1, DLAT, PDHX, SUCLA2, NARS2, TSFM, DNM1L, ETHE1, AIFM1*
CRBL magenta	*NDUFS4, DLD, PDHB, LIAS, SUCLG1, GFM1, GFM2, LRPPRC, SERAC1, MFF, SLC25A46, RANBP2*
FCTX yellow	*NDUFS4, NDUFAF4, DLD, PDHB, SLC25A4, GFM1, LRPPRC, SERAC1, MFF, SLC25A46, OPA1, RANBP2*
FCTX red	*NDUFA9, NDUFA12, NDUFC2, NDUFS3, TTC19, PDHA1, DLAT, PDHX, SUCLA2, MRPS34, NARS2, TACO1, DNM1L, ETHE1, AIFM1*
HIPP tan	*NDUFA2, LIAS, SUCLG1, GFM1, LRPPRC, ETHE1, RANBP2*
HIPP cyan	*NDUFA9, NDUFC2, NDUFS3, PDHA1, DLAT, PDHX, SUCLA2, NARS2, TSFM, DNM1L, AIFM1*
HIPP steel blue	*NDUFS2, NDUFAF4, SDHA, GFM2, SLC25A46, OPA1, ADAR*
HIPP white	*NDUFA10, LIPT1, FARS2, BTD, PTCD3, PNPT1, MFF, HIBCH*
MEDU yellow	*NDUFA9, NDUFA12, NDUFB8, NDUFS3, NDUFV1, NDUFA1, SDHA, COX8A, NDUFA4, DLAT, DLD, MECR, TACO1, TARS2, CLPB, OPA1, ADAR*
MEDU light cyan	*NDUFS2, NDUFS4, BCS1L, PDHB, LIAS, SUCLG1, RNASEH1, SSBP1, FARS2, GFM1, IARS2, LRPPRC, PTCD3, MFF, SERAC1, SLC25A46, RANBP2*
MEDU midnight blue	*NDUFC2, FOXRED1, NUBPL, PDHX, SLC25A19, SUCLA2, MRPS34, NARS2, ECHS1, AIFM1*
OCTX green yellow	*NDUFA9, NDUFA12, NDUFS3, FOXRED1, COX10, PDHX, SUCLA2, NARS2, TSFM*
OCTX black	*NDUFA2, NDUFS2, NDUFAF4, DLD, PDHB, BCS1L, SUCLG1, RNASEH1, LRPPRC, SLC25A46, RANBP2*
OCTX dark magenta	*COX8A, COX15, DLAT, SLC25A19, MRPS34, MORC2, CLPB*
PUTM grey60	*NDUFS4, UQCRQ, BCS1L, BTD FBXL4, LIAS, LIPT1, GFM1, GFM2, PNPT1, PTCD3, SERAC1, MFF*
PUTM red	*NDUFA2, NDUFS2, NDUFAF4, SDHA, DLD, PDHB, MECR, SUCLG1, SSBP1, IARS2, LRPPRC, SLC25A46, ETHE1, OPA1, ADAR, RANBP2*
PUTM brown	*NDUFA9, NDUFA12, NDUFC2, TTC19, COX4I1, PDHA1, DLAT, PDHX, SUCLA2, MRPS34, NARS2, DNM1L*
SNIG yellow	*NDUFA9, NDUFB8, NDUFC2, NDUFS7, NDUFV1, FOXRED1, SDHA, TTC19, DLAT, MECR, SLC25A19, SUCLA2, EARS2, TRMU, TSFM, MORC2, TARS2, CLPB, DNM1L*
SNIG salmon	*NDUFA2, NDUFA10, NDUFS2, NDUFS4, NDUFAF4, DLD, PDHB, LIAS, SUCLG1, RNASEH1, SSBP1, GFM1, PTCD3, LRPPRC, MFF, SERAC1, SLC25A46, OPA1, ADAR, RANBP2*
SNIG brown	*NDUFA1, NDUFA12, NDUFA13, NDUFS3, NDUFA4, NUBPL, PDHA1, PDHX, MRPS34, NARS2, LONP1, ECHS1*
TCTX green	*NDUFA9, NDUFA12, FOXRED1, TTC19, SLC25A19, PDHA1, DLAT, SUCLA2, TSFM, TACO1, AIFM1*
TCTX salmon	*FBXL4, BTD, TPK1, LIPT1, SSBP1, IARS2, PNPT1, PTCD3, HIBCH*
TCTX magenta	*NDUFA2, NDUFS2, NDUFAF4, NDUFA4, SDHA, DLD, GFM2, LRPPRC, SLC25A46, OPA1, ADAR*
THAL dark green	*NDUFA2, NDUFS4, DLD, PDHB, SUCLG1, RNASEH1, IARS2 FARS2, LRPPRC, SERAC1, MFF, SLC25A46, OPA1, RANBP2*
THAL light green	*NDUFA9, NDUFA12, NDUFC2, NDUFS3, NUBPL, DLAT, SUCLA2, GFM, NARS2, TSFM, ECHS1*
THAL magenta	*NDUFA4, NDUFS2, NDUFV1, FOXRED1, SDHA, PDHA1, SLC25A19, LIAS, GTPBP3, TRMU, AIFM1*
WHMT grey60	*NDUFA9, NDUFC2, TTC19, DLAT, DLD, PDHX, SUCLA2, NARS2, GFM2, LRPPRC, TSFM, SLC25A46, DNM1L, OPA1, ADAR, RANBP2*
WHMT pink	*NDUFA10, PDHB, BTD, LIAS, LIPT1, SUCLG1, FARS2, GFM1, PTCD3, PNPT1, FARS2, SERAC1, MFF*

Abbreviations: CRBL, cerebellum; FCTX, frontal cortex; HIPP, hippocampus; MEDU, medulla; OCTX, occipital cortex; PUTM, putamen; SNIG, substantia nigra; TCTX, temporal cortex; THAL, thalamus; WHMT, white matter.

expression of the same gene in all other cell types. The specificity values of Leigh syndrome genes in the respective cell types were obtained from a previous publication.[12]

2.4 | Synaptic association

To investigate the association of Leigh syndrome enriched modules with presynaptic and/or postsynaptic

structures, the modules listed in Table 2 were annotated in the CoExp website and then inputted in the synaptic gene ontologies (SynGO)[13] website (https://syngoportal.org/). The latter analysed the gene list for synaptic association, compared to a default background of all brain expressed genes (GTExV7 brain tissue), using a one-sided Fisher exact test. Once a term particularly enriched with the gene cluster was found, multiple testing corrections were applied using the FDR, and the resulting plots were coloured according to the enrichment *q*-values.

2.5 | Heritability enrichment analysis

A heritability study was performed to assess the potential co-morbidity of Leigh syndrome with neurological disorders characterised by overlapping clinical features (such as dystonia and seizures). Candidate genes for Parkinson disease,[14] epilepsy,[15] and schizophrenia[16] were obtained from the literature, and FDR-adjusted *p*-values were derived from the CoExp Gene Set Annotation tool. The CoExp modules in Table 2, containing more than 15 Leigh syndrome associated genes, were selected, and the FDR-adjusted *p*-values for the different neurological disorders were used for the plot, with a nominal enrichment threshold of $p < 0.05$.

2.6 | Data availability and URLs

Data used to generate Figures 1, 2, and 5 are available within the Supporting Information (Tables S1, S2, and S4, respectively). The level 1 (cell type) and level 2 (cell subtypes) specificity values for the Karolinska single-cell mouse RNA-Seq superset can be retrieved online from Nathan Skene's GitHub (https://github.com/NathanSkene/MAGMA_Celltyping). All gene groupings, enrichment annotations, and heritability analyses were performed using the CoExp website (www.rytenlab.com/coexp/Run/Catalog/), along with its supplementary CoExpNets packages (https://github.com/juanbot/CoExpNets). Gene lists from the CoExp tool were inputted into the Venny 2.1 online tool (https://bioinfogp.cnb.csic.es/tools/venny/) and the freely accessible SynGO website (https://www.syngoportal.org/) for additional analyses.

3 | RESULTS

3.1 | Leigh syndrome associated genes are highly enriched in the SNIG

Utilising human brain transcriptomic data from the UKBEC network within CoExp we sought to identify regions of the brain where Leigh syndrome associated genes were enriched. Of the 113 genes associated with Leigh syndrome (Table 1), 79 were represented within the UKBEC networks. Subsequently, regions of the brain found to have significant enrichment (FDR adjusted *p*-value <0.05) of these genes were identified (Table 3, Figure 1A). CoExp network modules with significant enrichment of Leigh syndrome associated genes include four modules of the hippocampus, three modules of the occipital cortex, pitamen, substantia nigra, and temporal cortex, and two modules of the cerebellum, frontal cortex, medulla, and white matter (Table 2). Modules with enrichment of the most Leigh syndrome associated genes include the red and yellow modules within the frontal cortex, yellow and light cyan modules of the medulla, the red module for the putamen, grey60 for white matter, and the yellow and salmon modules of the substantia nigra (Table 3). Over six-fold significant enrichment above background was observed in the salmon module of the substantia nigra. Significant enrichment four-fold above background was observed in the following four modules: frontal cortex red, putamen red, thalamus dark green, and white matter grey60.

