Alpha-synuclein overexpression induces epigenomic dysregulation of glutamate signaling and locomotor pathways

Samantha L. Schaffner[1], Zinah Wassouf[2,3,†], Diana F. Lazaro[2], Mary Xylaki[2], Nicole Gladish[1], David T. S. Lin[1], Julia MacIsaac[1],

Katia Ramadori[1,‡], Thomas Hentrich[4], Julia M. Schulze-Hentrich[4], Tiago F. Outeiro[2,3,5,6,§,*] and Michael S. Kobor [1,§,*]

[1]Department of Medical Genetics, Centre for Molecular Medicine and Therapeutics, British Columbia Children's Hospital Research Institute, University of British Columbia, Vancouver, BC, V5Z 4H4, Canada
[2]Department of Experimental Neurodegeneration, Center for Biostructural Imaging of Neurodegeneration, University Medical Center Göttingen, 37073 Göttingen, Germany
[3]German Centre for Neurodegenerative Diseases (DZNE), 37075 Göttingen, Germany
[4]Institute of Medical Genetics and Applied Genomics, University of Tübingen, 72076 Tübingen, Germany
[5]Max Planck Institute for Natural Sciences, 37075 Göttingen, Germany
[6]Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Framlington Place, Newcastle-upon-Tyne, NE2 4HH, UK
*To whom correspondence should be addressed at: Dr. Michael S. Kobor, PhD, Email: msk@bcchr.ca; Phone: +1 604-875-3803; Fax: +1 604-875-3840, 950 W 28th Ave, Vancouver, BC, V5Z 4H4, Canada

[†]Present address: Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridgeshire, UK

[‡] Present address: BC Cancer Genetics and Genomics Laboratory, 600 W 10th Ave #3305, Vancouver, BC, Canada.

[§] These authors share equal senior authorship.

[‖] New sequence accessions: GSE181126 (Gene Expression Omnibus).

Abstract

Parkinson's disease (PD) is a neurological disorder with complex interindividual etiology that is becoming increasingly prevalent worldwide. Elevated alpha-synuclein levels can increase risk of PD and may influence epigenetic regulation of PD pathways. Here, we report genome-wide DNA methylation and hydroxymethylation alterations associated with overexpression of two PD-linked alpha-synuclein variants (wild-type and A30P) in LUHMES cells differentiated to dopaminergic neurons. Alpha-synuclein altered DNA methylation at thousands of CpGs and DNA hydroxymethylation at hundreds of CpGs in both genotypes, primarily in locomotor behavior and glutamate signaling pathway genes. In some cases, epigenetic changes were associated with transcription. SMITE network analysis incorporating H3K4me1 ChIP-seq to score DNA methylation and hydroxymethylation changes across promoters, enhancers, and gene bodies confirmed epigenetic and transcriptional deregulation of glutamate signaling modules in both genotypes. Our results identify distinct and shared impacts of alpha-synuclein variants on the epigenome, and associate alpha-synuclein with the epigenetic etiology of PD.

Introduction

Alpha-synuclein (aSyn) plays crucial roles in neurodevelopment, neuronal health, and synaptic transmission (1). Mutations, multiplications, and single nucleotide polymorphisms (SNPs) in *SNCA*, the gene encoding aSyn, are associated with Parkinson's disease (PD), an age-associated, and therefore increasingly prevalent, neurological disorder (2,3). *SNCA* variants have also been linked to several other neurodegenerative conditions, including Lewy body dementia (LBD), multiple system atrophy (MSA), and Alzheimer's disease (4,5). The full range of functions of aSyn are still unclear, as a number of different cellular roles have been reported, including SNARE complex assembly, regulation of neuronal differentiation, glucose level, and dopamine biosynthesis, as well

as modulation of calmodulin activity/G protein-coupled receptor kinase activity (1). aSyn point mutations can impair one or several of these processes, and thus disrupt neuronal health. Interestingly, we recently reported that nuclear localization and phosphorylation modulate the pathological effects of aSyn (6). However, the molecular mechanisms leading to PD may differ according to the variant. For example, the A30P mutation occurs in the membrane binding domain of aSyn, which may affect its ability to act as a presynaptic chaperone and interact with membrane-bound receptors (7). Although A30P aSyn appears less likely to aggregate than wild-type (WT) aSyn, this mutation can still result in PD, possibly due to a loss of aSyn function (7). Conversely, multiplications of the *SNCA* gene can lead to excess protein production,

promoting aggregations and fibrils that impair synaptic function and can lead to neuronal death (8). Duplications and triplications of *SNCA* have been reported in patients with familial PD, LBD, and MSA (2,4).

Despite progress in the understanding the relations between aSyn dysfunction and PD, much remains to be elucidated. For example, although aSyn- and PD-associated transcriptional deregulation have been demonstrated in cell culture, animal models, and human patients, the mechanisms underlying these transcriptional changes are not yet fully understood (9–16). As different aSyn mutations can lead to PD through different mechanisms, it is also important to understand which aspects of these variants are unique and which are shared. Studies by our group and others have shown that expression of WT or A30P mutant aSyn in human neurons is associated with DNA damage and transcriptional dysregulation of genes involved in cell survival and DNA repair, which may be mediated by altered histone H3 acetylation (9,10). In addition, altered DNA methylation (DNAm) patterns have been observed in both the *SNCA* gene itself and genome-wide in blood, saliva, and brain of PD patients (17–21). Taken together, these observations suggest that genome-wide transcriptional and epigenetic dysregulation may be consequences of aSyn overexpression and could potentially be involved in PD susceptibility and pathogenesis.

Epigenetic studies of PD provide the opportunity to assess genetic and environmental influences on disease risk concurrently (22–28). DNA methylation (DNAm), which refers to the attachment of a methyl group to DNA on the 5′ carbon of cytosine, frequently in the context of cytosine-guanine (CpG) dinucleotides, is a well-studied epigenetic mark that changes during development and is influenced by genes and environment (22–25,29,30). Although DNAm patterns can be unrelated to mRNA expression patterns or laid down as a consequence of gene expression, in some cases DNAm can also impact transcription by altering the ability of transcription factors to bind to gene regulatory regions, changing chromosomal interactions, and influencing splicing (22,29–31). This can alter the regulation of pathways involved in disease susceptibility (32–34). In addition, DNAm patterns can be useful as biomarkers of aging, disease states, or exposures regardless of association with transcription (35–41).

Although DNAm is influenced by a myriad of factors that differ between individuals, the existing literature on DNAm in PD has primarily adopted a case–control approach (17–21). In addition, the role of DNA hydroxymethylation (DNAhm), which accounts for up to 40% of all modified cytosines in brain tissue, remains unclear (42). DNAhm is an oxidized form of DNAm, and an intermediate of the DNA demethylation process; however, DNAhm has also been suggested to be a stable, independent epigenetic mark in the brain (43). The stability of DNAhm and its ability to influence transcription by eliminating DNAm–protein interactions, introducing DNAhm–protein interactions, altering

chromatin state, and altering splicing suggest that it could impact the etiology of neurological diseases (44,45). It is important to distinguish DNAhm from DNAm, as conventional bisulfite conversion techniques measure both DNAhm and DNAm in one compound signal, which may bias interpretations (46). DNAhm is stable in postmortem formalin-fixed tissue, and some initial studies characterized DNAhm patterns in the brains of deceased PD patients (47). Bulk DNAhm levels are unchanged in the cortex, substantia nigra, and brainstem of PD patients, while DNAhm levels and expression of TET2, the enzyme that forms this mark, are elevated in neurons of the prefrontal cortex of PD patients (48,49). Importantly, DNAm and DNAhm patterns can be influenced by genetics, and genetic alterations at loci, such as *SNCA*, influence PD susceptibility, suggesting that it will be important to incorporate genetic considerations into future epigenetic studies of PD (50–52). To date, only a few studies have assessed the impact of genetic background on DNAm patterns in PD (53–55).

Accounting for the range of genetic, environmental, and lifestyle factors that can influence DNAm and DNAhm are major challenges in human epigenetic studies of PD (22,23,29,50). Furthermore, the use of postmortem tissue to study PD also typically limits the analysis to cases of advanced disease. Model systems, such as rodents or cell culture, can help to address these issues, providing contexts in which both genetics and environment can be tightly controlled, and allowing the examination of early impacts of neurological disease in the brain.

In this study, we characterized the influence of two molecularly distinct aSyn variants, WT and A30P aSyn, on the DNA methylome and hydroxymethylome of human dopaminergic neurons. We profiled genome-wide DNAm and DNAhm patterns in three groups of Lund human mesencephalic (LUHMES) cells differentiated into dopaminergic neurons: control LUHMES cells, LUHMES cells overexpressing WT aSyn at levels similar to those seen in *SNCA* multiplication carriers (WT aSyn cells), and LUHMES cells overexpressing A30P aSyn (A30P aSyn cells). In addition, we integrated our DNAm and DNAhm data with transcriptomic data from the same cells, incorporating H3K4me1 ChIP-seq to score DNA(h)m changes across gene regulatory features (9). WT and A30P aSyn expression were associated with thousands of DNA(h)m changes, particularly affecting genes that regulate locomotor behavior and glutamate signaling pathways. These observations suggested that familial PD-associated aSyn mutations have widespread epigenomic effects, and may contribute to molecular and transcriptional dysregulation associated with the etiology and heterogeneity of PD.

Results
aSyn overexpression altered genome-wide DNAm patterns in dopaminergic neurons

We used differentiated LUHMES cells as a model system in which to evaluate the epigenome-wide impacts of
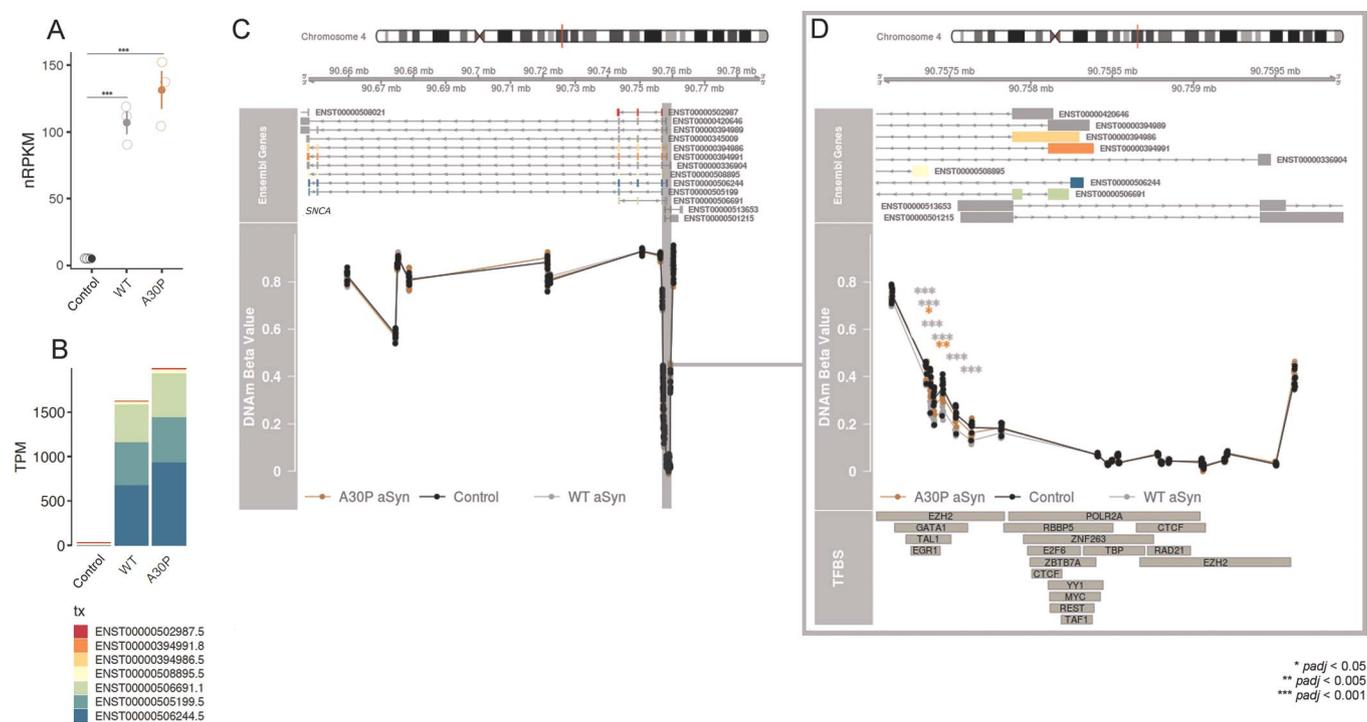
Figure 1. aSyn overexpression was associated with decreased DNAm at the *SNCA* 5′ UTR. (A) Gene-level *SNCA* mRNA expression in nRPKM. *p*-values: *DESeq*2 Wald test with Benjamini–Hochberg adjustment. WT aSyn vs. control and A30P aSyn vs. control *padj* < 2.2 × 10^{-308}. (B) Transcript-level *SNCA* mRNA expression in TPM. See Supplementary Material, Table S1 for statistics. (C) Top: hg19 coordinates are shown for the *SNCA* gene, with Ensembl transcripts below. Boxes represent exons, lines represent introns. Bottom: DNAm beta values are shown for each sample, colored by genotype. Only EPIC probes specific to endogenous *SNCA* are shown. Control cells: $n = 7$ biological replicates; WT aSyn cells: $n = 8$; A30P aSyn cells: $n = 8$. (D) hg19 coordinates, DNAm beta values, and transcription factor binding sites for the *SNCA* 5′ UTR/intron 1/TSS upstream region. *p*-values: *limma* empirical Bayes moderated *t* test with Benjamini–Hochberg adjustment. Bottom: ENCODE transcription factor ChIP-seq clusters.

aSyn expression in dopaminergic neurons. Genome-wide DNAm and DNAhm were analyzed in the same cells as used in a previously published transcriptomic analysis by our group (GSE89115, GSE181126) (9). Previously, we verified dopaminergic neuronal differentiation in these cells by immunostaining, which demonstrated the presence of tyrosine hydroxylase (*TH*)-positive axonal and dendritic networks, as well as endogenous aSyn expression (Supplementary Material, Fig. S1). Proliferating LUHMES cells were infected with lentiviral constructs containing IRES-GFP, WT aSyn-IRES-GFP, or A30P aSyn-IRES-GFP, and positive clones were selected by fluorescence activated cell sorting (FACS). As expected, the levels of *SNCA* mRNA and aSyn protein expression were higher in WT aSyn cells and A30P aSyn cells than in control cells, as determined by RNA-seq (approximately 4-fold in both genotypes) (Fig. 1A, B, Supplementary Material, Table S1) and immunoblotting (approximately 6-fold in both genotypes, see ref. 9). Elevated *SNCA* mRNA expression was primarily driven by three transcripts in WT and A30P aSyn cells (Fig. 1B, Supplementary Material, Table S1).