Of the four brain regions with the most Leigh syndrome associated genes identified, namely, the substania nigra, putamen, medulla, and thalamus, the modules with the most Leigh syndrome associated genes were compared to one another to elucidate any common genes (Figure 1B). Of the genes significantly enriched, 19 were in common between the four brain regions, namely *NDUFS2*, *NDUFS4*, *NDUFA9*, *NDUFA12*, *NDUFC2*, *SDHA*, *LRPPRC*, *PDHB*, *DLAT*, *DLD*, *SUCLA2*, *SUCLG1*, *NARS2*, *GFM1*, *OPA1*, *SERAC1*, *MFF*, *LIAS*, and *RANBP2*. Gene associations unique to the substantia nigra were *NDUFS7*, *NDUFA10*, *NDUFA13*, *EARS2*, *LONP1*, and *MORC2*. Putaminal-unique genes were *UQCRQ*, *COX4I1*, *FBXL4*, *PNPT1*, *GFM2*, *LIPT1*, *BTD*, and *ETHE1*. Medulla unique genes include *COX8A*, *NUBPL*, and *TACO1*. Two genes unique to the thalamus are *SLC24A46* and *GTPBP3*. The substantia nigra was found to have enrichment of greater than 64% of the Leigh syndrome associated genes present in the UKBEC network. Of note, nearly 10% of the genes included in the Venny (i.e., target genes expressed in the substantia nigra, putamen, medulla, or thalamus) are only expressed in the substantia nigra and not the putamen, medulla, or thalamus.

3.2 | Leigh syndrome genes are significantly enriched in interneurons and pyramidal cells

To build upon the CoExp findings that Leigh syndrome associated genes are enriched in neuronal tissue,

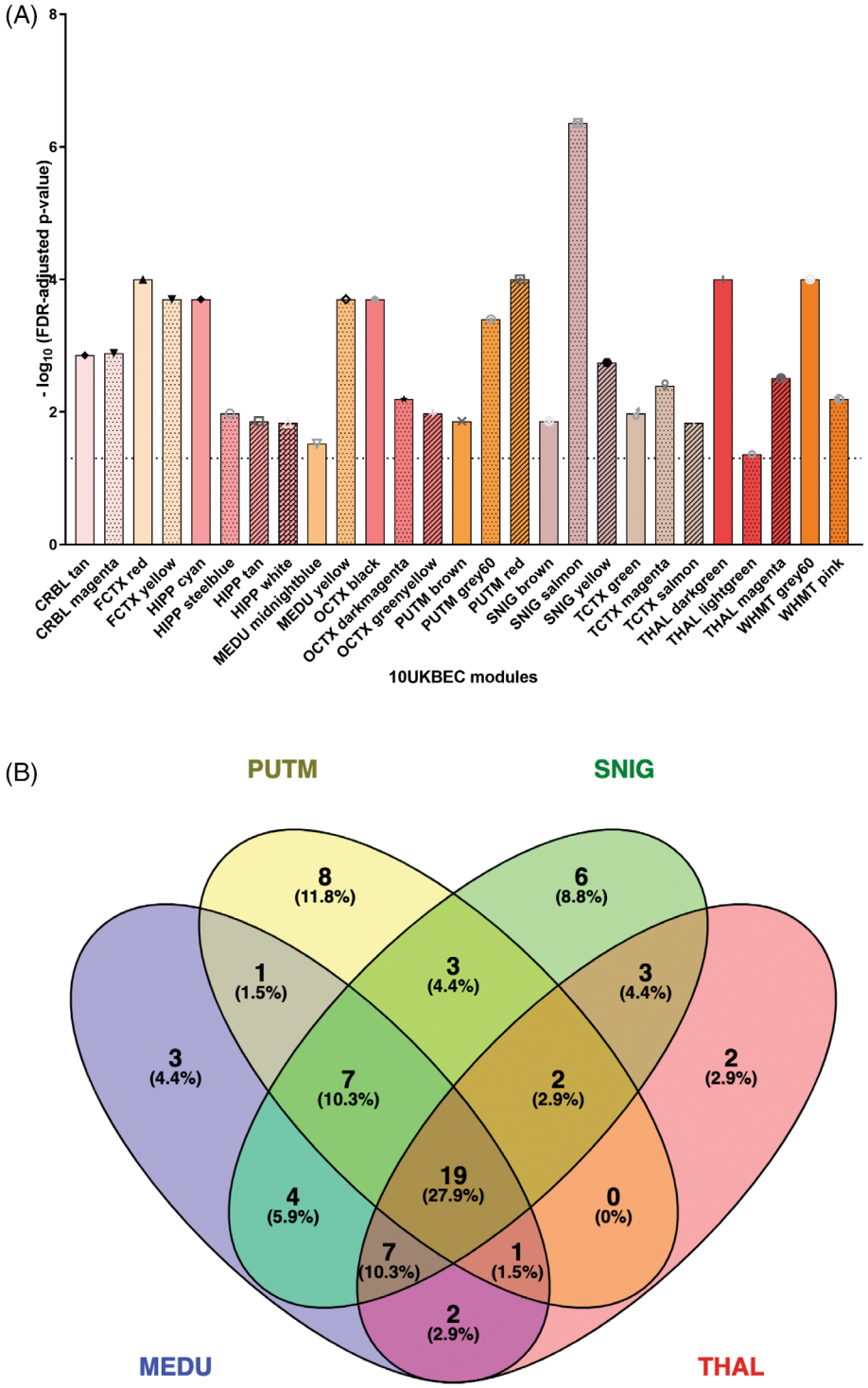

FIGURE 1 Expression of Leigh syndrome associated genes within the enriched UKBEC modules. (A) Expression of Leigh syndrome associated genes in UKBEC network co-expression modules. (B) Venny diagram showing that 64% of the genes within four regions of interest are enriched within the substantia nigra with one or more overlapping modules, and almost 10% of the genes are only expressed in the substantia nigra. This suggests that the substantia nigra is likely to be the most affected in patients with Leigh syndrome, owing to its high number of Leigh syndrome-associated genes. The dashed line marks the approximate FDR threshold cut-off for nominal enrichment (~1.301). MEDU, medulla; PUTM, putamen; SNIG, substantia nigra; THAL, thalamus

particularly with 64% of genes being enriched within the substantia nigra, we interrogated the Karolinska Institute single-cell RNA-Seq mouse brain data set using EWCE.[11] Cells present in level 1 (cell types) of this data set include interneurons, pyramidal cells of the HIPP (CA1) and somatosensory cortex (SS), astrocytes, glial cells, microglia, oligodendrocytes, astrocytes (ependymal), and endothelial mural cells.

EWCE analysis of the level 1 (higher level cell types) data set revealed that Leigh syndrome associated genes are significantly enriched in interneurons, pyramidal CA1, and pyramidal SS cell types (FDR adjusted $p = 0.0162$, FDR adjusted $p = 0.0027$, FDR adjusted $p = 0.0036$, respectively, Figure 2A). Notably, OXPHOS enzyme complex subunits and assembly factors were found to be among the top 10 genes in all three cell types where Leigh

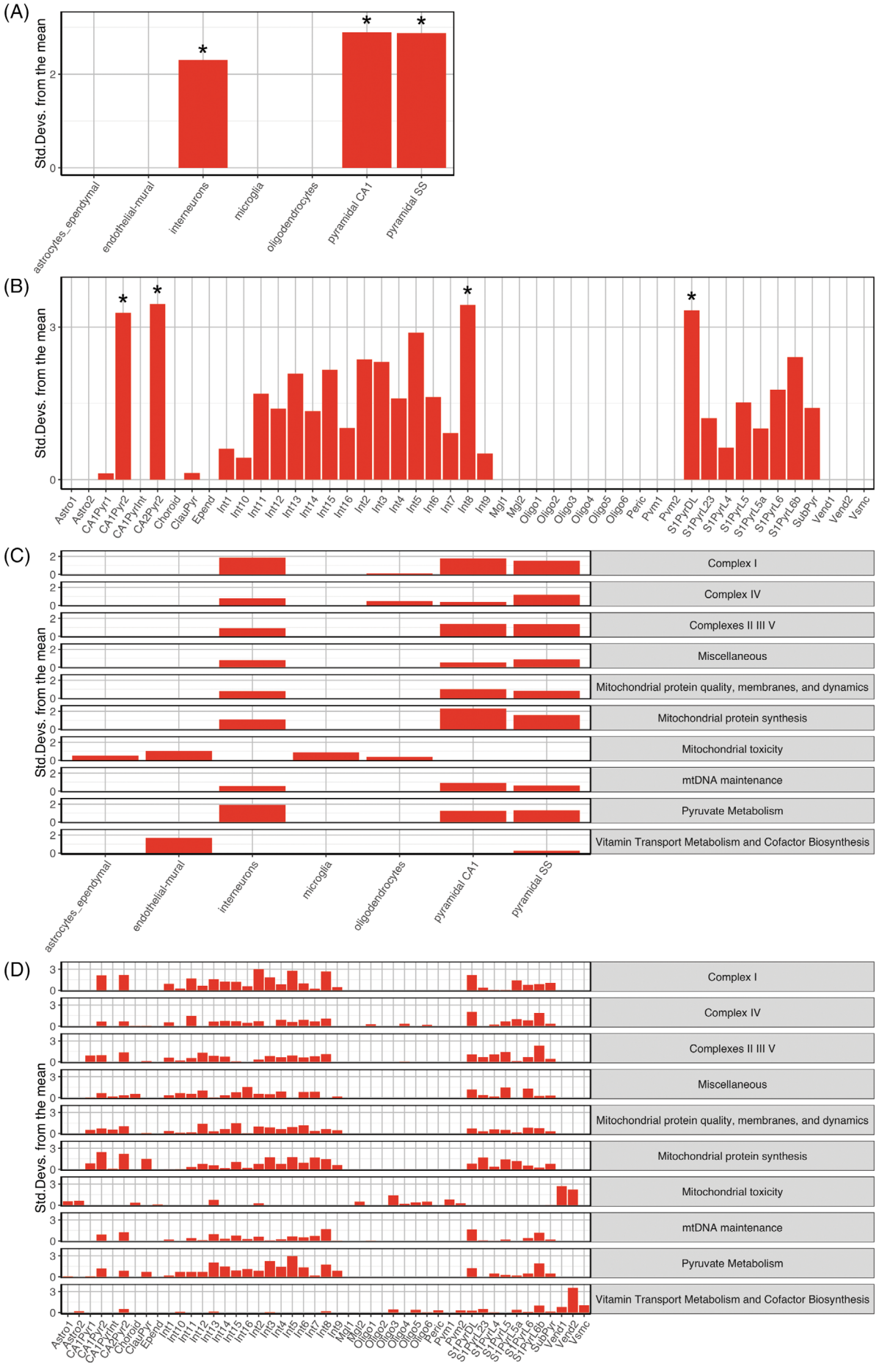

FIGURE 2 Legend on next page.

syndrome associated genes were significantly enriched. EWCE analysis of the cell subtypes (level 2, Figure 2B) revealed significant enrichment of Leigh associated genes in type 2 pyramidal cells of CA1 (FDR adjusted $p = 0.0031$) and CA2Pyr2 (FDR adjusted $p = 0.0026$). Significant enrichment was also observed in interneuron type 8 (Int8, FDR adjusted $p = 0.0033$) and deep layer pyramidal cells of S1 (S1PyrDL, FDR adjusted $p = 0.0015$).

Further, we sought to determine whether there was differential enrichment of genes based on the function of their products. Following EWCE analysis of the genes subdivided into groups based on similar functions (Table 1), a pattern emerged whereby genes encoding proteins and assembly factors of Complexes I–IV, or involved in mitochondrial protein quality, synthesis, membranes, and dynamics were expressed in interneurons and the CA1 and SS pyramidal cells (Figure 2C). Genes encoding products that are involved in pyruvate metabolism and mtDNA maintenance are similarly expressed. In contrast, genes encoding proteins whose dysfunction results in mitochondrial toxicity were found to be expressed in astrocytes and ependymal cells, endothelial and mural cells, microglia, and oligodendrocytes (Figure 2C). Leigh syndrome associated genes functioning in vitamin transport metabolism and cofactor biosynthesis were expressed in endothelial and mural cells, with some expression in pyramidal SS cells (Figure 2C). EWCE analysis of the level 2 cell types revealed a similar pattern to that observed for the level 1 data set, where genes associated with mitochondrial toxicity, vitamin transport metabolism and cofactor biosynthesis had an expression profile different from the rest of the functional groups (Figure 2D). It should be noted that EWCE findings for the functional groupings did not yield any statistically significant results after correcting for multiple testing when examining both cell types and subtypes, however some trends were observed.

The genes with the highest specificity in CA1Pyr2 cells were *NDUFAF4*, *NDUFA12*, *MTFMT*, *PDSS2*, *SUCLA2*, *SUCLG1*, *NARS2*, *GTPBP3*, *EARS2*, and *TACO1* (Figure 3A). In CAPyr2, highest specificity genes included *NDUFA10*, *NDUFA12*, *NDUFAF5*, *TTC19, PDSS2*, *SUCLA2*, *GTPBP3*, *MTFMT*, *MRPS34*, and *TSFM* (Figure 3B). Within Int8 cells, *NDUFAF5*, *NDUFV2*, *BCS1L*, *COX4I1*, *SURF1*, *MECR*, *SUCLA2*, *SUCLG1*, *SSBP1*, and *MRPS34* genes possessed the highest specificity (Figure 3C). Finally, within S1PyrDL cells, genes with the highest specificity were *NDUFA2*, *NDUFA4*, *NDUFA1*, *NDUFS4*, *NDUFS8*, *NDUFAF4*, *UQCRQ*, *COX4I1*, *MECR*, and *GFM2* (Figure 3D).

Examining the difference between nuclear and mitochondrially encoded genes associated with Leigh syndrome was inconclusive since none of the mitochondrially encoded genes associated with Leigh syndrome were represented in the data set used for the EWCE analysis. In view of this, we utilised the STRING database[17] to identify proteins interacting with products from our mitochondrial encoded target gene list and generated an extended 'mitochondrial' gene list (Table S3). EWCE analysis of the extended gene list revealed a similar pattern of expression in interneurons, pyramidal CA1, and pyramidal SS cells; however, the findings were not statistically significant (Figure S2). Of note, the STRING-database identified 19 interactors (*MT-CYB*, *MT-ND4L*, *MT-ATP8*, *NDUFA6*, *NDUFS5*, *NDUFS6*, *NDUFAB1*, *UQCRQ2*, *COX5A*, *COX5B*, *COX6A1*, *COX6B1*, *COX6C*, *COX7A2*, *COX7C*, *ATP5A1*, *ATP5B*, *ATP5D*, and *ATP5F1*) which are not currently identified as causes of Leigh syndrome but are possible candidate genes for Leigh syndrome.

3.3 | Leigh enriched modules are primarily associated with pre-synaptic structures

To understand the role of Leigh syndrome associated genes at the synaptic level and investigate the UKBEC CoExp modules further, we utilised the SynGO database (https://www.syngoportal.org/), an evidence-based and expert-curated resource for synapse function and gene enrichment analysis (Figure 4). CoExp modules containing the largest number of Leigh syndrome associated genes were found to be significantly enriched in synaptic structures. Of the modules of interest, four (TCTX magenta, TCTX green, SNIG yellow, and MEDU yellow; see Table S5 for a detailed explanation of the colour coding of modules) exhibited enrichment in both pre- and post-synaptic structures, with greater association with pre-synaptic structures, suggesting a potential synaptic function for the genes enriched in these modules. The hippocampus and white matter module of interest (HIPP steel blue and WHMT grey60, respectively) demonstrated enrichment in only pre-synaptic structures. We were

FIGURE 2 Leigh syndrome associated genes are significantly enriched in interneurons and pyramidal CA1 and SS neurons. Enrichment of Leigh syndrome associated genes in level 1 (A) and level 2 (B) cell types from the mouse brain single-cell RNA-Seq Karolinksa institute superset (reps = 10 000, * denotes significance where Benjamini–Hochberg FDR adjusted $p < 0.05$, controlled for GC content and length of genes). EWCE analysis of enrichment of Leigh syndrome associated genes grouped based on gene product function in level 1 (C) and level 2 (D) cell types from the Karolinska mouse brain superset (reps = 100 000, controlled for GC content and length of genes)