To investigate the impact of WT and A30P aSyn expression on genome-wide DNAm patterns, we assessed DNAm differences between control, WT aSyn, and A30P aSyn cells at 813589 EPIC probes. An effect size threshold of |delta beta| ≥ 0.05 was selected to be well above technical noise (maximum 2.2% RMSE between technical replicates), and probes were considered to

be statistically significant at a Benjamini–Hochberg adjusted *p* (*padj*) ≤ 0.05, corrected for multiple comparisons. First, we examined whether overexpression of aSyn was associated with reduced DNAm at the first intron of *SNCA*, as reported previously in PD patients (20,21). We expected DNAm to be lower at this region in aSyn-expressing LUHMES cells as intron 1 has been established as a putative promoter for *SNCA* containing GATA binding sites, and enzymatic removal of DNAm in this region is associated with increased levels of *SNCA* mRNA and protein expression (56,57). In addition, to specifically examine DNAm levels at the endogenous gene without confounding by CpGs also located on the transgene, we removed five CpGs mapping to the cDNA sequence used in the aSyn constructs (NM_000345) and the remaining 35 *SNCA* CpGs are shown in Fig. 1C, D. As expected, both aSyn genotypes had significantly lower levels of DNAm at several CpGs in the region corresponding to intron 1 for the majority of *SNCA* transcripts, and upstream of the transcription start site (TSS) for several other transcripts (Fig. 1C, D). This region covered binding sites for GATA1, EZH2, TAL1, and EGR1 (Fig. 1D).

We next assessed genome-wide DNAm changes induced by overexpression of each aSyn variant. First, we used a site-specific approach, with linear modeling applied to all CpGs passing quality control (QC) filters (Fig. 2, Table 1). In comparison to controls, WT aSyn cells had 18 521 sites with decreased DNAm and 10 812
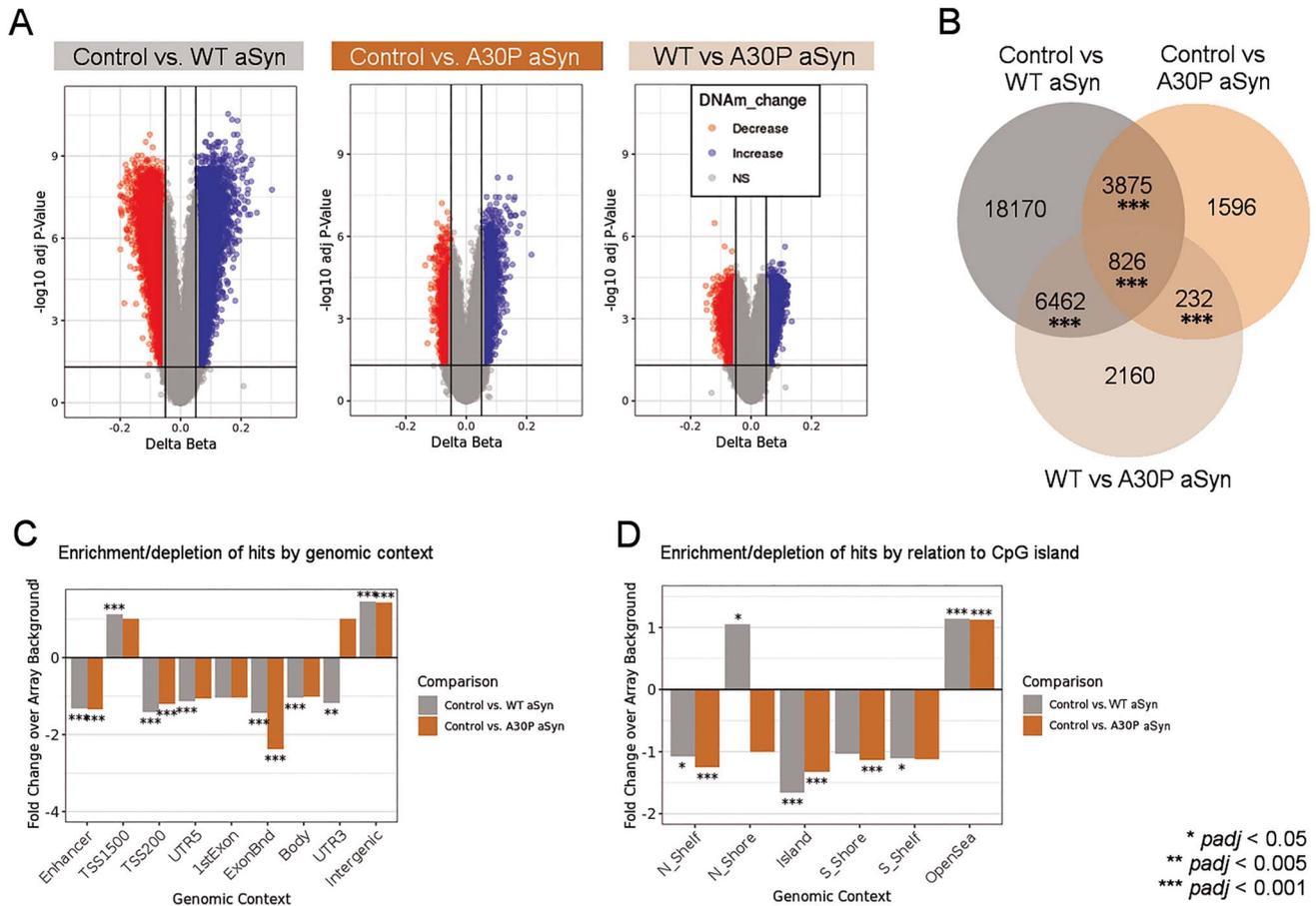
Figure 2. Overexpression of WT and A30P mutant aSyn altered the DNA methylome of dopaminergic neurons. (A) Volcano plots comparing DNAm patterns between control ($n = 7$ biological replicates) and WT aSyn ($n = 8$) LUHMES cells, control and A30P aSyn ($n = 8$) LUHMES cells, and WT aSyn and A30P aSyn LUHMES cells. Colored points passed thresholds of delta beta ≥ 0.05 and *padj* ≤ 0.05 (*limma* empirical Bayes moderated *t* test with Benjamini–Hochberg adjustment). (B) Number of differentially methylated probes unique to each comparison and probes shared between comparisons (*padj* < 0.001, 10 000 permutations). (C) Relative enrichment/depletion of differentially methylated sites across genomic contexts, permuted against array background (10 000 iterations). (D) Relative enrichment/depletion of differentially methylated sites by relation to CpG island, permuted against array background (10 000 iterations).

sites with increased DNAm (*padj* ≤ 0.05) (Fig. 2A, Table 1, Supplementary Material, Table S2). In A30P aSyn cells, 3091 probes had decreased DNAm and 3438 probes had increased DNAm compared to control cells (Fig. 2A, Table 1, Supplementary Material, Table S3). In comparison to WT aSyn cells, A30P aSyn cells had 3014 sites with decreased DNAm and 6666 sites with increased DNAm (Fig. 2A, Supplementary Material, Table S4). DNAm changes in all comparisons were significantly skewed in one direction, and the bias in effect direction was seen across all genomic features ($p < 2.2e^{-16}$) (Supplementary Material, Table S5). To confirm an example from our findings by pyrosequencing, we selected a region of the *TUBA8* gene that included the top two probes with the greatest changes in DNAm across both genotypes; differential DNAm at these two positions, one additional position on the array, and four positions not on the array were all confirmed (Table 1, Supplementary Material, Fig. S2, Supplementary Material, Table S5).

The proportion of differentially methylated probes shared across all comparisons and any two of three

comparisons was greater than expected by chance, and in most cases, the differences from controls were more pronounced for WT aSyn cells than for A30P aSyn cells (*padj* < 0.001, 10 000 permutations) (Fig. 2B, Table 1). Interestingly, WT and A30P aSyn overexpression primarily affected CpGs outside of genes and enhancers as defined by H3K4me1 ChIP-seq (Fig. 2C). However, differentially methylated probes were enriched in TSS1500 regions for both genotypes (statistically significant in WT only), and trends for differentially methylated probe feature enrichment differed between genotypes for 3' UTRs and CpG North Shores (Fig. 2C, D).

Following the investigation of differential DNAm patterns at individual CpG sites, we additionally examined genotype-dependent DNAm differences at comethylated regions (CMRs) (Supplementary Material, Fig. S3). CMRs are defined as regions with correlated DNAm across individuals, and such regional patterns of DNAm are more likely to be biologically meaningful than DNAm at single CpGs (58). We defined 57 941 custom DNAm CMRs in LUHMES neurons and applied

Table 1. Top 5 unique and shared differentially methylated probes for each comparison, ranked by delta beta

Control vs. WT aSyn unique					
Chr	chr5	chr4	chr7	chr4	chr1
Pos	42 994 123	151 462 241	44 614 305	151 503 546	57 884 286
Probe ID	cg14979301	cg09064570	cg05758978	cg22265605	cg01074356
WT aSyn adjusted p-value	$2.54e^{-06}$	$5.36e^{-07}$	0.000236254	$4.86e^{-07}$	0.000191117
WT aSyn delta beta	−0.1995039	0.19595166	−0.186602	0.1851907	0.18111904
A30P aSyn adjusted p-value	0.09078044	0.12926655	0.05169001	0.25104086	0.12524333
A30P aSyn delta beta	−0.0646483	0.0490894	−0.1084993	0.03763804	0.06857821
Gene		LRBA	DDX56	LRBA	DAB1
Genomic context		Body	TSS200	Body	5′ UTR
Relation to CpG island	N_Shore	OpenSea	S_Shore	N_Shore	N_Shelf

Control vs. A30P aSyn unique					
Chr	chr10	chr22	chr21	chr14	chr1
Pos	131 694 689	45 809 619	44 915 811	65 006 281	245 755 356
Probe ID	cg26275858	cg05072848	cg14177086	cg21035222	cg18199753
WT aSyn adjusted p-value	0.914887877	0.006470176	0.265419174	0.008640058	0.065585536
WT aSyn delta beta	0.00315511	0.04852953	−0.0369103	0.04451941	−0.0481363
A30P aSyn adjusted p-value	$7.93e^{-05}$	$1.91e^{-05}$	0.002633198	$2.86e^{-07}$	0.000666635
A30P aSyn delta beta	−0.1182818	0.11255114	−0.1123709	0.10928006	−0.1059641
Gene	EBF3	SMC1B		HSPA2	KIF26B
Genomic context	Body	TSS200		TSS1500	Body
Relation to CpG island	Island	Island	OpenSea	N_Shore	S_Shore

Control vs. WT aSyn and Control vs. A30P aSyn shared					
Chr	chr22	chr22	chr10	chr12	chr22
Pos	18 593 586	18 593 441	94 455 710	12 849 159	18 593 609
Probe ID	cg27391512	cg25835669	cg07856667	cg13417420	cg08715837
WT aSyn adjusted p-value	$1.69e^{-08}$	$1.38e^{-09}$	$2.08e^{-09}$	$1.17e^{-07}$	$2.38e^{-08}$
WT aSyn delta beta	0.30304448	0.23725831	0.23219774	0.25294376	0.2215025
A30P aSyn adjusted p-value	$4.62e^{-06}$	$1.44e^{-07}$	$7.46e^{-07}$	0.0005165	$1.06e^{-05}$
A30P aSyn delta beta	0.21622949	0.18841661	0.16977899	0.13573927	0.1532297
Average delta beta	0.25963699	0.21283746	0.20098836	0.19434151	0.1873661
Gene	TUBA8	TUBA8		GPR19	TUBA8
Genomic context	5′ UTR	TSS200		TSS200	5′ UTR
Relation to CpG island	Island	Island	Island	Island	Island

linear modeling to the composite beta values calculated for each CMR, representing a weighted average of CMR probe DNAm levels. Similar trends were seen in CMR analysis and site-specific analysis; the number of DNAm changes was greater in WT aSyn cells vs. control (3755 CMRs with $padj \leq 0.05$ and |delta composite beta| ≥ 0.05) than in either A30P aSyn cells vs. control (548 CMRs) or A30P aSyn cells vs. WT aSyn cells (765 CMRs) (Supplementary Material, Fig. S3A, B). Approximately 40% of the total differentially methylated probes discovered with the site-specific approach mapped to CMRs in any comparison, and overlaps between differentially methylated CMRs and individual CpGs were more likely than expected by chance ($padj < 0.001$, 10 000 permutations) (Supplementary Material, Fig. S3A). Trends for differential CMR methylation across gene features were also similar to the site-specific approach (Supplementary Material, Fig. S3C, D). A CMR covering the *SNCA* first intron region was included among the differentially methylated CMRs in the control vs. WT aSyn and control vs. A30P aSyn comparisons, passing the composite delta beta cutoff only in WT aSyn neurons (chr4: 90 757 139–90 757 814; control vs. WT aSyn

$padj = 9.96e^{-10}$, delta composite beta = −0.083; control vs. A30P aSyn $padj = 0.0029$, delta composite beta = −0.032).

aSyn overexpression altered genome-wide DNAhm patterns to a lesser extent than DNAm patterns

After assessing genome-wide DNAm alterations, we examined differential DNAhm at 233 440 sites that passed our detection threshold of 3.6% (calculated according to the 95% quantile of negative hydroxymethylation (hmC) values after subtraction; see Methods). Fewer changes were observed in DNAhm for each comparison compared to DNAm, most of which were increases; these increases in DNAhm were significant when tested for skewness ($p < 2.2e^{-16}$) (Fig. 3, Table 2, Supplementary Material, Tables S6, S7). Only one probe was differentially hydroxymethylated in comparison between WT aSyn and A30P aSyn cells (cg11833293 in the *SHANK2* 3′ UTR, $padj = 0.045$, delta beta = 0.089) (Fig. 3A, B). The overlaps of 16 CpG sites between control vs. WT aSyn and control vs. A30P aSyn were not greater than expected by chance (Fig. 3B, $padj > 0.05$, 10 000 permutations). Both genotypes also had fewer differen-
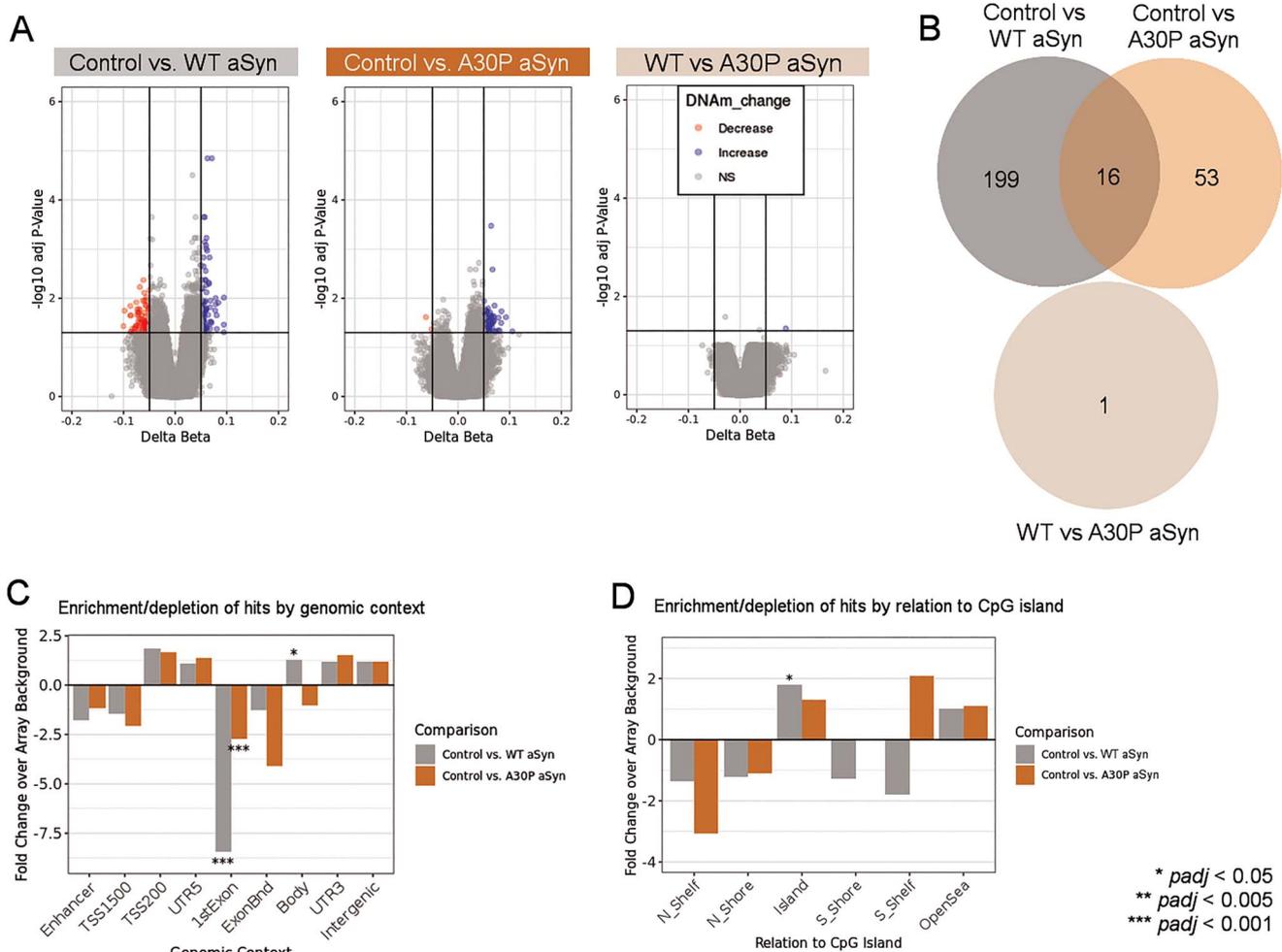
Figure 3. WT and A30P aSyn overexpression were correlated with increased DNAhm levels in dopaminergic neurons. (A) Volcano plots comparing DNAhm patterns between control ($n=7$ biological replicates) and WT aSyn ($n=8$) LUHMES cells, control and A30P aSyn ($n=8$) LUHMES cells, and WT aSyn and A30P aSyn LUHMES cells. Colored points passed thresholds of delta beta ≥ 0.05 and $padj \leq 0.05$ (*limma* empirical Bayes moderated *t* test with Benjamini–Hochberg adjustment). (B) Number of differentially hydroxymethylated probes unique to each comparison and probes shared between comparisons. All $padj = 1$ (10 000 permutations). (C) Relative enrichment/depletion of differentially hydroxymethylated sites across genomic contexts, permuted against array background (10 000 iterations). (D) Relative enrichment/depletion of differentially hydroxymethylated sites by relation to CpG island, permuted against array background (10 000 iterations).

tially hydroxymethylated probes in the first exon than expected by chance (Fig. 3C).

When defining CMRs in the DNAhm data and applying linear modeling, we also found greater increases in CMR DNAhm in control vs. WT aSyn and control vs. A30P aSyn comparisons (Supplementary Material, Fig. S4). Although fewer CMRs were found in DNAhm data overall, differentially hydroxymethylated CMRs were also more likely to overlap with differentially hydroxymethylated CpGs from site-specific analysis than expected by chance (Supplementary Material, Fig. S4A).

WT and A30P aSyn altered DNAm at locomotor and glutamate signaling pathway genes

To identify possible functional consequences associated with the DNAm and DNAhm changes in each genotype, we performed Gene Ontology (GO) enrichment analysis on all differentially methylated or hydroxymethylated genes within each comparison using over-representation analysis in ermineR (59). Thirty-four GO biological

processes were enriched in differentially methylated sites shared between the control vs. WT aSyn and control vs. A30P aSyn comparisons (Fig. 4A, $padj < 0.05$). A total of 442 CpG sites were annotated to the top GO term, 'locomotory behavior,' and differentially methylated in at least one genotype (420 sites in WT aSyn cells and 89 sites in A30P aSyn cells) (Fig. 4B–E). A total of 129 differentially methylated sites were annotated to the second highest ranking GO term, 'glutamate receptor signaling pathway' (127 sites in WT aSyn cells and 18 sites in A30P aSyn cells) (Supplementary Material, Fig. S5). No significant multifunctionality-corrected enrichments were observed for differentially methylated or hydroxymethylated sites in any of the other groups.

A subset of WT aSyn-induced DNAm and DNAhm changes were also observed in WT aSyn transgenic mice

To assess the relevance of our findings *in vivo*, the differential DNA(h)m results from control vs. WT aSyn

Table 2. Top 5 unique and shared differentially hydroxymethylated probes for each comparison, ranked by delta beta

Control vs. WT aSyn unique					
Chr	chr11	chr1	chr11	chr5	chr15
Pos	11 959 869	157 010 623	119 191 624	134 825 791	99 329 769
Probe ID	cg21029705	cg11184358	cg12762029	cg01889143	cg08307952
WT aSyn adjusted *p*-value	$3.71e^{-02}$	$1.80e^{-02}$	0.009721248	$4.94e^{-02}$	0.034644699
WT aSyn delta beta	−0.100144845	−0.098018504	0.0944902	0.094449407	0.09426821
A30P aSyn adjusted *p*-value	0.8044695	0.2341719	0.2673974	0.2917405	0.6017781
A30P aSyn delta beta	−0.0315162	−0.06167421	0.03536456	0.0533201	0.02102634
Gene	USP47	ARHGEF11	MCAM		IGF1R
Genomic context	Body	Body	3′ UTR		5′ UTR
Relation to CpG island	OpenSea	OpenSea	Island	Island	OpenSea
Control vs. A30P aSyn unique					
Chr	chr3	chr19	chr6	chr21	chr2
Pos	173 001 443	54 720 832	20 697 104	33 267 395	192 759 994
Probe ID	cg03417233	cg12555719	cg25475839	cg04815049	cg14971567
WT aSyn adjusted *p*-value	0.538342	0.2605252	0.1508893	0.16631	0.4370172
WT aSyn delta beta	0.0383568	0.03308958	0.0366181	0.0432397	0.01959213
A30P aSyn adjusted *p*-value	$4.75e^{-02}$	$2.44e^{-02}$	0.018507248	$4.64e^{-02}$	0.024865866
A30P aSyn delta beta	0.105273858	0.092858583	0.084182475	0.080982712	0.078224072
Gene	RP11-324C10.1	LILRB3	CDKAL1	HUNK	AC098617.2
Genomic context	5′ UTR	3′ UTR	Body	Body	5′ UTR
Relation to CpG island	OpenSea	OpenSea	OpenSea	OpenSea	OpenSea
Control vs. WT aSyn and Control vs. A30P aSyn shared					
Chr	chr12	chr16	chr7	chr8	chr7
Pos	56 121 485	24 834 981	1 281 200	108 510 324	158 863 255
Probe ID	cg22249612	cg14908168	cg02471897	cg16318522	cg00128168
WT aSyn adjusted *p*-value	$2.23e^{-02}$	$9.93e^{-03}$	$1.76e^{-02}$	$3.76e^{-02}$	$3.98e^{-02}$
WT aSyn delta beta	0.080308633	0.078234846	0.059501066	0.053512554	0.054876613
A30P aSyn adjusted *p*-value	$3.19e^{-02}$	$3.88e^{-02}$	$2.32e^{-02}$	0.02916036	$3.19e^{-02}$
A30P aSyn delta beta	0.083408157	0.059676009	0.064654602	0.067996928	0.066623524
Average delta beta	0.08185839	0.06895543	0.06207783	0.06075474	0.06075007
Gene	CD63	TNRC6A		ANGPT1	VIPR2
Genomic context	Body	Body		TSS200	Body
Relation to CpG island	N_Shore	OpenSea	Island	OpenSea	OpenSea

LUHMES cells were compared to those in the hippocampus of 12-month-old mice with high-level (sixfold over normal) expression of WT human aSyn from a BAC construct (14). Hippocampal DNAm and DNAhm levels were profiled by reduced representation bisulfite sequencing (RRBS) with the same paired bisulfite and oxidative bisulfite conversion approach as used for LUHMES cells (Schaffner et al., manuscript in preparation). At the DNAm level, 4037 of 8391 differentially methylated genes in WT aSyn LUHMES cells were also covered in the mouse RRBS data. Of these, 44 genes were differentially methylated in control vs. WT aSyn mice ($padj \leq 0.05$ and |delta beta| ≥ 0.1, chosen to match the minimum $10\times$ read coverage filter applied during RRBS QC). The number of genes differentially methylated within this subset of 4037 in mice was greater than expected by chance ($padj < 0.001$, 10 000 permutations). These included *Grm3*, encoding a metabotropic glutamate receptor, as well as several genes encoding transcriptional regulators, cell adhesion proteins, glucose and ion transporters, and cell signaling proteins (Supplementary Material, Table S8). At the DNAhm level, 50 of 136 differentially methylated LUHMES cell genes were covered by RRBS, with 9

being differentially hydroxymethylated (greater than expected by chance, $padj < 0.001$, 10 000 permutations). Differentially hydroxymethylated genes in WT aSyn mice and LUHMES cells had similar functions to differentially methylated genes, including transcription regulation, cell adhesion and signaling, as well as a folate transporter (*Slc19a1*) and vacuolar ATPase (*Atp6v0a1*; Supplementary Material, Table S9).

aSyn impacted epigenetic and transcriptional regulation of glutamate, NOTCH, insulin, PDGF, and SHH signaling network genes