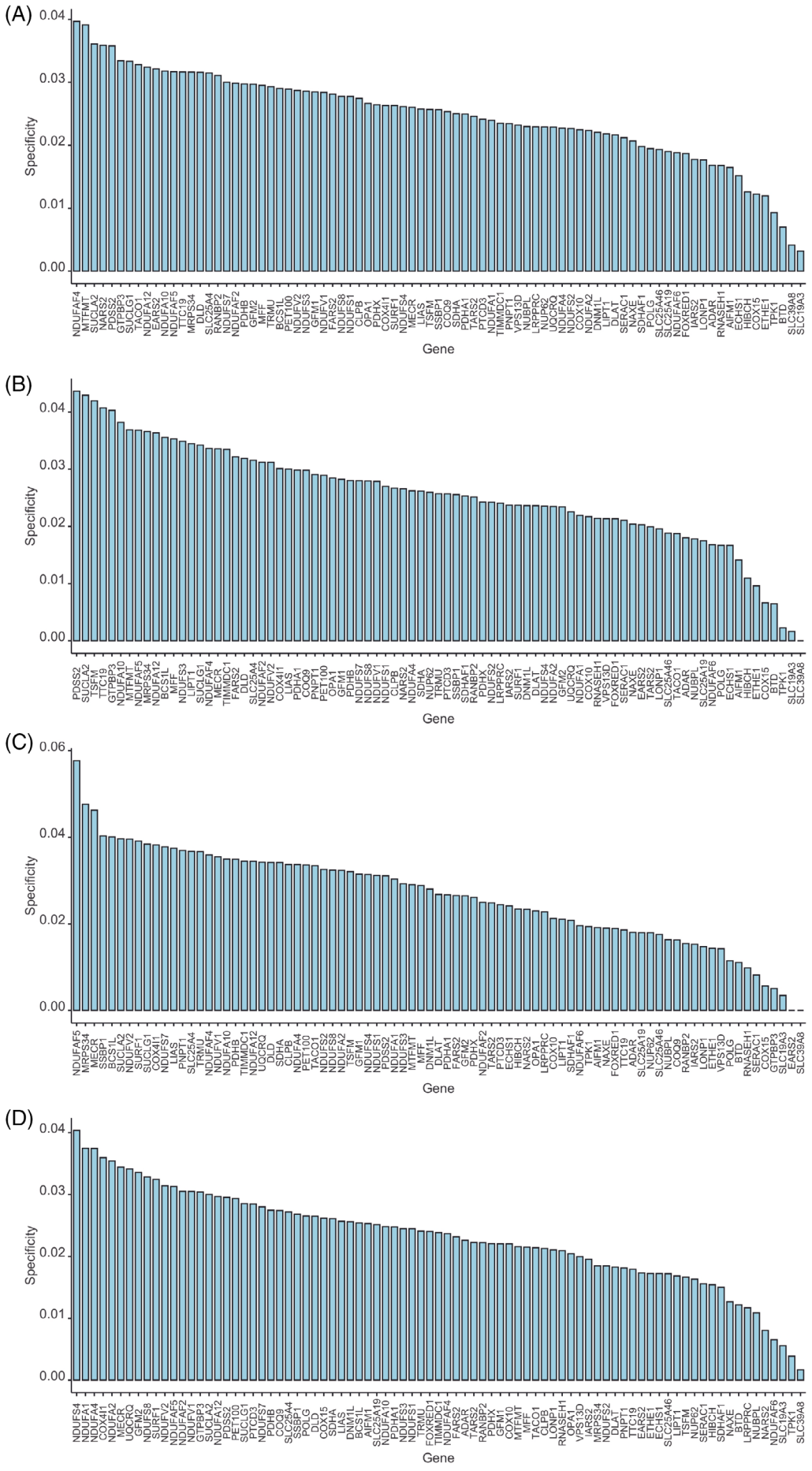

FIGURE 3 Legend on next page.

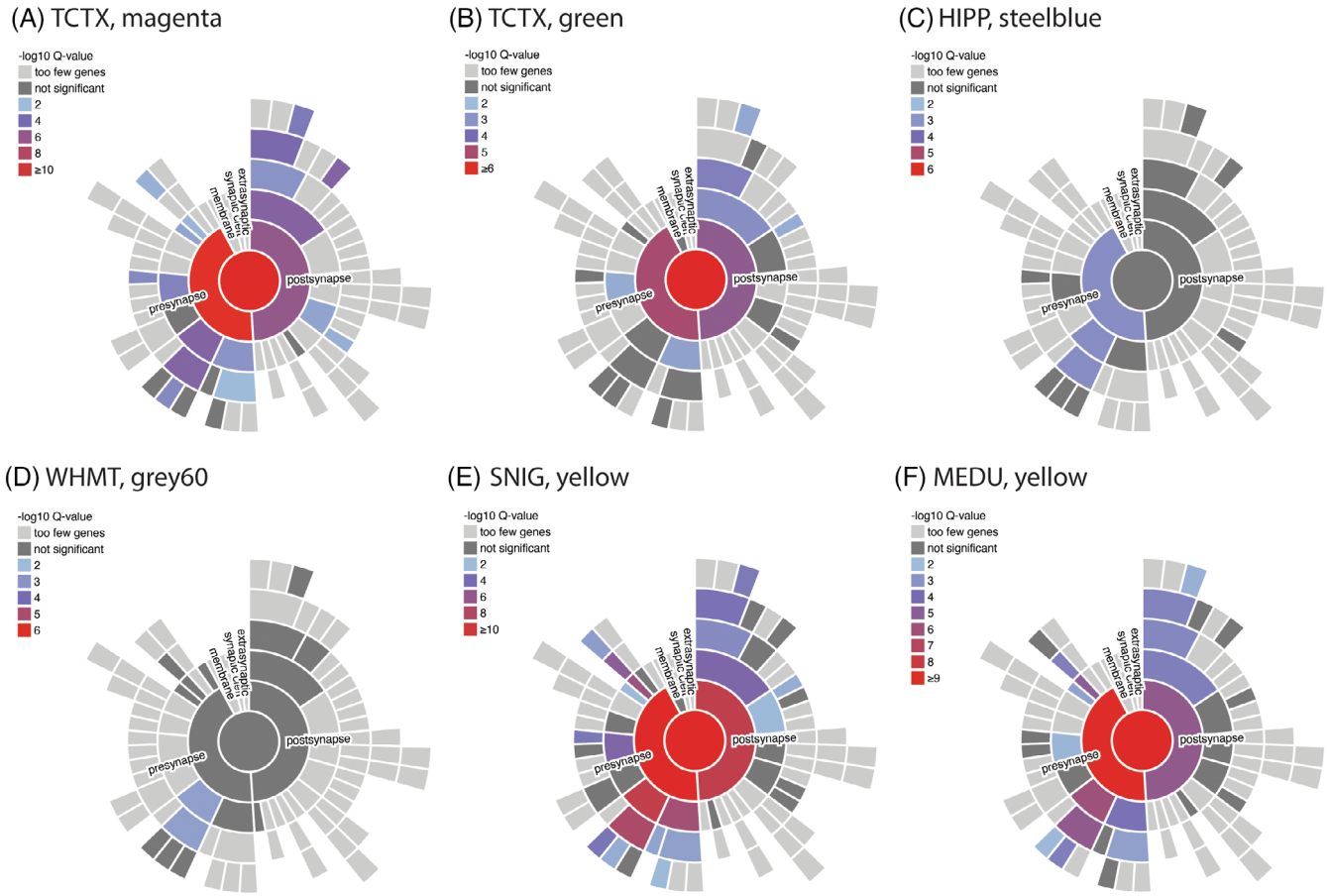

FIGURE 4 Differential synaptic associations of Leigh syndrome genes in the enriched UKBEC modules, as observed using the SynGO database. Visualisations of the enrichment of SynGO ontology terms for synaptic location (pre- or post-synaptic) within the Leigh syndrome linked UKBEC modules (TCTX magenta and green, HIPP steelblue, WHMT grey60, SNIG yellow, and MEDU yellow). This diagram shows that only some of the UKBEC modules demonstrating enrichment of Leigh syndrome candidate genes are associated with either presynaptic or postsynaptic structures, or both. The colour coding is based on enrichment based on *q*-values (−log*p*-values), and concentric rings indicate gene function hierarchy at the synapse, with more general terms centrally and specific terms peripherally. HIPP, hippocampus; MEDU, medulla; SNIG, substantia nigra; TCTX, temporal cortex; WHMT, white matter. See Table S5 for details on the colour coding of the brain modules

unable to determine whether the remaining modules with enrichment of Leigh syndrome associated genes exhibited pre- or post-synaptic enrichment, owing to the smaller number of differentially expressed genes in these modules.

3.4 | Co-enrichment of genes associated with Leigh syndrome, epilepsy, Parkinson disease, and schizophrenia

Identification of UKBEC modules in which Leigh syndrome associated genes were enriched allowed us to examine the genetic mechanism of Leigh syndrome compared to other neurological disorders. Owing to some similar clinical features, for example, dystonia and seizures, we compared Leigh syndrome to Parkinson disease, epilepsy, and schizophrenia (Figure 5). Modules with significant enrichment of Leigh syndrome associated genes (−log*p*-value [FDR corrected] > 1.301, or FDR adjusted *p*-value <0.05) were plotted (Figure 5A). Parkinson disease-causing genes were found to be enriched in substantia nigra yellow (FDR adjusted $p = 0.0271$) and white matter grey60 (FDR adjusted $p = 0.049$) modules (Figure 5B). Epilepsy-associated

FIGURE 3 Plot of level 2 cell type specificity values from Karolinska mouse single-cell RNA-Seq superset. Specificity values for Leigh syndrome associated genes within level 2 cell subtypes, CA1 type 2 pyramidal (A), CA2 type 2 pyramidal (B), Interneuron 8 (C), and S1 deep layer pyramidal (D), were obtained from Skene et al., and were generated by dividing the average expression of a gene in a particular cell type by the average expression of the same gene in all other cell types