We next queried genes and pathways associated with changes in both the epigenome and the transcriptome in WT and A30P aSyn LUHMES neurons. These loci represent candidates where DNAm and/or DNAhm may be involved in regulation of gene expression, and their identification is important to understand the molecular etiology of PD. First, we identified CpG sites with alterations in both DNAm and DNAhm; in both genotypes, there were more CpGs with changes to both modifications than expected by chance (109 CpGs in WT aSyn cells, 34 CpGs in A30P aSyn cells, $padj < 0.001$, 1000 permutations)
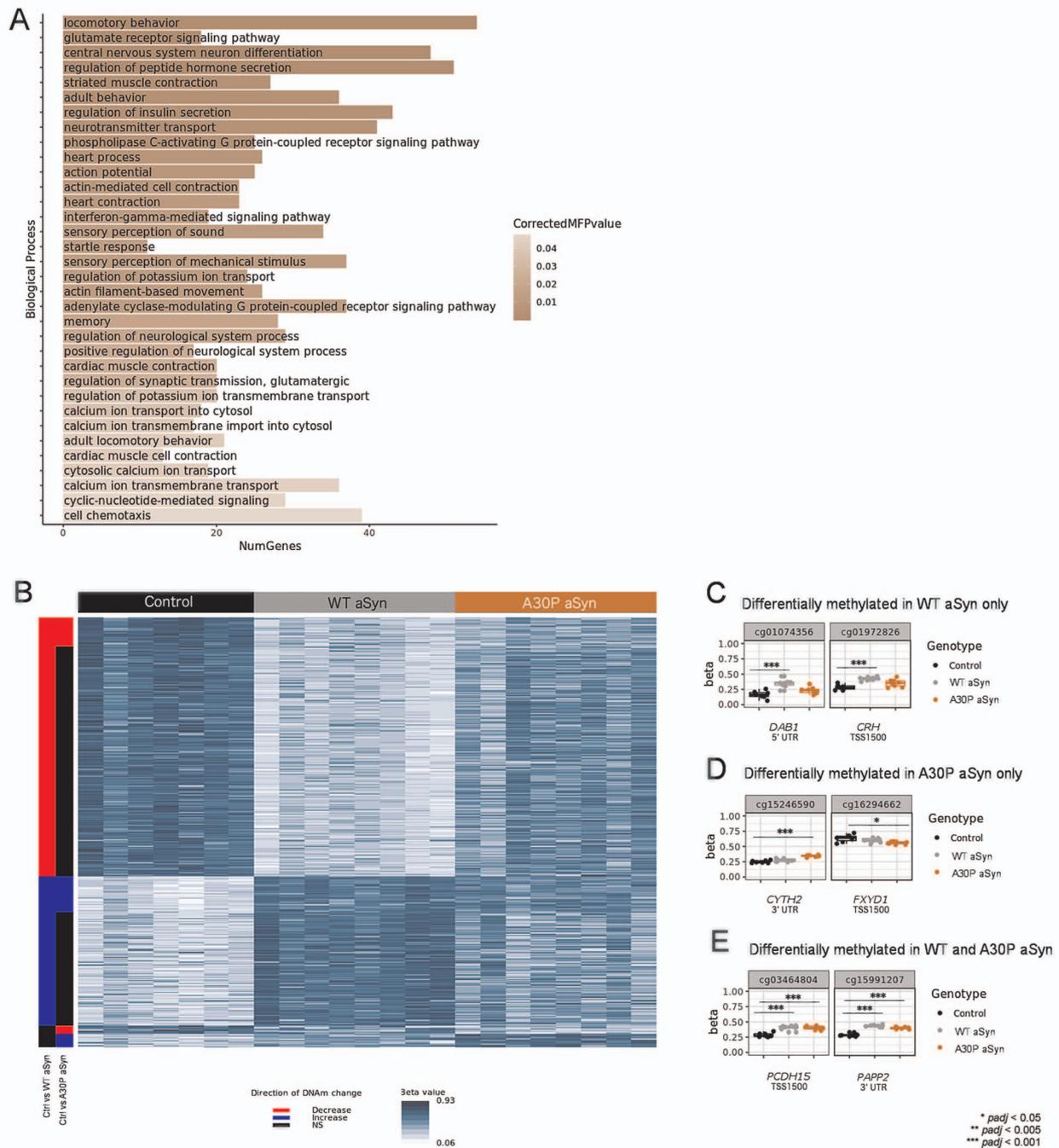
Figure 4. Differentially methylated genes in both genotypes were enriched for locomotor behavior and glutamate signaling functions. (A) GO biological process terms enriched at multifunctionality-corrected, multiple test-corrected hypergeometric *p*-value (CorrectedMFPvalue) < 0.05 in genes shared between control ($n = 7$ biological replicates) vs. WT aSyn ($n = 8$) and control vs. A30P aSyn ($n = 8$) DNAm analyses. (B) Heat map showing beta values for all probes annotated to the 'locomotory behavior' pathway and differentially methylated in at least one comparison. Row labels (left to right): probes that passed significance thresholds (absolute delta beta ≥ 0.05 and *padj* ≤ 0.05, *limma* empirical Bayes moderated *t* test with Benjamini–Hochberg adjustment) in control vs. WT aSyn comparison, probes that passed significance thresholds in control vs. A30P aSyn comparison. (C) Representative examples of 353 total locomotory behavior-related CpG sites differentially methylated only in WT aSyn cells (*DAB1 padj* $= 1.91e^{-4}$, *CRH padj* $= 9.34e^{-5}$). (D) Representative examples of 22 total locomotory behavior-related CpG sites differentially methylated only in A30P aSyn cells (*CYTH2 padj* $= 1e^{-5}$, *FXYD1 padj* $= 0.028$). (E) Representative examples of 67 total locomotory behavior-related CpG sites differentially methylated in both genotypes (*PCDH15* WT aSyn *padj* $= 7.8e^{-6}$, *PCDH15* A30P aSyn *padj* $= 8.61e^{-5}$, *PAPP2* WT aSyn *padj* $= 1.68e^{-9}$, *PAPP2* A30P aSyn *padj* $= 3.55e^{-7}$). *p*-values: *limma* empirical Bayes moderated *t* test with Benjamini–Hochberg adjustment.

(Supplementary Material, Fig. S6). This was likely to be a biological effect, as DNAhm can be an intermediate of DNA demethylation; thus, CpG sites with increased DNAhm are likely to have decreased DNAm (60). Indeed, all of the CpGs with changes to both DNAm and DNAhm

had opposite effect directions for each modification (Supplementary Material, Fig. S6).

Next, we compared the DNAm and DNAhm hits identified in this analysis against our previously published differential expression results from the same cells (863

differentially expressed genes in WT aSyn cells and 1315 genes in A30P aSyn cells, |log2FC| > 0.5 and *padj* < 0.01; validated by qPCR, see ref. 9). Of 10 798 genes with data for all three modifications in WT aSyn cells, seven showed differential DNAm, DNAhm, and expression, including an ionotropic glutamate receptor (*GRIK2*) (Supplementary Material, Table S10). In A30P cells, two of 10 831 genes had concurrent changes to these modifications: *TSPEAR* (cg04837104 and cg11631644 with altered DNAm, delta beta = 0.07 and 0.06, *padj* = 0.006 and 0.003; cg12061357 with altered DNAhm, delta beta = 0.06, *padj* = 0.0003; mRNA log2FC = −1.21, *padj* < 1e⁻⁴) and *SGPP2* (cg27350348 with altered DNAm and DNAhm, DNAm delta beta = −0.06, DNAm *padj* = 0.02, DNAhm delta beta = 0.07, DNAhm *padj* = 0.05; mRNA log2FC = 0.76, *padj* < 1e⁻⁴). The relations between DNAm, DNAhm, and expression varied according to the gene and the CpG site in question (Supplementary Material, Table S10). For both WT and A30P aSyn LUHMES cells, the overlap between differentially methylated, hydroxymethylated, and expressed genes was not greater than expected by chance (WT *padj* = 0.431, A30P *padj* = 0.171, 10 000 permutations).

We additionally employed SMITE, which captures modules of functionally related genes with changes to at least one of DNAm, DNAhm, or gene expression in each comparison (61). We included H3K4me1 ChIP-seq data from the same LUHMES cells used for DNA(h)m and mRNA analyses in our SMITE workflow, allowing us to consider DNAm and DNAhm changes at enhancers in addition to the default promoter and gene body regions, and weighted the importance of each modification (expression: 0.4; DNAm: 0.35; DNAhm: 0.25). Eighteen modules were identified in WT aSyn cells, encompassing functions including DNA damage repair, cell cycle control, neuronal differentiation, chaperone activity, and cell signaling (insulin, glutamate, and PPAR) (Fig. 5A, Supplementary Material, Fig. S7). Twenty-four modules were identified in A30P aSyn cells, which had similar functions to the WT modules as well as some additional modules related to the urea cycle, sumoylation, PDGF signaling, and K⁺ channel transport (Fig. 5B, Supplementary Material, Fig. S8). Modules from both comparisons were primarily driven by gene expression, with some contributions from DNAm and/or DNAhm (Fig. 5). This effect was consistent even when DNAm or DNAhm was weighted higher in the SMITE model (Supplementary Material, Tables S11, S12).

Glutamate receptor signaling modules discovered using SMITE had more concurrent changes to DNAm, DNAhm, and transcription in both genotypes. Two modules centered around *GRIK2* were each significantly altered in WT and A30P cells, consistent with the direct comparison of differentially (hydroxy)methylated versus differentially expressed genes. *GRIK2* promoter DNAm and mRNA levels were negatively correlated in both genotypes (Fig. 6). Several differentially methylated *GRIK2* promoter CpGs in both genotypes also belonged to a CMR (chr6: 101846707–101 846 916), which contained five individually significant CpGs from the control vs. WT aSyn comparison and one individually significant CpG from the control vs. A30P aSyn comparison.

Discussion

It is important to elucidate the mechanisms underlying the development and progression of PD in different individuals to develop preventative strategies and treatments. Here, we explored whether DNAm and DNAhm contributed to interindividual differences in PD etiology among carriers of aSyn variants. We assessed the impacts of overexpressing WT or A30P aSyn on genome-wide DNAm and DNAhm patterns in dopaminergic neurons, the primary cell type affected in PD. Our results showed that overexpressing either aSyn variant induced thousands of DNAm changes and hundreds of DNAhm changes in pathways related to PD and neurodegeneration, and that WT aSyn particularly impacted glutamate receptor signaling genes at both epigenetic and transcriptional levels. The distinct characteristics of each aSyn protein may explain why both similar and unique effects on DNA(h)m were observed. This study enhances our understanding of the wide-ranging genomic impacts of different aSyn forms and provides further insights into the possible molecular mechanisms of PD.

The association of aSyn expression with differential DNAm of genes associated with locomotor behavior observed here was consistent with previous reports that aSyn plays a role in dopamine biosynthesis and that dopaminergic neuron activity has impacts on locomotor phenotypes (62,63). The altered DNAm, DNAhm, and expression of neuronal differentiation genes in both genotypes also agreed with previous work in *SNCA* triplication PD patient-derived induced pluripotent stem cells (iPSCs), which indicated reduced expression of genes involved in neuronal differentiation and impaired differentiation capacity associated with aSyn overexpression (64). Several genes identified in our SMITE network analysis have altered epigenetic regulation in the brains of PD patients, including *DLG4* (differentially methylated in PD cingulate gyrus) and *MAGI2* (differentially methylated in the frontal cortex and blood of PD patients, and shows increased H3K27ac in the prefrontal cortex of PD patients) (13,15,17). Glutamate signaling pathway genes have also been implicated in PD, primarily through excessive glutamatergic transmission associated with excitotoxicity in dopaminergic neurons (65). Consistent with these observations, one of the differentially methylated genes shared between WT aSyn-expressing LUHMES neurons and mouse hippocampus was a metabotropic glutamate receptor, and differentially methylated and hydroxymethylated genes in WT aSyn-expressing mice had similar functions to genes with differential methylation, hydroxymethylation, and/or expression in our SMITE analysis. These observations suggested that although
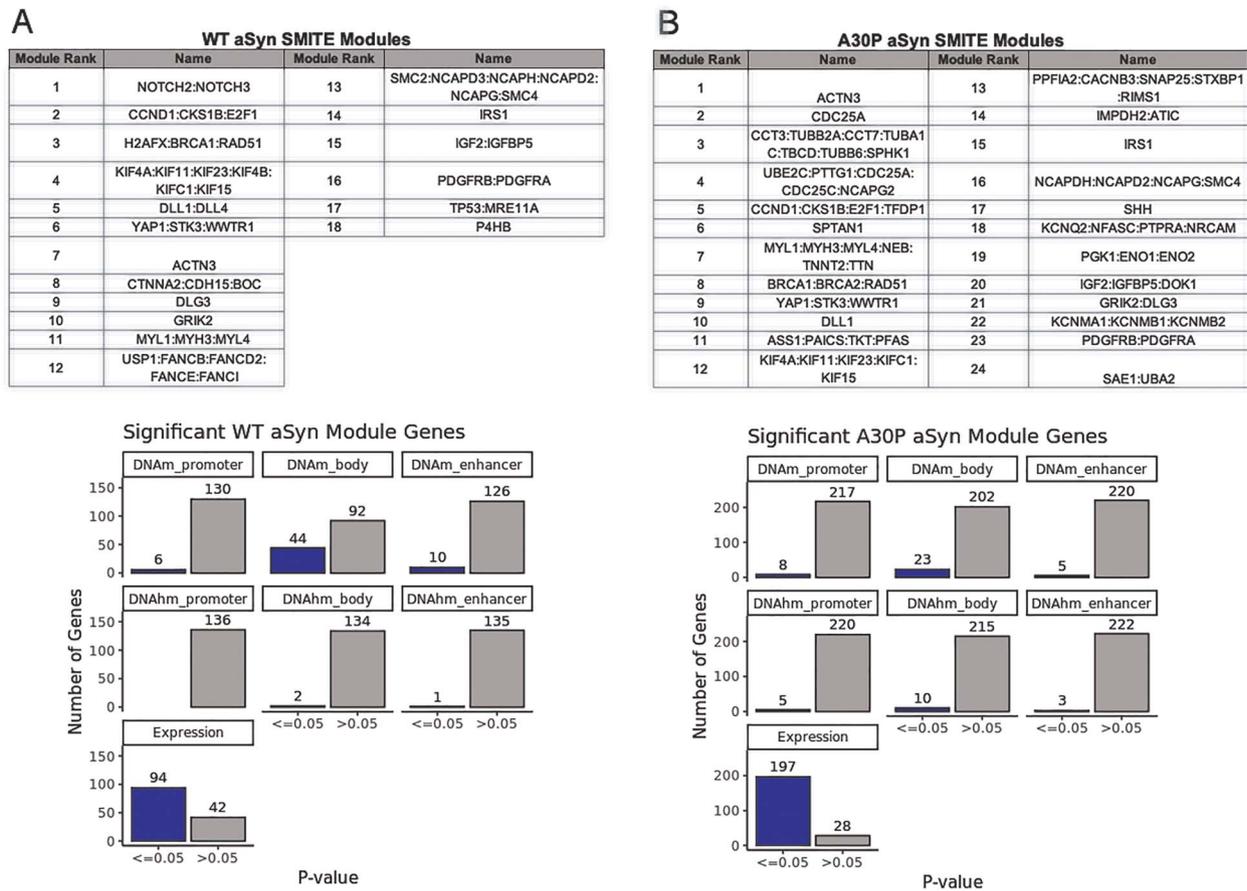
Figure 5. Modules from SMITE analysis were driven by gene expression, with contributions from DNAm and/or DNAhm. (A) Top: Names and ranks of 18 SMITE modules in control vs. WT aSyn LUHMES cells. Bottom: Distribution of significant ($p \leq 0.05$) and non-significant ($p > 0.05$) SMITE combined p-values (Fisher's method) for each score category, across 136 genes with at least one significant category in control vs. WT aSyn analysis. (B) Top: Names and ranks of 24 SMITE modules in control vs. A30P aSyn LUHMES cells. Bottom: Distribution of significant ($p \leq 0.05$) and non-significant ($p > 0.05$) SMITE combined p-values (Fisher's method) for each score category, across 225 genes with at least one significant category in control vs. A30P aSyn analysis. DNA(h)m data: $n = 7$ control LUHMES cells and $n = 8$ WT aSyn and A30P aSyn LUHMES cells, expression data: $n = 3$/group (9), H3K4me1 ChIP-seq data: $n = 3$/group. All samples are distinct biological replicates.

DNAm and DNAhm patterns are highly cell type-specific and can change with age and environment, some of the major signaling, transcriptional, and metabolic pathways altered by aSyn expression in this model are more likely to have functional relevance across different physiological contexts, possibly including the human brain. In addition, many of the CpGs identified in site-specific differential (hydroxy)methylation analyses belonged to CMRs, including *SNCA* and *GRIK2*, suggesting that these CpGs are likely to impact regulation of glutamate receptor genes and other pathways affected by aSyn as a unit (58).

Ionotropic and metabotropic glutamate receptor expression are at least in part epigenetically controlled, as transcription of both receptor types is modulated by H3K4 methylation across the human lifespan (66). Signaling cascades activated as a result of glutamate receptor stimulation can also result in gene expression changes, possibly including expression of glutamate receptor genes themselves (67). Therefore, it is possible that aSyn-induced changes in DNAm of the promoters of glutamate receptor genes may be one mechanism

underlying the regulation of their expression and subsequent glutamate release, thus facilitating excitotoxicity. It is also possible that existing high levels of glutamate release in the context of aSyn expression resulted in activation of signaling cascades in this study, which altered glutamate receptor expression, and DNAm changes occurred after this expression change. Although it is not possible to determine which of these scenarios is true based on our data, this study provided reasonable support for a role of aSyn-mediated epigenetic deregulation of glutamate receptor expression, with potential downstream impacts on synaptic activity and excitotoxicity. Indeed, preclinical studies have highlighted glutamate receptors as therapeutic targets for PD, and regulation of glutamate signaling genes may be influenced by lifestyle factors, such as dietary exposure to neurotoxins or tea polyphenols (26,68–70). Consistent with our findings, aSyn overexpression was also recently shown to induce glutamate release from mouse neurons and astrocytes *in vitro* and *in vivo* and to activate glutamate receptors (71). While we did not
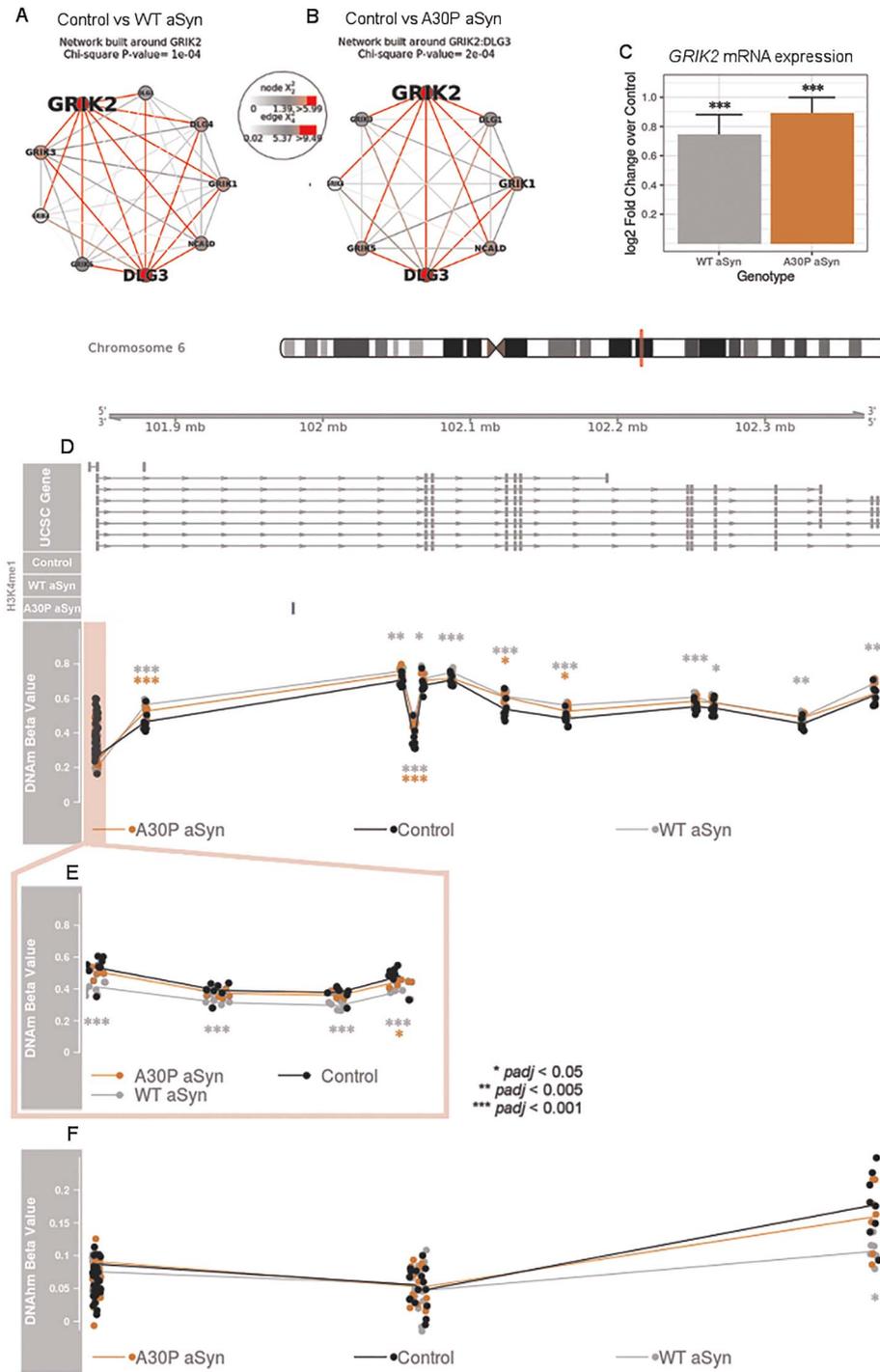
Figure 6. *GRIK2* was differentially methylated, hydroxymethylated, and expressed in LUHMES cells overexpressing either WT or A30P aSyn. (A) GRIK2 protein–protein interaction network showing differential DNAm, DNAhm, and expression in control vs. WT aSyn cells for the underlying genes. (B) GRIK2:DLG3 protein–protein interaction network showing differential DNAm, DNAhm, and expression in control vs. A30P cells for the underlying genes. (C) Relative expression of *GRIK2* mRNA in each genotype, normalized relative to control cells (9). WT aSyn *padj* = 1.46e^{-6}, A30P aSyn *padj* = 2.71e^{-15} (*DESeq2* Wald test with Benjamini–Hochberg adjustment). Error bars: SE. (D) Top: UCSC hg19 coordinates are shown, with *GRIK2* transcripts below. Middle: LUHMES cell H3K4me1 consensus peaks are shown in blue. Bottom: DNAm beta values are shown for each sample, colored by genotype. (E) DNAm at the *GRIK2* TSS200 region. (F) DNAhm levels for EPIC array probes across the *GRIK2* gene. *p*-values: *limma* empirical Bayes moderated *t* test with Benjamini–Hochberg adjustment. DNA(h)m data: *n* = 7 control LUHMES cells and *n* = 8 WT aSyn and A30P aSyn LUHMES cells, expression data: *n* = 3/group (9), H3K4me1 ChIP-seq data: *n* = 3/group. All samples are distinct biological replicates.

directly examine glutamatergic transmission in this study, our results provide evidence from an epigenetic perspective to expand existing research on the role of glutamate signaling in the etiology and prevention of PD, and support a potential role for aSyn in this pathway.

In addition to these core similarities in the pathways affected by each aSyn variant, WT and A30P aSyn also had unique effects on the DNA methylome and hydroxymethylome. Broadly, many more DNAm and DNAhm changes were seen in WT aSyn cells than in A30P cells in comparison with controls. This was somewhat surprising in light of our previous work, which showed greater effects at the transcriptional level in A30P cells (9). These differences were unlikely to be related to aSyn protein levels, as both WT and A30P LUHMES neurons expressed similar levels of aSyn (9). It was also unexpected that there were almost no changes to non-CpG methylation with aSyn overexpression (one differentially methylated CH probe in WT vs. A30P aSyn cells) (Supplementary Material, Table S3) despite known enrichment of non-CpG methylation in neurons (72,73).

Several mechanisms may explain the greater number of DNAm changes observed in WT aSyn cells than in A30P aSyn cells. First, DNA damage occurs to a greater extent in WT aSyn LUHMES neurons than in A30P aSyn LUHMES neurons, and it is possible that some DNAm changes in WT aSyn cells are a reflection of this damage and/or subsequent repair attempts (9,74). Following DNA damage, protective mechanisms in areas of active transcription maintain DNA repair and restore epigenetic modifications, which may also explain why we observed more DNAm changes in intergenic regions but fewer in intragenic regions than expected by chance in both genotypes (74,75). Second, differences in characteristics of each protein could also explain the greater number of DNAm alterations in WT cells. WT aSyn could influence DNAm through various mechanisms, including binding to membrane-bound G protein-coupled receptors and initiation of signaling cascades; facilitation of SNARE complex assembly at presynaptic membranes, which could allow neurotransmitters to activate signaling cascades at postsynaptic neurons; and binding to DNA (6,8,76). The A30P mutation prevents aSyn from binding to membranes, thus removing some of these avenues of influence (7). Finally, WT aSyn can sequester Dnmt1 from the nucleus in mice, resulting in global loss of DNAm in the brain (77). Differences in DNMT1 sequestration between aSyn variants could reduce the capacity for maintenance of DNAm in a nonspecific manner. aSyn aggregation, which is more likely with the WT protein, has also been shown to increase this sequestration (77). Although A30P aSyn is more likely to be localized to the nucleus than WT aSyn, the ability of A30P aSyn to sequester Dnmt1 has not been assessed (76–78). In addition to influencing DNAm loss in WT aSyn cells and gain in A30P aSyn cells, this potential difference in nuclear aSyn may also affect the amount of DNAm at transcription-associated regions due to its ability to bind DNA (6,76–79).

In contrast to DNAm, the patterns of differential hydroxymethylation were similar between aSyn genotypes, and fewer DNAhm changes were seen at our statistical thresholds. This was expected due to the fetal origin of the LUHMES cell line (43,80). Some of these DNAhm patterns in differentiated LUHMES cells may still be informative for later life stages as certain genes, particularly cell type-specific genes, retain similar brain DNAhm signatures through early development into adulthood (72). The overall increase in DNAhm with aSyn overexpression observed here was consistent with previous reports showing that DNAhm levels are increased in neurons and cerebellar white matter from PD patients (48,49). The opposite direction of change for DNAm and DNAhm at the same loci was also expected, as DNAm must be oxidized for the formation of DNAhm (60). From the results of this study, it is not possible to determine whether this DNAhm was a transient intermediate or represented a stable epigenetic mark; however, correlating DNAhm patterns with gene expression levels has been shown to provide some insight regarding loci where it is associated with transcription (43).