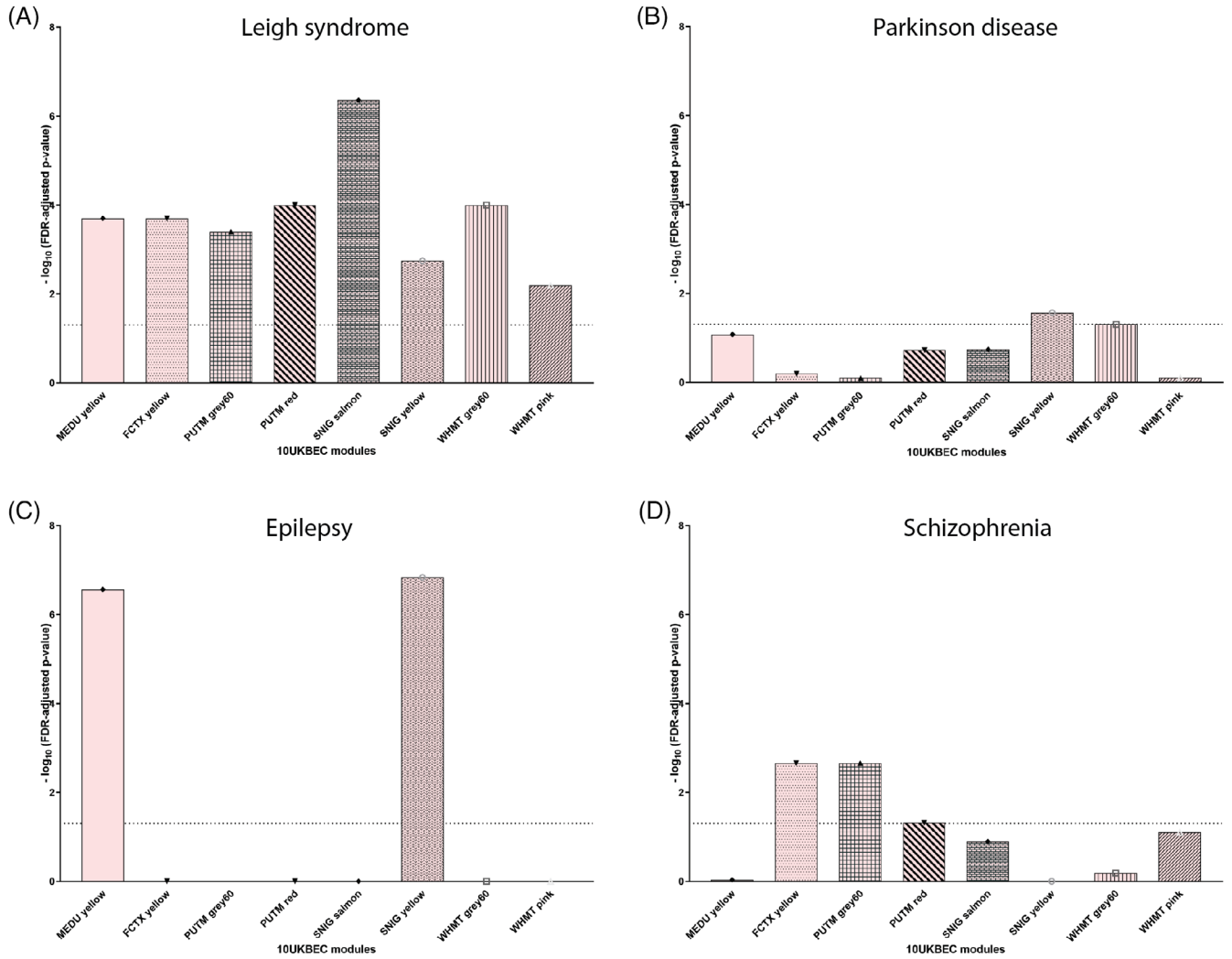

FIGURE 5 Disease heritability enrichment analysis of Parkinson disease, epilepsy, and schizophrenia candidate genes in Leigh syndrome enriched UKBEC modules. The *y*-axis is the $-\log_{10}$ adjusted *p*-values. The dashed line marks the approximate FDR threshold cut-off for nominal enrichment (~1.301). This figure shows that the medulla 'yellow' and substantia nigra 'yellow' modules are both enriched for Leigh syndrome (A) and epilepsy (C) genes, suggesting a possible degree of relatedness in their mechanisms in these two brain regions. The substantia nigra 'yellow' module is also enriched for Parkinson disease (B) linked genes. Furthermore, frontal cortex yellow, putamen grey60 and red were enriched in schizophrenia (D)

genes are highly enriched within the 10UKBEC medulla yellow (FDR adjusted $p = 1.4591\,e{-}7$) and substantia nigra yellow (FDR adjusted $p = 6.8265\,e{-}14$) modules (Figure 5C). Genes involved in schizophrenia were found to be significantly enriched in the 10UKBEC frontal cortex yellow (FDR adjusted $p = 0.0022$) and putamen grey60 (FDR adjusted $p = 0.0022$) and red (FDR adjusted $p = 0.0482$) modules (Figure 5D).

4 | DISCUSSION

Leigh syndrome is a rare, frequently rapidly progressing neurodegenerative disorder that generally affects children and young infants and often leads to early death. Our current study sought to determine the cell-specific expression of Leigh syndrome associated genes in the brain. Here we demonstrate that Leigh syndrome associated genes are enriched in the putamen and substantia nigra. Cell-specific analysis of mouse brain transcriptomics revealed significant enrichment in interneurons and hippocampal pyramidal cells. The genes were found to be of relatively similar specificity, with a number of the same genes in the regions with the highest specificity, particularly complex I and IV subunits, the complex I assembly factor *NDUFAF5,* and genes needed for mtDNA maintenance (especially *SUCLA2*) and mitochondrial gene expression (*GTPBP3*, *MTFMT*, and *MRPS34*). Further, we found that Leigh syndrome associated genes were associated with pre-synaptic structures. Finally,

heritability analysis of the CoExp modules revealed a similar enrichment profile in some CoExp modules for epilepsy-associated genes and genes associated with Leigh syndrome.

Although Leigh syndrome presents with considerable clinical, biochemical, and genetic heterogeneity, the neuropathological features are consistent. Patients exhibit neurodevelopmental delays and progressive decline of the central nervous system with focal bilateral lesions primarily involving the basal ganglia.[5] Lesions in the thalamus, brainstem, spinal cord, and cerebellum may also be present.[18] Of the genes associated with Leigh syndrome represented within the UKBEC network, we found that almost two thirds were enriched within the substantia nigra. This is consistent with clinical presentation of children with Leigh syndrome where necrotising lesions are observed within the basal ganglia, brainstem, and thalamus.[19] CoExp analyses revealed enrichment of Leigh syndrome associated genes in all 10 brain regions within the UKBEC network: cerebellum, frontal cortex, hippocampus, medulla, occipital cortex, putamen, substantia nigra, temporal cortex, thalamus, and white matter. The modules with enrichment of the largest number of genes associated with Leigh syndrome included modules of the medulla, substantia nigra, and putamen. Indeed, pathological lesions in these regions, namely putamen and substantia nigra, are commonly identified on MRI scans in patients with Leigh syndrome and are the most affected areas on neuropathological analysis.[18]

EWCE analysis revealed significant enrichment of Leigh syndrome associated genes in interneurons, hippocampal pyramidal CA1 neurons, and SS cells of the somatosensory cortex. Subtype analysis revealed significant enrichment within type 2 pyramidal neurons of the CA1 and CA2 regions of the brain, interneuron type 8, and the deep layer pyramidal neurons of the primary somatosensory cortex. The HIPP possesses a tri-synaptic circuitry involving neurons of the dentate gyrus, CA3, and CA1 areas. CA1 pyramidal cells of the HIPP represent a key output node of hippocampal memory circuits, with widespread projections.[20] Specifically, type 2 CA1 pyramidal cells have been found to associate with mitochondrial function and have also previously been correlated with the rate of firing and length of projections of cortical neurons.[11,21] Further, a close relationship between oxygen consumption and gamma wave oscillations within the HIPP was suggested as there was correlation between spatial expression of complex I subunits and gamma oscillations in mice and rats.[21] These oscillations are thought to function in sensory processing and memory formation.[22] Late onset Leigh syndrome presents with memory loss symptoms, which is consistent with our findings of enrichment of Leigh syndrome associated genes within CA1 and CA2 hippocampal neurons. Loss of mitochondrial function may result in inadequate or loss of fast neuronal network gamma oscillations, which occur in hippocampal pyramidal neurons, and thus impaired memory formation. The basal ganglia, THAL, cerebellum, brainstem, and spinal cord are commonly cited as regions of degeneration observed in Leigh syndrome. However, abnormalities in the HIPP, in addition to the corpus callosum, cortex, and cerebellar vermis, were identified in a case report of a child with Leigh syndrome as a result of a compound heterozygous mutation in *LRPPRC*.[23] Abnormal findings in the hippocampal tails were also observed in a patient with mutations in *NDUFAF8*.[24] Thus, our observation of enrichment in the pyramidal cells of the HIPP corroborates the findings of these case reports and suggests a role for Leigh syndrome associated genes in the HIPP.

The somatosensory cortex receives and processes afferent somatosensory input and contributes to the integration of both somatic and proprioceptive input required for skilled movement. Therefore, mutations in genes enriched in SS pyramidal cells of the somatosensory cortex could be contributing to the ataxia, dystonia, and movement disorders observed in patients with Leigh syndrome.