It is important to understand the interplay between DNAm, DNAhm, and gene expression to gain insight into the molecular etiology of PD, and cell culture models, such as LUHMES cells, provide appropriate platforms for concurrent multi-omics profiling. The relatively weak correlations between DNA(h)m changes and expression changes in this study ($< 0.1\%$ of genes in either genotype) were expected and were consistent with the literature, where correlations with expression at only 0.6%–15% of CpG sites in blood and up to 0.3% of CpG sites in brain have been reported (81–85). Therefore, it was also unsurprising that there was no significant enrichment for the overlap between differentially methylated, differentially hydroxymethylated, and differentially expressed genes. However, the few loci where transcriptional changes were correlated with DNA(h)m changes may still have either functional relevance for PD or represent potential biomarkers of neurodegeneration-relevant pathways. For the subset of genes with changes to both DNAm and mRNA expression, DNAm changes may have occurred first, thus impacting expression of disease-relevant genes, or the reverse may be true, i.e. DNAm changes may reflect alterations in gene expression (81).

When changes in DNAm precede changes in transcription, DNAm may represent a target for modulation of gene expression to prevent, for example, upregulation of glutamate signaling genes in individuals with aSyn multiplications. Indeed, an epigenetic editing approach using site-directed DNAm via a dCas9-DNMT3A fusion construct was reported previously to successfully reduce aSyn levels in SNCA-triplication patient-derived iPSC dopaminergic neurons, demonstrating therapeutic possibilities for manipulation of DNAm in modulating aberrant transcription (86). Conversely, when DNAm changes follow transcriptional changes, DNAm is more likely to represent a biomarker of expression.

Finally, DNAm alterations that did not affect gene expression at this cross-sectional time point may have an impact on expression at later times or upon

exposure to particular stimuli (87). Further studies involving exposure of aSyn-expressing neurons to substances and stimuli relevant to PD risk and/or neuroprotection may be useful to examine whether the role of DNA(h)m in the toxicity of aSyn is related to an influence on the ability of the cell to respond appropriately to either risk or protective factors for PD.

The assessment of whether changes in DNA(h)m associated with aSyn overexpression are correlated with transcriptional changes will provide insights regarding whether the roles of cytosine modifications in the toxicity of aSyn are mediated by effects on gene expression. However, it should be noted that this study was not designed to elucidate the mechanisms underlying differential methylation in certain loci in our aSyn-expressing LUHMES neurons. In the general context of the complex and multi-layered regulation of chromatin biology, and given the specific limitations of existing epigenome-wide association study approaches, it is only possible to speculate about the pathways and signals targeting specific CpG loci across the human genome, including the observed alterations in DNAm levels at the endogenous SNCA gene itself (33).

This study provided novel insights into the epigenomic impact of aSyn overexpression and set the stage for future studies in samples derived from PD patients. Nevertheless, it should be taken into consideration that integration of the SNCA transgene occurred at random positions in the genome in our cells, which could affect DNAm levels surrounding the integration positions. To minimize this variability and represent a range of integration positions, we used eight replicates of each LUHMES cell genotype and seven replicates of control LUHMES cells for EPIC array profiling, with each replicate consisting of a pool of thousands of cells. Cell culture has also been shown to influence DNAm (88–90). We reduced the effects of such variability by using replicates obtained from a range of passages and culture dishes, and regressed out the effects of passage before analysis of differential DNA(h)m. In addition, as expected, determination of DNAhm using a combination of oxidative bisulfite- and bisulfite-converted samples introduced technical noise, which prompted us to reduce the number of sites for DNAhm analysis (91). Finally, the LUHMES cells used in this study are an approximation of dopaminergic neurons in the brains of PD patients, but are subject to the inherent limitations of any *in vitro* cellular model. However, LUHMES cells and other tissue culture models of aSyn overexpression (primarily iPSCs and neuroblastoma lines) have been used successfully in previous studies to uncover biochemical mechanisms of PD, including mitochondrial dysfunction, increased H3 acetylation, and cell–cell transfer of aSyn protein (9,92,93). LUHMES cells provide a useful platform in which to study the biology of aSyn and possibly gain insight into preclinical stages of PD, prior to aSyn aggregation and neurotoxicity on a large scale. As aSyn overexpression has been shown previously to affect

cellular physiology even at the neural precursor stage, such studies are still valuable (94). In addition, our LUHMES cells have an advantage over iPSC studies in that the cells are of neural origin, reducing the potential for epigenetic changes associated with iPSC reprogramming (89,90).

Overall, this study demonstrated that WT and A30P aSyn overexpression each had significant impacts on the DNA methylome of human dopaminergic neurons, and that only a small proportion of these DNAm changes occur concomitantly with DNAhm and changes in expression of neurodegeneration-related genes. This study contributes to our understanding of the molecular genetic etiology of PD and lays groundwork for future studies to investigate whether epigenetic and transcriptional changes associated with increased expression or mutations in aSyn may be reversible by modification of environmental and/or lifestyle factors.

Materials and Methods
Cell culture

LUHMES cells were a kind gift from Prof. Marcel Leist and were cultured and differentiated as described previously (9,80,95); the methods below are reproduced from (9) with permission. Briefly, proliferating cells were cultured in Advanced Dulbecco's modified Eagle's medium/F12 (DMEM/F12, Gibco, Thermo Fisher Scientific, Waltham, MA, USA) supplemented with N2 (Gibco), 2 mM L-glutamine (Gibco), and 40 ng/mL recombinant basic fibroblast growth factor (bFGF, R&D Systems Inc., Minneapolis, MN, USA), in pre-coated flasks (Corning Inc., Corning, NY, USA) with 50 μg/mL PLO (Sigma-Aldrich, St. Louis, MO, USA), at 37°C in a humidified 5% CO_2 atmosphere. Differentiation was achieved by replacing the proliferating medium with DMEM/F12 supplemented with N2 (Gibco), 2 mM L-glutamine (Gibco), 1 μg/mL tetracycline (Sigma-Aldrich), 1 mM dibutyryl cAMP (cAMP; Sigma-Aldrich), and 2 ng/mL recombinant human GDNF (R&D Systems). After 2 days in culture (pre-differentiation), cells were trypsinized and seeded into plates pre-coated with 50 μg/mL PLO and 1 μg/mL fibronectin (Sigma-Aldrich). On day 5 of differentiation, one third of the culture medium was replaced with fresh medium. Cells were harvested after 8 days of differentiation.

Generation of aSyn-expressing LUHMES cells

Expression of aSyn in LUHMES cells was achieved as described previously (9,95); the methods below are reproduced from (9) with permission. Full-length human aSyn c-DNA (SNCA, NM_000345) and A30P aSyn were subcloned into the second-generation bicistronic lentiviral vector, pWPI (Tronolab, Lausanne, Switzerland), under the control of the chicken β-actin (CBA) promoter. All cloned sequences were verified by direct sequencing (StarSeq, Mainz, Germany). The pWPI lentiviral vector

containing only the internal ribosomal entry site (IRES)-green fluorescent protein (GFP) cassette was used as a positive control for infection in all experiments. Second-generation lentiviral particles were generated as described previously (96). Briefly, 293 T cells were transiently co-transfected with the modified transfer vector plasmids and second-generation packing system (Tronolab). Supernatant was collected 48-h post-transfection, concentrated by PEG-it Virus Precipitation Solution (System Biosciences, Palo Alto, CA, USA), and resuspended in Panserin 402 (Pan-Biotech, Aidenbach, Germany). Lentiviral vector titers were measured quantitatively by qPCR using primer sequences specific for the woodchuck hepatitis virus post-transcriptional regulatory element (WPRE), as described previously (97). To generate cells expressing aSyn, proliferating LUHMES cells were infected with equimolar amounts of IRES-GFP and WT aSyn-IRES-GFP or A30P aSyn-IRES-GFP lentivirus. Cells positive for green fluorescence were selected by FACS (FACSCalibur; BD Biosciences, Franklin Lakes, NJ, USA).

Immunocytochemistry

Immunocytochemistry was performed as described previously (9,95); the methods below are reproduced from (9) with permission. Cells were grown on coverslips and fixed at different time points of differentiation with 4% paraformaldehyde for 40 min at room temperature (RT). After washing with 1× phosphate buffered saline (PBS), the cells were permeabilized with 0.5% Triton/PBS for 15 min, and finally blocked with 1.5% normal goat serum (NGS)/PBS (blocking buffer) for 1 h at RT. Cells were incubated with anti-aSyn antibody (mouse, 1:1000; BD Transduction Laboratories, Franklin Lakes, NJ, USA) and anti-TH antibody (rabbit, 1:1000; Millipore, Billerica, MA, USA) for 3 h. The cells were then washed with 1× PBS and probed for 2 h at RT with secondary antibodies (rabbit or mouse, 1:1000 Alexa 488 and 1:1000 Alexa 555; Life Technologies, Carlsbad, CA, USA) diluted in blocking buffer. Finally, the nuclei were stained with DAPI and mounted with Mowiol (Sigma-Aldrich). The images were acquired on a Leica 6000B microscope (Leica Microsystems, Wetzlar, Germany).

DNA extraction and bisulfite conversion

Cell pellets were thawed on dry ice and homogenized with a 20G needle in Qiagen Buffer RLT Plus (Qiagen, Valencia, CA, USA) with β-mercaptoethanol. DNA extractions were performed with a Qiagen AllPrep DNA/RNA Mini Kit in accordance with the manufacturer's instructions. DNA quantity and purity were assessed by spectrophotometry.

DNA samples of 1.5 μg were split into two equal aliquots of 750 ng for sodium bisulfite (BS) conversion or oxidative bisulfite (oxBS) conversion using the NuGEN TrueMethyl oxBS Module (NuGEN Technologies, San Carlos, CA, USA). One aliquot per sample was treated with the oxidation protocol, while the second aliquot was subjected to mock oxidation. All samples were then sodium bisulfite converted according to the manufacturer's instructions.

DNA methylation and DNA hydroxymethylation analysis

Aliquots of 750 ng per sample of BS- or oxBS-converted DNA were run on Infinium HumanMethylationEPIC (EPIC) BeadChips (Illumina, San Diego, CA, USA) according to the manufacturer's instructions, producing data for 853 307 CpG and 2880 CNG sites (where N represents any other nucleotide) (data available at https://www.ncbi.nlm.nih.gov/geo/ with accession No. GSE181126). Beta values were generated from raw intensity signals using GenomeStudio software (Illumina) and exported into R ver. 3.6 (R Foundation for Statistical Computing, Vienna, Austria) for data analysis. Filtering was performed to remove 59 internal SNP control probes and 43 254 cross-hybridizing probes (98). The pfilter function in the wateRmelon package was additionally used to remove sites with a detection $p > 0.05$ in 1% of samples (7685) and sites with a bead count < 3 in 5% of samples (1626) (98). This left 813 589 EPIC probes in the final dataset. The dasen function in wateRmelon was used to normalize the oxBS and BS data separately (99). Batch effects from chip, position, and passage were removed using the ComBat function in the SVA package (100). The hmC values were calculated by subtracting the normalized oxBS beta values from the normalized BS beta values. A threshold for hmC detectability (3.6%) was calculated using the 95% quantile of negative hmC values generated after subtraction; all probes with a mean hmC level below this threshold were discarded (91).

Differential methylation between control ($n = 7$), WT aSyn ($n = 8$), and A30P aSyn ($n = 8$) cells was calculated by linear regression using limma with the model DNA(h)m ~ Genotype (101). Individual sites were considered significant at FDR ≤ 0.05 and |delta beta| ≥ 0.05 (calculated by subtracting the mean beta value across all replicates of the group in question from the reference group, e.g. WT aSyn beta − Control beta). The skewness.norm.test function from the normtest R package was used to assess delta beta skewness, with 1000 Monte Carlo simulations (https://cran.r-project.org/web/packages/normtest/normtest.pdf).

CMRs were defined separately in pre-processed DNAm and DNAhm data from LUHMES neurons using the cmr function from the CoMeBack R package with a maximum distance of 1 kb, minimum Pearson's correlation coefficient of 0.4, and minimum of two CpGs per CMR (58). Composite betas were calculated with the cmr_comp function, using the default method (weighted average of the first principal component of CMR probe DNA(h)m levels). As in site-specific analysis, linear modeling with limma was applied to the DNAm and DNAhm composite betas, with one composite beta

value and one CpG probe ID representing each CMR. The overlap in differentially (hydroxy)methylated probes from the CMR and site-specific analyses was permuted 10 000 times by randomly sampling a set of CpG probes from the preprocessed EPIC DNAm or DNAhm data the same size as the number of significant site-specific probes, randomly sampling a set of representative CMR probes from the DNAm or DNAhm CMRs the same size as the number of significant CMRs, and calculating how many iterations had greater or smaller overlap than the real value.

All code is publicly available at github.com/samschaf/LUHMES.

Pyrosequencing

1 μL of BS or oxBS-converted DNA per 30-μL reaction was used to perform PCR with HotStarTaq DNA Polymerase (Qiagen) for 45 cycles with an annealing temperature of 58°C. PCR samples were visualized on a 1% agarose gel post-amplification to ensure integrity.

Pyrosequencing assays were designed with PyroMark® Assay Design 2.0 software (Qiagen) (Supplementary Material, Table S5). 10 μL of PCR reaction mixture, 70 μL of binding buffer, and 12 μL of sequencing reaction mix per sample were prepared with PyroMark® Gold Q96 Reagents (Qiagen) according to the manufacturer's instructions. Sequencing was performed with the PyroMark® Q96 ID (Qiagen). Differential DNAm/DNAhm between groups was assessed by Welch's two-sample t test (two-tailed), using the t.test() function in R.

Generation of aSyn-expressing mice and genome-wide DNAm/DNAhm profiling

Generation of WT aSyn-expressing mice and DNA extraction were performed as described previously (14). Hippocampal DNA from four WT and four transgenic mice raised under standard housing conditions to the age of 12 months was profiled by RRBS for genome-wide DNAm and DNAhm. RRBS libraries were prepared following the protocol of Gu *et al.* (102), using 400 ng of genomic DNA per sample. 0.4 ng of unmethylated phage lambda DNA per sample was added during the *Msp*I digestion step as a control for bisulfite conversion efficiency, and *Msp*I reactions were performed with incubation at 37°C for 16 h. Digested DNA was end-repaired, and then purified using AMPure XP magnetic beads (Beckman Coulter Life Sciences, Fullerton, CA, USA) at 2× concentration, and 0.1 μM methylated adapters (IDT, Coralville, IA, USA) were used in ligation reactions. Ligated DNA was purified with 1.5× magnetic beads and quantified using a Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific). Eight samples per lane were pooled in equimolar fashion, and each pool was split into two aliquots for oxBS and BS conversion using the NuGEN TrueMethyl oxBS Module (NuGEN Technologies). One aliquot was subjected to oxidation, while the second aliquot was subjected to mock oxidation; the bisulfite conversion protocol was

then applied to both. Libraries were then amplified by PCR with Pfu Turbo Cx HotStart DNA polymerase for 14 cycles (Agilent Technologies, Santa Clara, CA, USA). Size selection (200–550 bp) was performed using a 0.55× AMPure beads wash (Agencourt Bioscience, Beverly, MA, USA) followed by two 1.2× washes for cleanup. The final libraries were eluted in 25 μL of Qiagen EB Buffer (Qiagen), and quality was assessed by PCR, spectrophotometry, fluorometry, and DNA high-sensitivity chip (Agilent Technologies). 75 bp paired-end sequencing was conducted with Illumina HiSeq 2500 by the Michael Smith Genome Sciences Centre (Vancouver, BC, Canada).

BSMAP was used to align reads to the mm10 and phage lambda genomes and to calculate methylation ratios (103). Sequence data was assessed for quality before and after alignment using FastQC software (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Methylation ratio data was then read into R ver. 3.6.3 and filtered for: a minimum read depth of 10×; a maximum read depth of the 99.9% quantile of coverage; and sites covered in all samples. oxBS beta values were used to represent DNAm, while BS − oxBS values were used to represent DNAhm, with the 95% quantile of negative hmC values generated after subtraction used as a detectability threshold (91). We removed cytosines with ≤ 10% variability between groups to reduce the multiple testing burden for differential DNA(h)m analysis.

Differential DNA(h)m analysis was conducted with the betaRegression function from the BiSeq R package, using beta-binomial regression with the Wald test for significance (104). The p-values were adjusted for multiple comparisons using the Benjamini–Hochberg method. Cytosines were considered differentially methylated at $padj \leq 0.05$ and |delta beta| ≥ 0.1.

To determine whether more genes differentially methylated in WT aSyn-expressing LUHMES neurons were also differentially methylated in mice than expected by chance, we first subset the RRBS data to genes differentially methylated in LUHMES cells. We considered the 'hit list' of differentially methylated cytosines in mice as the number of cytosines within this subset passing $padj \leq 0.05$ and |delta beta| ≥ 0.1. Next, we randomly selected a set of cytosines the same size as this 'hit list' from the subsetted RRBS background, and determined the genes to which they mapped. The number of differentially methylated genes in the random selection was compared to the number of differentially methylated genes in the 'hit list.' This process was repeated 10 000 times in each of the DNAm and DNAhm datasets. Enrichment and depletion p-values were calculated by dividing the number of iterations with more or fewer differentially methylated genes than the real number by the number of permutations.

Gene ontology enrichment

Genes were assigned to CpGs if they belonged to the longest UCSC RefGene transcript annotated to each

site, resulting in 30 325 unique coding and non-coding gene annotations for the DNAm dataset and 23 190 unique gene annotations for the DNAhm dataset. Gene Ontology enrichment analysis was performed with the over-representation (ORA) method in ermineR 1.0.1.9, using differentially methylated genes identified with the effect size and significance thresholds described above as input (59). Genes that were differentially methylated or hydroxymethylated in both WT aSyn and A30P aSyn cells in either direction were removed from the input list and analyzed separately. All code is publicly available at github.com/samschaf/LUHMES.

RNA sequencing and differential expression analysis

RNA-seq of LUHMES cells used for multi-omic integration analysis in this study was described previously (9,95) (data available at https://www.ncbi.nlm.nih.gov/geo/ with accession No. GSE89115); the methods below are reproduced from (9) with permission. Total RNA from differentiated LUHMES cells was extracted and purified using an RNeasy mini kit (Qiagen) in accordance with the manufacturer's instructions. Three biological replicates were used for RNA-seq experiments. Sequencing and RNA quality analyses were performed as described previously (105). Briefly, RNA quality was assessed using RNA 6000 Nano chips run on a 2100 Bioanalyzer (Agilent Technologies). Libraries were prepared using a TruSeq RNA Sample Preparation v2 kit (Illumina). The library quality was checked using High-Sensitivity DNA chips on a Bioanalyzer (Agilent Technologies). The sample concentration was measured with a Qubit dsDNA HS Assay Kit and adjusted to 2 nM before sequencing (50 bp) on a HiSeq 2000 sequencing system (Illumina) in accordance with the manufacturer's instructions.

Differential gene expression analysis was performed as described previously (105). Briefly, RNA-seq data were aligned to the genome using STAR with the default options, which generated mapping files (BAM format). Differential expression read counts were generated using featureCounts, and samples were compared for differential expression using DESeq2 (106,107). Genes with $p \leq 0.05$ and mean read count ≥ 10 were considered to be differentially expressed.

Transcript-level analysis of *SNCA* mRNA was performed with Salmon v1.6.0 and the hsapiens gentrome (genome + transcriptome) derived from Gencode v39 using the following parameters: —numGibbsSamples 20 —seqBias —gcBias —validateMappings (108). Scaling abundances to final values was done in R with tximeta v1.10.0 (109). Differential expression of transcripts was calculated with the fishpond R package (110).

Chromatin immunoprecipitation sequencing (ChIP-seq)

Chromatin and DNA were cross-linked by adding formaldehyde to the culture dish to a final concentration of 1% and incubating for 10 min at 25°C. The cross-linking

reaction was quenched with glycine at a final concentration of 0.125 M at 25°C. Cells were then collected, washed with cold PBS, and pelleted by centrifugation at 900 rcf for 5 min at 4°C. Cell pellets were resuspended in RIPA SDS buffer (140 mM NaCl, 1 mM EDTA at pH 8.0, 1% Triton X-100, 0.1% sodium deoxycholate, 10 mM Tris-Cl at pH 8.0, 1% SDS) supplemented with Roche complete protease inhibitors, incubated at 4°C for 10 min, and then used for shearing in Covaris® instrument. The sheared chromatin was cleared by centrifugation at 16000 rcf for 5 min and an aliquot was used to check DNA size on an Agilent 2100 Bioanalyzer (Agilent Technologies). Immunoprecipitation was performed when DNA was sheared to fragments of 150–300 bp.

For chromatin immunoprecipitation (ChIP), samples were diluted 10× in IP buffer (150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 50 mM Tris–HCl, pH 8.0, 2 mM EDTA) supplemented with Roche complete protease inhibitors (Roche, Basel, Switzerland), pre-cleared with protein A magnetic beads (Dynabeads; Invitrogen, Carlsbad, CA, USA) for 1 h at 4°C, and then used for immunoprecipitation with histone modification-specific ChIP-grade antibody (anti H3 monomethyl K4 antibody, ab8895; Abcam, Cambridge, UK). After overnight incubation at 4°C, protein A magnetic beads were added and incubated for 2 h at 4°C. Beads were then washed twice with cold IP buffer with 0.1% SDS and protease inhibitors, 3× with wash buffer (100 mM Tris–HCl, pH 8.0, 500 mM LiCl, 1% v/v NP-40, 1% w/v sodium deoxycholate, 2 mM EDTA), and 2× with TE. After the last wash, supernatant was removed and beads were resuspended in 1 mM Tris (pH 8.0) and RNase A (0.1 μg/μL), and incubated 30 min at 37°C. To reverse the cross-linking, beads were incubated overnight at 65°C in 100 mM Tris–HCl (pH 8.0), 20 mM EDTA, 2% SDS, and proteinase K (0.5 μg/μL). The DNA was purified with SureClean (Bioline, Taunton, MA, USA) and DNA concentration was measured with a Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific).

Libraries were generated with 3 ng of ChIPed DNA using a NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs, Ipswich, MA, USA), size was determined with Agilent Bioanalyzer DNA high-sensitivity (Agilent Technologies), and sequencing was performed on an Illumina NovaSeq 6000 instrument (Illumina) according to the manufacturer's instructions. Read quality of ChIP-seq data was assessed using FastQC (v0.11.5) and reads were aligned to the human genome version 38 using Bowtie2 (v2.0.2) (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/, 111). MACS2 (v2.1.2) was used to call broad peaks from BAM files (112).

Multi-omic integration

The overlap between differentially methylated, hydroxymethylated, and expressed genes in each comparison was permuted 10 000 times to test whether it was greater or smaller than expected by chance. For this purpose, a set of EPIC probes from the preprocessed DNAm data

the same size as the number of differentially methylated probes was randomly selected, a set of EPIC probes from the preprocessed DNAhm data the same size as the number of differentially hydroxymethylated probes was randomly selected, and a set of genes from the RNA-seq data the same size as the number of differentially expressed genes was randomly selected. The gene-level overlap between randomly sampled DNAm, DNAhm, and RNA-seq data was then calculated, and this number was compared to the real number to calculate enrichment and depletion *p*-values.

DNAm, DNAhm, and RNA-seq data were also integrated using SMITE in R ver. 3.6, with adjusted *p*-values as significance input, delta betas as effect size input for DNAm and DNAhm, and log2FC values as effect size input for RNA-seq (62). Genomic regions were split into enhancers (± 5000 bp from H3K4me1 ChIP-seq peaks), promoters (± 1000 bp from TSS), and gene bodies (remaining regions). DNAm and DNAhm *p*-values were weighted across genomic regions according to their significance using Stouffer's method, creating a combined *p*-value for each region. The *p*-values from all modifications were then used to create weighted gene-level scores, with application of the following weights: expression, 0.4; enhancer DNAm, 0.125; promoter DNAm, 0.125; body DNAm, 0.1; enhancer DNAhm, 0.125; promoter DNAhm, 0.125; body DNAhm, 0.1. Weights were chosen to return modules likely to contain differentially expressed genes, with equal probability of contribution from differential DNAm and/or differential DNAhm (total weight of 0.25 for each methylation type). Enhancers and promoters were weighted slightly higher than gene bodies due to the higher likelihood that DNAm at these regions would influence gene expression. Alternative SMITE models with equal weights for each modification, DNAm weighted highest, DNAhm weighted highest, or DNAm and expression weighted higher than DNAhm were also tested, and produced overall similar results (Supplementary Material, Tables S11, S12). Scores were annotated to a REACTOME protein–protein interaction network, and a spin-glass algorithm was used to identify modules that had altered DNAm, DNAhm, and expression for each comparison. All code is publicly available at github.com/samschaf/LUHMES.

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgements

Funding

References

1. Emamzadeh, F.N. (2016) Alpha-synuclein structure, functions, and interactions. *J. Res. Med. Sci.*, **9**, 21–29.

2. Lesage, S. and Brice, A. (2009) Parkinson's disease: from monogenic forms to genetic susceptibility factors. *Hum. Mol. Genet.*, **18**, R48–R59.

3. GBD (2015) Neurological disorders collaborator group (2017) global, regional, and national burden of neurological disorders during 1990–2015: a systematic analysis for the global burden of disease study 2015. *Lancet Neurol.*, **16**, 877–897.

4. Morato Torres, C.A., Wassouf, Z., Zafar, F., Sastre, D., Outeiro, T.F. and Schüle, B. (2020) The role of alpha-synuclein and other Parkinson's genes in neurodevelopmental and neurodegenerative disorders. *Int. J. Mol. Sci.*, **21**, 1–32.

5. Gracia, P., Camino, J.D., Volpicelli-Daley, L. and Cremades, N. (2020) Multiplicity of α-synuclein aggregated species and their possible roles in disease. *Int. J. Mol. Sci.*, **21**, 1–27.

6. Pinho, R., Paiva, I., Jercic, K.G., Fonesca-Ornelas, L., Gerhardt, E., Fahlbusch, C., Garcia-Esparcia, P., Kerimoglu, C., Pavlou, M.A.S., Villar-Pique, A. *et al.* (2019) Nuclear localization and phosphorylation modulate pathological effects of alpha-synuclein. *Hum. Mol. Genet.*, **28**, 31–50.

7. Jo, E., Fuller, N., Rand, R.P., St George-Hyslop, P. and Fraser, P.E. (2002) Defective membrane interactions of familial Parkinson's disease mutant A30P alpha-synuclein. *J. Mol. Biol.*, **315**, 799–807.

8. Volpicelli-Daley, L.A., Luk, K.C., Patel, T.P., Tanik, S.A., Riddle, D.M., Stieber, A., Meaney, D.F., Trojanowski, J.Q. and Lee, V.M.Y. (2011) Exogenous α-synuclein fibrils induce Lewy body pathology leading to synaptic dysfunction and neuron death. *Neuron*, **72**, 57–71.