We did not find any enrichment in glial cells of Leigh syndrome associated genes in most of the functional groups analysed. Some level of expression within astrocytes and oligodendrocytes was observed in genes whose products are involved in mitochondrial toxicity. Lesions observed in Leigh syndrome are characterised by necrosis, gliosis, demyelination, spongiosis, and capillary proliferation.[19] The lack of enrichment of Leigh syndrome associated genes in glial cells may be because glycolysis is the main source of energy in astrocytes as opposed to oxidative metabolism in neurons.[25,26] Further, due to impaired mitochondrial functioning and thus neuronal damage and inflammation, reactive gliosis occurs. Romero-Morales et al.[27] have demonstrated impaired corticogenesis in cerebral organoids generated from induced pluripotent stem cells derived from Leigh syndrome patient fibroblasts. Organoids harbouring mutations in *DLD*, PDH, or an *MT-ATP6*/PDH double mutant had increased astrocyte marker levels at both the mRNA and protein level. Previous genetic fate mapping has suggested that mature astroglial cells can de-differentiate and resume proliferation.[28,29] Thus, in Leigh syndrome, the upregulation of astrocytes can be due to a reactive process in response to neuronal damage or may be a result of increased cells differentiating down the astroglial lineage.

Functional grouping of Leigh syndrome associated genes revealed that mitochondrial toxicity and vitamin

transport metabolism and co-factor biosynthesis demonstrated an expression pattern different from that of all other functional groups. Furthermore, genes associated with mitochondrial toxicity were expressed in both subtypes of astrocytes, those distributed throughout the cortex and those closely associated with the pia.[11] The genes within the mitochondrial toxicity functional grouping include *ETHE1* and *SQOR,* which are involved in sulphide metabolism, *ECHS1* and *HIBCH,* which are associated with valine degradation, and finally, *NAXE* and *SLC39A8,* which provide products needed for mitochondrial detoxification. Recent research has indicated that mitochondria are key in regulating astrocyte functions such as, calcium (Ca^{2+}) signalling, fatty-acid metabolism, glutamate-glutamine cycling, antioxidant production, and neuroinflammatory activation.[30,31] These observations lead to the hypothesis that in Parkinson disease, mitochondrial dysfunction, through a gain-of-function mechanism, results in an overactive inflammatory function of astrocytes, thereby negatively altering dopaminergic neuronal health.

Enrichment of Leigh syndrome associated genes in the interneurons, pyramidal CA1 and SS cells, was found to be similar across most of the genes of interest. It is worth noting that genes encoding OXPHOS enzymes or assembly factors made up half of the genes in the top 10 genes in both pyramidal cell types and three of the interneuron cell types, suggesting that Leigh syndrome is related to a primary mitochondrial dysfunction. Furthermore, *SLC19A3* and *SLC39A8* were the least enriched in these three cell types. Interestingly, these two genes encode non-mitochondrial proteins, further suggesting that there is an underlying primary mitochondrial dysfunction in neuronal cells in patients with Leigh syndrome. This could be suggestive of a yet undiscovered mechanism by which *SLC19A3* and *SLC39A8* cause mitochondrial dysfunction or that there is a different pathomechanism by which *SLC19A3* and *SLC39A8* mutations result in a Leigh-like syndrome as opposed to Leigh syndrome caused by a primary mitochondrial dysfunction. Given that Leigh syndrome presents on a spectrum with a wide range of disorders and the underlying pathology is due to aberrant mitochondrial function caused by mutations in nuclear or mitochondrial encoded proteins, the finding that no single gene drives the enrichment is as expected. Furthermore, given that the genes encode intracellular proteins and subunits of mitochondrial proteins, which are ubiquitously expressed in all human tissues with the exception of mature erythrocytes, such a finding is not unexpected.

The Karolinska Institute dataset utilised for the EWCE analysis comprises mouse brain data, not human. Despite that, our EWCE results are important since Leigh syndrome has been demonstrated in mouse models.[32,33] Indeed, these mouse models are widely utilised to investigate the disease mechanism of Leigh syndrome, as well as to develop novel therapies for Leigh syndrome.[34,35] The transcriptomic nature of our study allows us to infer the pathology of Leigh syndrome based on the assumption that aberrant gene functioning underlies the pathology of the disease. As discussed above, mutations in genes, either through a loss-of-function and decreased expression or a gain-of-function, result in altered mitochondrial activity throughout the brain and thus the variable clinical features observed in Leigh syndrome. It is difficult to determine whether it is solely mitochondrial dysfunction due to mutations or pathogenic gain of function mutations that contribute to the features observed in Leigh. We suggest it may be a combination of both, however, our data only provide support for the former theory, and future studies could elucidate this further.

Neurons are energy demanding cells that require significant amounts of ATP to generate action potentials, and thus normal mitochondrial function is essential for this purpose. Multiple neurodegenerative pathways, including those implicated in Alzheimer,[36] Parkinson,[37,38] and Huntington diseases,[39] are thought to converge on energy failure and thus result in early degeneration of the synaptic terminal, or the pre-synaptic terminal.[40] Canonically, mitochondria are thought to function to provide ATP to maintain the electrochemical gradient and recycling of synaptic vesicles as well as a Ca^{2+} buffer to allow for tight spatial and temporal control of neurotransmitter release.[41] Furthermore, synaptic development and subsequent pruning and plasticity is key for normal functioning of neuronal circuitry. One study found that *NDUFS4* silencing led to significantly decreased expression of synaptophysin, a pre-synaptic protein. This was further found to lead to shorter neurite lengths and suggested a role for NDUFS4-modulated mitochondrial activity in neuroplasticity.[42] An important factor for adequate development and remodelling of synapses is protein synthesis, another energy demanding process dependent on mitochondrial function. Further, the anti-apoptotic factor BCL-X_L (B-cell lymphoma extra-large) has been shown to drive presynaptic maturation via increased activity of dynamin related protein 1 (encoded by *DNM1L*), a protein needed for mitochondrial fission and thereby mitochondrial activity in the presynaptic terminal.[43,44] Both dominant and recessive mutations in *DNM1L* have been linked to Leigh syndrome. Our analyses revealed enrichment of *DNM1L* in the yellow module of the substntia nigra, which was found to have significant association with pre-synaptic structures. Previous reports have demonstrated defective bioenergetics[45,46] in cells derived from patients with

Leigh syndrome and more recently, impaired Ca^{2+} homeostasis,[47] illustrating the importance of mitochondria at neuronal synapses. Despite the above, it is well documented that mitochondria are not present at all pre-synaptic terminals,[48,49] and thus it remains unclear whether mitochondria are essential at synapses. We saw enrichment of genes associated with pre-synaptic function in the white matter grey60 module, which is interesting as white matter is primarily composed of axons of neurons. However, literature has demonstrated the presence of neurons in neo-cortical WHMT. The role of these neurons is unclear, but roles in the pathophysiology of schizophrenia and Alzheimer's disease have been proposed.[50] Our findings show that certain CoExp modules are significantly associated with pre-synaptic structures while others are not, suggesting that perhaps the expression pattern mirrors that of mitochondrial distribution within neurons. A limitation of this study in determining whether the remaining modules are associated with pre- or post-synaptic structures is the small number of genes associated with these modules.

The inheritance pattern of Leigh syndrome can be either maternally inherited in the case of defective mitochondrial genes or autosomal recessive, X-linked, or dominant if the defects are in nuclear-encoded genes. Epilepsy associated genes were found to be enriched in the yellow CoExp modules of the SNIG and medulla, modules in which Leigh syndrome genes are also enriched. This correlates with previous studies showing the comorbidity of epilepsy with neurometabolic disorders, including Leigh syndrome. A recent systematic review and meta-analysis of 385 patients with Leigh syndrome reported in five studies demonstrated that epilepsy occurs in a third of patients.[51] Previous literature has also consistently cited epilepsy as one of the clinical features present that may be present in patients with Leigh syndrome, suggesting that there may be a common underlying neuropathological mechanism between epilepsy and Leigh syndrome.[52]

Dystonia, spasticity, ataxia, and tremor are neurological signs that may be present in Leigh syndrome and may also be observed in Parkinson disease. For this reason, we tested whether the genetic basis of Parkinson disease and Leigh syndrome may be linked. Enrichment in the yellow module of the substantia nigra and grey60 module of white matter was observed for Parkinson disease, both regions that were also enriched for Leigh syndrome associated genes. It is well known that Parkinson disease is due to a loss of dopaminergic neurons in the pars compacta of the substantia nigra,[53,54] and many patients with Leigh syndrome have lesions involving the basal ganglia,[18] suggesting there may be a common underlying pathophysiology between the parkinsonism observed in some patients with Leigh syndrome and Parkinson disease. Despite this, no definitive heritability pattern was elucidated, suggesting that the similar clinical features observed may be due to different pathological mechanisms resulting in the same phenotype, or that genes that are not well represented in the studied networks may confer a different inheritance pattern and module enrichment that may reveal further insights. One case with adult-onset Leigh syndrome and a mutation in *MTFMT* had atypical parkinsonism.[55] The authors suggested that the dopaminergic neurons in patients with Leigh syndrome may be more susceptible to parkinsonism and Parkinson disease due to impaired OXPHOS, leading to extrapyramidal symptoms. Currently, little is known about dopamine transport and signalling in patients with Leigh syndrome, and thus future research is warranted to elucidate potential links between Parkinson disease and Leigh syndrome.

Finally, psychiatric symptoms can be present in patients with Leigh syndrome and there is some evidence suggesting that mitochondrial dysfunction is key in the pathophysiology of schizophrenia.[56] For these reasons, we examined the enrichment pattern of schizophrenia associated genes. We found significant enrichment in frontal cortex yellow and putamen grey60 and red for schizophrenia associated genes, modules in which significant enrichment of Leigh syndrome associated genes was also identified. Across the three disorders we compared modules demonstrating enrichment of disease associated genes, it can be noted that the findings for Parkinson disease, epilepsy, and schizophrenia, when combined, do result in a pattern similar to that observed for Leigh syndrome. This suggests that the underlying pathology of Leigh syndrome may perhaps be predisposing affected patients to develop parkinsonian phenotypes, psychiatric disorders, or seizures.

4.1 | Limitations

To our knowledge, this is the first study examining nervous tissue cell-specific expression analysis of the genes associated with Leigh syndrome. Despite the novelty of our study, it possesses limitations. First, our study is limited by the availability of high-quality, cell-specific brain transcriptomic data. The UKBEC CoExp networks utilised in our analyses did not possess data on all of the Leigh syndrome associated genes within all tissues of interest. More specifically, none of the 16 mitochondrially encoded genes linked to Leigh syndrome were present within the data set, and thus our synaptic enrichment and heritability study data are incomplete as the level of enrichment may be an underestimation. For this reason, we were also

unable to detect these genes within the specificity analysis. We also acknowledge that the UKBEC data set is an adult human brain data set, while Leigh syndrome is primarily paediatric in onset. We are limited in that there are no publicly available paediatric brain data sets and thus perhaps the findings of our study may be underestimated. Second, for our EWCE analysis we utilised mouse brain data for its completeness, however once again not all of the Leigh syndrome associated genes were present within the data set, particularly none of the mitochondrially encoded genes. We believe it is for this reason we were unable to identify any significant associations for mitochondrially encoded genes in the functional analysis. We were able to demonstrate enrichment in neuronal cells as opposed to astroglial cells, however, we do appreciate there may be species differences. We were unable to determine whether the spatial expression within brain cells was related to the mode of inheritance of genes associated with Leigh syndrome. As previously mentioned, this was due to poor representation of the mitochondrially encoded genes within the UKBEC network as well as within the mouse single-cell RNA-Seq data set used for EWCE. Given this, our findings could potentially underestimate the underlying mitochondrial aetiology of Leigh syndrome. Third, given that Leigh syndrome is genetically heterogeneous, with mutations in many different genes encoding very different products having been implicated, it is likely that there are other mutations that cause Leigh syndrome that have yet to be identified. Thus, the findings of our study may change over time as new Leigh syndrome causing gene defects are identified.

5 | CONCLUSIONS

Taken together, our study enabled us to identify specific cells and brain regions that demonstrate enrichment of genes that are associated with Leigh syndrome. We demonstrated that genes associated with the development of Leigh syndrome are expressed within brain regions most affected in patients with Leigh syndrome, namely, the putamen, substantia nigra, thalamus, and medulla. Transcriptomic analyses revealed significant enrichment of Leigh syndrome associated genes in neurons, particularly interneurons and pyramidal cells of the HIPP and somatosensory cortex. More specifically, genes with products involved in mitochondrial toxicity and vitamin transport metabolism and cofactor biosynthesis revealed a different expression pattern with enrichment in astroglial cells as opposed to neurons for the other functional groups. Taking the human and mouse brain RNA-Seq data together, our findings are suggestive of altered, more specifically, decreased neuronal expression of genes associated with Leigh syndrome as the underlying pathology for Leigh syndrome. We also show that these genes are preferentially associated with presynaptic structures. Finally, we show that Leigh syndrome associated genes share some enrichment modules with Parkinson disease, epilepsy, and schizophrenia.

ACKNOWLEDGEMENTS

The authors thank Professor Mina Ryten for helpful discussions. All research at Great Ormond Street Hospital NHS Foundation Trust and UCL Great Ormond Street Institute of Child Health is made possible by the NIHR Great Ormond Street Hospital Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

FUNDING INFORMATION

Chandika Soondram is the recipient of a Laidlaw Scholarship. Shamima Rahman acknowledges grant funding from Great Ormond Street Hospital Children's Charity, the Lily Foundation, and the National Institute of Health Research (NIHR) Great Ormond Street Hospital Biomedical Research Centre.

CONFLICT OF INTEREST

The authors declare no conflicts of interest, except that Shamima Rahman is Editor in Chief of the Journal of Inherited Metabolic Disease and JIMD Reports.

ORCID

Azizia Wahedi https://orcid.org/0000-0001-9329-9432
Alan E. Murphy https://orcid.org/0000-0002-2487-8753
Nathan Skene https://orcid.org/0000-0002-6807-3180
Shamima Rahman https://orcid.org/0000-0003-2088-730X

REFERENCES

1. Rahman S, Blok RB, Dahl HHM, et al. Leigh syndrome: clinical features and biochemical and DNA abnormalities. *Ann Neurol*. 1996;39(3):343-351. doi:10.1002/ANA.410390311
2. Darin N, Oldfors A, Moslemi AR, Holme E, Tulinius M. The incidence of mitochondrial encephalomyopathies in childhood: clinical features and morphological, biochemical, and DNA abnormalities. *Ann Neurol*. 2001;49(3):377-383. doi:10.1002/ana.75
3. Sofou K, De Coo IFM, Isohanni P, et al. A multicenter study on Leigh syndrome: disease course and predictors of survival. *Orphanet J Rare Dis*. 2014;9(1):52. doi:10.1186/1750-1172-9-52
4. Alves CAPF, Teixeira SR, Martin-Saavedra JS, et al. Pediatric Leigh syndrome: neuroimaging features and genetic correlations. *Ann Neurol*. 2020;88(2):218-232. doi:10.1002/ana.25789
5. Lake NJ, Compton AG, Rahman S, Thorburn DR. Leigh syndrome: one disorder, more than 75 monogenic causes. *Ann Neurol*. 2016;79(2):190-203. doi:10.1002/ana.24551

6. Rahman J, Rahman S. Mitochondrial medicine in the omics era. *Lancet*. 2018;391(10139):2560-2574. doi:10.1016/S0140-6736(18)30727-X
7. García-Ruiz S, Gil-Martínez AL, Cisterna A, et al. CoExp: a web tool for the exploitation of co-expression networks. *Front Genet*. 2021;12:630187. doi:10.3389/FGENE.2021.630187
8. Trabzuni D, Ryten M, Walker R, et al. Quality control parameters on a large dataset of regionally dissected human control brains for whole genome expression studies. *J Neurochem*. 2011;119(2):275-282. doi:10.1111/j.1471-4159.2011.07432.x
9. Mencacci NE, Reynolds R, Ruiz SG, et al. Dystonia genes functionally converge in specific neurons and share neurobiology with psychiatric disorders. *Brain*. 2020;143(9):2771-2787. doi:10.1093/BRAIN/AWAA217
10. Skene NG, Grant SGN. Identification of vulnerable cell types in major brain disorders using single cell transcriptomes and expression weighted cell type enrichment. *Front Neurosci*. 2016;10:16. doi:10.3389/FNINS.2016.00016/BIBTEX
11. Zeisel A, Moz-Manchado AB, Codeluppi S, et al. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science*. 2015;347(6226):1138-1142. doi:10.1126/SCIENCE.AAA1934
12. Skene NG, Bryois J, Bakken TE, et al. Genetic identification of brain cell types underlying schizophrenia. *Nat Genet*. 2018;50(6):825-833. doi:10.1038/S41588-018-0129-5
13. Koopmans F, van Nierop P, Andres-Alonso M, et al. SynGO: an evidence-based, expert-curated knowledge base for the synapse. *Neuron*. 2019;103(2):217-234.e4. doi:10.1016/j.neuron.2019.05.002
14. Ferrari R, Kia DA, Tomkins JE, et al. Stratification of candidate genes for Parkinson's disease using weighted protein-protein interaction network analysis. *BMC Genomics*. 2018;19(1):452. doi:10.1186/s12864-018-4804-9
15. Wang J, Lin ZJ, Liu L, et al. Epilepsy-associated genes. *Seizure*. 2017;44:11-20. doi:10.1016/j.seizure.2016.11.030
16. Farrell MS, Werge T, Sklar P, et al. Evaluating historical candidate genes for schizophrenia. *Mol Psychiatry*. 2015;20(5):555-562. doi:10.1038/mp.2015.16
17. Szklarczyk D, Gable AL, Nastou KC, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res*. 2021;49:D605-D612. doi:10.1093/nar/gkaa1074
18. Cavanagh JB, Harding BN. Pathogenic factors underlying the lesions in Leigh's disease. Tissue responses to cellular energy deprivation and their clinico-pathological consequences. *Brain*. 1994;117(6):1357-1376. doi:10.1093/BRAIN/117.6.1357
19. Leigh D. Subacute necrotizing encephalomyelopathy in an infant. *J Neurol Neurosurg Psychiatry*. 1951;14(3):216-221. doi:10.1136/JNNP.14.3.216
20. Soltesz I, Losonczy A. CA1 pyramidal cell diversity enabling parallel information processing in the hippocampus. *Nat Neurosci*. 2018;21(4):484-493. doi:10.1038/s41593-018-0118-0
21. Kann O, Huchzermeyer C, Kovács R, Wirtz S, Schuelke M. Gamma oscillations in the hippocampus require high complex i gene expression and strong functional performance of mitochondria. *Brain*. 2011;134(2):345-358. doi:10.1093/brain/awq333
22. Colgin LL, Moser EI. Gamma oscillations in the hippocampus. *Phys Ther*. 2010;25(5):319-329. doi:10.1152/PHYSIOL.00021.2010/ASSET/IMAGES/LARGE/PHY0051000300002.JPEG
23. Piro E, Serra G, Antona V, et al. Novel LRPPRC compound heterozygous mutation in a child with early-onset Leigh syndrome French-Canadian type: case report of an Italian patient. *Ital J Pedatr*. 2020;46(1):1-7. doi:10.1186/s13052-020-00903-7
24. Alston CL, Veling MT, Heidler J, et al. Pathogenic bi-allelic mutations in NDUFAF8 cause Leigh syndrome with an isolated complex I deficiency. *Am J Hum Genet*. 2020;106(1):92-101. doi:10.1016/j.ajhg.2019.12.001
25. Bélanger M, Allaman I, Magistretti PJ. Brain energy metabolism: focus on astrocyte-neuron metabolic cooperation. *Cell Metab*. 2011;14(6):724-738. doi:10.1016/J.CMET.2011.08.016
26. Rose J, Brian C, Pappa A, Panayiotidis MI, Franco R. Mitochondrial metabolism in astrocytes regulates brain bioenergetics, neurotransmission and redox balance. *Front Neurosci*. 2020;14:1155. doi:10.3389/FNINS.2020.536682/BIBTEX
27. Romero-Morales A, Rastogi A, Temuri H, et al. Human iPSC-derived cerebral organoids model features of Leigh syndrome and reveal abnormal corticogenesis. *Development*. 2022;149(20):dev199914. doi:10.2139/ssrn.3611282
28. Robel S, Berninger B, Götz M. The stem cell potential of glia: lessons from reactive gliosis. *Nat Rev Neurosci*. 2011;12(2):88-104. doi:10.1038/nrn2978
29. Robel S, Mori T, Zoubaa S, et al. Conditional deletion of β1-integrin in astroglia causes partial reactive gliosis. *Glia*. 2009;57(15):1630-1647. doi:10.1002/glia.20876
30. Bantle CM, Hirst WD, Weihofen A, Shlevkov E. Mitochondrial dysfunction in astrocytes: a role in Parkinson's disease. *Front Cell Dev Biol*. 2021;8:608026. doi:10.3389/fcell.2020.608026
31. Russo I, Kaganovich A, Ding J, et al. Transcriptome analysis of LRRK2 knock-out microglia cells reveals alterations of inflammatory- and oxidative stress-related pathways upon treatment with α-synuclein fibrils. *Neurobiol Dis*. 2019;67:129-178. doi:10.1016/j.nbd.2019.05.012
32. Kruse SE, Watt WC, Marcinek DJ, Kapur RP, Schenkman KA, Palmiter RD. Mice with mitochondrial complex I deficiency develop a fatal encephalomyopathy. *Cell Metab*. 2008;7(4):312-320. doi:10.1016/j.cmet.2008.02.004
33. Quintana A, Kruse SE, Kapur RP, Sanz E, Palmiter RD. Complex I deficiency due to loss of Ndufs4 in the brain results in progressive encephalopathy resembling Leigh syndrome. *Proc Natl Acad Sci USA*. 2010;107(24):10996-11001. doi:10.1073/pnas.1006214107
34. Johnson SC, Yanos ME, Kayser EB, et al. mTOR inhibition alleviates mitochondrial disease in a mouse model of Leigh syndrome. *Science*. 2013;342(6165):1524-1528. doi:10.1126/science.1244360
35. Corrà S, Cerutti R, Balmaceda V, Viscomi C, Zeviani M. Double administration of self-complementary AAV9NDUFS4 prevents Leigh disease in Ndufs4−/− mice. *Brain*. 2022;145(10):3405-3414. doi:10.1093/brain/awac182
36. Scheff SW, Price DA, Schmitt FA, Dekosky ST, Mufson EJ. Synaptic alterations in CA1 in mild Alzheimer disease and mild cognitive impairment. *Neurology*. 2007;68(18):1501-1508. doi:10.1212/01.WNL.0000260698.46517.8F
37. Cheng H-C, Ulane CM, Burke RE. Clinical progression in Parkinson's disease and the neurobiology of axons. *Ann Neurol*. 2010;67:715-725. doi:10.1002/ana.21995
38. Chu Y, Morfini GA, Langhamer LB, He Y, Brady ST, Kordower JH. Alterations in axonal transport motor proteins in sporadic and experimental Parkinson's disease. *Brain*. 2012;135(7):2058-2073. doi:10.1093/BRAIN/AWS133

39. Li H, Li S-H, Yu Z-X, Shelbourne P, Li X-J. Huntingtin aggregate-associated axonal degeneration is an early pathological event in Huntington's disease mice. *J Neurosci*. 2001;21(21):8473-8481. doi:10.1523/JNEUROSCI.21-21-08473.2001
40. Pathak D, Berthet A, Nakamura K. Energy failure: does it contribute to neurodegeneration? *Ann Neurol*. 2013;74:506-516. doi:10.1002/ana.24014
41. Devine MJ, Kittler JT. Mitochondria at the neuronal presynapse in health and disease. *Nat Rev Neurosci*. 2018;19(2):63-80. doi:10.1038/nrn.2017.170
42. Shil SK, Kagawa Y, Umaru BA, et al. Ndufs4 ablation decreases synaptophysin expression in hippocampus. *Sci Rep*. 2021;11(1):10969. doi:10.1038/s41598-021-90127-4
43. Smirnova E, Shurland DL, Ryazantsev SN, van der Bliek AM. A human dynamin-related protein controls the distribution of mitochondria. *J Cell Biol*. 1998;143(2):351-358. doi:10.1083/JCB.143.2.351
44. Li H, Chen Y, Jones AF, et al. Bcl-xL induces Drp1-dependent synapse formation in cultured hippocampal neurons. *Proc Natl Acad Sci USA*. 2008;105(6):2169-2174. doi:10.1073/PNAS.0711647105
45. Ma H, Folmes CDL, Wu J, et al. Metabolic rescue in pluripotent cells from patients with mtDNA disease. *Nature*. 2015;524(7564):234-238. doi:10.1038/nature14546
46. Lorenz C, Lesimple P, Bukowiecki R, et al. Human iPSC-derived neural progenitors are an effective drug discovery model for neurological mtDNA disorders. *Cell Stem Cell*. 2017;20(5):659-674.e9. doi:10.1016/j.stem.2016.12.013
47. Galera-Monge T, Zurita-Díaz F, Canals I, et al. Mitochondrial dysfunction and calcium dysregulation in leigh syndrome induced pluripotent stem cell derived neurons. *Int J Mol Sci*. 2020;21(9):3191. doi:10.3390/ijms21093191
48. Kang JS, Tian JH, Pan PY, et al. Docking of axonal mitochondria by syntaphilin controls their mobility and affects short-term facilitation. *Cell*. 2008;132(1):137-148. doi:10.1016/J.CELL.2007.11.024
49. Obashi K, Okabe S. Regulation of mitochondrial dynamics and distribution by synapse position and neuronal activity in the axon. *Eur J Neurosci*. 2013;38(3):2350-2363. doi:10.1111/EJN.12263
50. Alix JJ, Domingues AM. White matter synapses: form, function, and dysfunction. *Neurology*. 2011;76(4):397-404. doi:10.1212/WNL.0b013e3182088273
51. Chang X, Wu Y, Zhou J, Meng H, Zhang W, Guo J. A meta-analysis and systematic review of Leigh syndrome: clinical manifestations, respiratory chain enzyme complex deficiency, and gene mutations. *Medicine*. 2020;99(5):e18634. doi:10.1097/MD.0000000000018634
52. Rahman S. Advances in the treatment of mitochondrial epilepsies. *Epilepsy Behav*. 2019;101:106546. doi:10.1016/j.yebeh.2019.106546
53. del Rey NLG, Quiroga-Varela A, Garbayo E, et al. Advances in Parkinson's disease: 200 years later. *Front Neuroanat*. 2018;12:113. doi:10.3389/FNANA.2018.00113/BIBTEX
54. Poewe W, Seppi K, Tanner CM, et al. Parkinson disease. *Nat Rev Dis Prim*. 2017;3(1):1-21. doi:10.1038/nrdp.2017.13
55. Hemelsoet DM, Vanlander AV, Smet J, et al. Leigh syndrome followed by parkinsonism in an adult with homozygous c.626C>T mutation in MTFMT. *Neurol Genet*. 2018;4((6)):e298. doi:10.1212/NXG.0000000000000298
56. Kung L, Roberts RC. Mitochondrial pathology in human schizophrenic striatum: a postmortem ultrastructural study. *Synapse*. 1999;31(1):67-75. doi:10.1002/(SICI)1098-2396(199901)31:1<67::AID-SYN9>3.0.CO;2-#

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Wahedi A, Soondram C, Murphy AE, Skene N, Rahman S. Transcriptomic analyses reveal neuronal specificity of Leigh syndrome associated genes. *J Inherit Metab Dis*. 2022;1-18. doi:10.1002/jimd.12578