9. Paiva, I., Pinho, R., Pavlou, M.A., Hennion, M., Wales, P., Schütz, A., Rajput, A., Szego, E.M., Kerimoglu, C., Gerhardt, E. *et al.* (2017) Sodium butyrate rescues dopaminergic cells from alpha-synuclein-induced transcriptional deregulation and DNA damage. *Hum. Mol. Genet.*, **26**, 2231–2246.

10. Baptista, M.J., O'Farrell, C., Daya, S., Ahmad, R., Miller, D.W., Hardy, J., Farrer, M. and Cookson, M.R. (2003) Co-ordinate transcriptional regulation of dopamine synthesis genes by alpha-synuclein in human neuroblastoma cell lines. *J. Neurochem.*, **85**, 957–968.

11. Henderson, A.R., Wang, Q., Meechoovet, B., Siniard, A.L., Naymik, M., De Both, M., Huentelman, M.J., Caselli, R.J., Driver-Dunckley, E. and Dunckley, T. (2021) DNA methylation and expression profiles of whole blood in Parkinson's disease. *Front. Genet.*, **12**, 1–17.

12. Hentrich, T., Wassouf, Z., Reiss, O. and Schulze-Hentrich, J.M. (2018) *SNCA* overexpression disturbs hippocampal gene expression trajectories in midlife. *Aging*, **10**, 4024–4041.

13. Toker, L., Tran, G.T., Sundaresan, J., Tysnes, O., Alves, G., Haugarvoll, K., Nido, G.S., Dölle, C. and Tzoulis, C. (2021) Genome-wide histone acetylation analysis reveals altered transcriptional regulation in the Parkinson's disease brain. *Mol. Neurodegener.*, **16**, 1–20.

14. Wassouf, Z., Hentrich, T., Samer, S., Rotermund, C., Kahle, P.J., Ehrlich, I., Riess, O., Casadei, N. and Schulze-Hentrich, J.M. (2018) Environmental enrichment prevents transcriptional disturbances induced by alpha-synuclein overexpression. *Front. Cell. Neurosci.*, **12**, 112.

15. Young, J.I., Sivasankaran, S.K., Wang, L., Ali, A., Mehta, A., Davis, D.A., Dykxhoorn, D.M., Petito, C.K., Beecham, G.W., Martin, E.R. *et al.* (2019) Genome-wide brain DNA methylation analysis suggests epigenetic reprogramming in Parkinson disease. *Neurol. Genet.*, **5**, e42.

16. Pinho, R., Guedes, L.C., Soreq, L., Lobo, P.P., Mestre, T., Coelho, M., Rosa, M.M., Gonçalves, N., Wales, P., Mendes, T. *et al.* (2016) Gene expression differences in peripheral blood of Parkinson's disease patients with distinct progression profiles. *PLoS One*, **11**, e0157852.

17. Masliah, E., Dumaop, W., Galasko, D. and Desplats, P. (2013) Distinctive patterns of DNA methylation associated with Parkinson disease: identification of concordant epigenetic changes in brain and peripheral blood leukocytes. *Epigenetics*, **8**, 1030–1038.

18. Chuang, Y., Paul, K.C., Bronstein, J., Bordelon, Y., Horvath, S. and Ritz, B. (2017) Parkinson's disease is associated with DNA methylation levels in human blood and saliva. *Genome Med.*, **9**, 76.

19. Rawlik, K., Rowlatt, A. and Tenesa, A. (2016) Imputation of DNA methylation levels in the brain implicates a risk factor for Parkinson's disease. *Genetics*, **204**, 771–781.

20. Schmitt, I., Kaut, O., Khazneh, H., de Boni, L., Ahmad, A., Berg, D., Klein, C., Fröhlich, H. and Wüllner, U. (2015) L-dopa increases α-synuclein DNA methylation in Parkinson's disease patients *in vivo* and *in vitro*. *Mov. Disord.*, **30**, 1794–1801.

21. de Boni, L., Riedel, L., Schmitt, I., Kraus, T.F.J., Kaut, O., Piston, D., Akbarian, S. and Wüllner, U. (2015) DNA methylation levels of alpha-synuclein intron 1 in the aging brain. *Neurobiol. Aging*, **36**, 3334.e7–3334.e11.

22. Zhang, D., Cheng, L., Badner, J.A., Chen, C., Chen, Q., Luo, W., Craig, C.W., Redman, M., Gershon, E.S. and Liu, C. (2010) Genetic control of individual differences in gene-specific methylation in human brain. *Am. J. Hum. Genet.*, **86**, 411–419.

23. Teh, A.L., Pan, H., Chen, L., Ong, M., Dogra, S., Wong, J., MacIsaac, J., Mah, S.M., McEwen, L.M., Saw, S. *et al.* (2014) The effect of genotype and in utero environment on interindividual variation in neonate DNA methylomes. *Genome Res.*, **24**, 1064–1074.

24. Gaunt, T.R., Shihab, H.A., Hemani, G., Min, J.L., Woodward, G., Lyttleton, O., Zheng, J., Duggirala, A., McArdle, W.L., Ho, K. *et al.* (2015) Systematic identification of genetic influences on methylation across the human life course. *Genome Biol.*, **17**, 61.

25. Wang, H., Lou, D. and Wang, Z. (2019) Crosstalk of genetic variants, allele-specific DNA methylation, and environmental factors for complex disease risk. *Front. Genet.*, **9**, 695.

26. Vallegra, C.L., Zhang, F., Fowdar, J., McRae, A.F., Qui, T., Nabais, M.F., Zhang, Q., Kassam, I., Henders, A.K., Wallace, L. *et al.* (2020) Analysis of DNA methylation associates the cystine-glutamate antiporter SLC7A11 with risk of Parkinson's disease. *Nat. Commun.*, **11**, 1238.

27. Go, R.C.P., Corley, M.J., Ross, G.W., Petrovitch, H., Masaki, K.H., Maunakea, A.K., He, Q. and Tiirikainen, M.I. (2020) Genome-wide epigenetic analyses in Japanese immigrant plantation workers with Parkinson's disease and exposure to organochlorines reveal possible involvement of glial genes and pathways involved in neurotoxicity. *BMC Neurosci.*, **21**, 31.

28. Paul, K.C., Horvath, S., Del Rosario, I., Bronstein, J.M. and Ritz, B. (2021) DNA methylation biomarker for cumulative lead exposure is associated with Parkinson's disease. *Clin. Epigenetics*, **13**, 59.

29. Lim, I.Y., Lin, X. and Karnani, N. (2017) Implications of Genotype and Environment on Variation in DNA Methylation. In Patel, V. and Preedy, V. (eds), *Handbook of Nutrition, Diet, and Epigenetics*. Springer, Cham, pp. 1–20.

30. Jones, P.A. (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.*, **13**, 484–492.

31. Zhang, L., Xie, W.J., Liu, S., Meng, L., Gu, C. and Gao, Y.Q. (2017) DNA methylation landscape reflects the spatial organization of chromatin in different cells. *Biophys. J.*, **113**, 1395–1404.

32. Jin, Z. and Liu, Y. (2018) DNA methylation in human diseases. *Genes Dis.*, **5**, 1–8.

33. Lappalainen, T. and Greally, J. (2017) Associating cellular epigenetic models with human phenotypes. *Nat. Rev. Genet.*, **18**, 441–451.

34. Landgrave-Gómez, J., Mercado-Gómez, R. and Guevera-Guzmán, R. (2015) Epigenetic mechanisms in neurological and neurodegenerative diseases. *Front. Cell. Neurosci.*, **9**, 58.

35. Fransquet, P.D., Lacaze, P., Saffery, R., McNeil, J., Woods, R. and Ryan, J. (2018) Blood DNA methylation as a potential biomarker of dementia: a systematic review. *Alzheimers Dement.*, **14**, 81–103.

36. Garagnani, P., Bacalini, M.G., Pirazzini, C., Gori, D., Giuliani, C., Mari, D., Di Blasio, A.M., Gentilini, D., Vitale, G., Collino, G. *et al.* (2012) Methylation of ELOV2 gene as a new epigenetic marker of age. *Aging Cell*, **11**, 1132–1134.

37. Shenker, N.S., Ueland, P.M., Polidoro, S., van Veldhoven, K., Ricceri, F., Brown, R., Flanagan, J.M. and Vineis, P. (2013) DNA methylation as a long-term biomarker of exposure to tobacco smoke. *Epidemiology*, **24**, 712–716.

38. Levenson, V.V. (2010) DNA methylation as a universal biomarker. *Expert Rev. Mol. Diagn.*, **10**, 481–488.

39. Salameh, Y., Bejaoui, Y. and El Hajj, N. (2020) DNA methylation biomarkers in aging and age-related diseases. *Front. Genet.*, **11**, 171.

40. Horvath, S. (2013) DNA methylation age of human tissues and cell types. *Genome Biol.*, **14**, R115.

41. Levine, M.E., Lu, A.T., Quach, A., Chen, B.H., Assimes, T., Bandinelli, S., Hou, L., Baccarelli, A.A., Stewart, J.D. and Li, Y., Whitsel E.A., Wilson J.G., Reiner A.P., Aviv A., Lohman K., Liu Y., Ferrucci L., Horvath S. (2018) An epigenetic biomarker of aging for lifespan and healthspan. *Aging*, **10**, 276162, 573, 591.

42. Kinde, B., Gabel, H.W., Gilbert, C.S., Griffith, E.C. and Greenberg, M.E. (2015) Reading the unique DNA methylation landscape of the brain: non-CpG methylation, hydroxymethylation, and MeCP2. *Proc. Natl. Acad. Sci. U. S. A.*, **112**, 6800–6806.

43. Bachman, M., Uribe-Lewis, S., Yang, X., Williams, M., Murrell, A. and Balasubramanian, S. (2014) 5-Hydroxymethylcytosine is a predominantly stable DNA modification. *Nat. Chem.*, **6**, 1049–1055.

44. Szulwach, K.E., Li, X., Li, Y., Song, C., Wu, H., Dai, Q., Irier, H., Upadhyay, A.K., Gearing, M., Levery, A.I. *et al.* (2012) 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. *Nat. Neurosci.*, **14**, 1607–1616.

45. Kumar, S., Chinnusamy, V. and Mohapatra, T. (2018) Epigenetics of modified DNA bases: 5-methylcytosine and beyond. *Front. Genet.*, **9**, 640.

46. Nestor, C., Ruzov, A., Meehan, R.R. and Dunican, D.S. (2010) Enzymatic approaches and bisulfite sequencing cannot distinguish between 5-methylcytosine and 5-hydroxymethylcytosine in DNA. *Biotechniques*, **48**, 317–319.

47. Bak, S.T., Staunstrup, N.H., Starnawska, A., Daugaard, T.F., Nyengaard, J.R., Nyegaard, M., Børglum, A., Mors, O., Dorph-Petersen, K. and Nielsen, A.L. (2016) Evaluating the feasibility of DNA methylation analyses using long-term archived brain formalin-fixed paraffin-embedded samples. *Mol. Neurobiol.*, **55**, 668–681.

48. Kaut, O., Kuchelmeister, K., Moehl, C. and Wüllner, U. (2019) 5-Methylcytosine and 5-hydroxymethylcytosine in brains of patients with multiple system atrophy and patients with Parkinson's disease. *J. Chem. Neuroanat.*, **96**, 41–48.

49. Marshall, L.L., Killinger, B.A., Ensink, E., Li, P., Li, K.X., Lubben, N., Weiland, M., Wang, X., Gordevicius, J., Coetzee, G.A. et al. (2020) Epigenomic analysis of Parkinson's disease neurons identifies Tet2 loss as neuroprotective. *Nat. Neurosci.*, **23**, 1203–1214.

50. Spiers, H., Hannon, E., Schalkwyk, L.C., Bray, N.J. and Mill, J. (2017) 5-Hydroxymethylcytosine is highly dynamic across human fetal brain development. *BMC Genomics*, **18**, 1–14.

51. Cheung, W.A., Shao, X., Morin, A., Siroux, V., Kwan, T., Ge, B., Aïssi, D., Chen, L., Vasquez, L., Allum, F. et al. (2017) Functional variation in allelic methylomes underscores a strong genetic contribution and reveals novel epigenetic alterations in the human epigenome. *Genome Biol.*, **18**, 1–21.

52. Klein, C. and Westenberger, A. (2012) Genetics of Parkinson's disease. *C. S. H. Persp. Med.*, **2**, a008888.

53. International Parkinson's Disease Genomics Consortium and Wellcome Trust Case Control Consortium (2011) A two-stage meta-analysis identifies several new loci for Parkinson's disease. *PLoS Genet.*, **7**, e1002142.

54. Nalls, M.A., Pankratz, C., Lill, C., Do, C.B., Hernandez, D.G., Saad, M., DeStefano, A.L., Kara, E., Bras, J., Sharma, M. et al. (2014) Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat. Genet.*, **46**, 989–993.

55. Kia, D.A., Zhang, D., Guelfi, S., Manzoni, C., Hubbard, L., Reynolds, R.H., Botía, J., Ryten, M., Ferrari, R., Lewis, P.A. et al. (2021) Identification of candidate Parkinson disease genes by integrating genome-wide association study, expression, and epigenetic data sets. *JAMA Neurol.*, **78**, 464–472.

56. Jowaed, A., Schmitt, I., Kaut, O. and Wüllner, U. (2010) Methylation regulates alpha-synuclein expression and is decreased in Parkinson's disease patients' brains. *J. Neurosci.*, **30**, 6355–6359.

57. Scherzer, C.R., Grass, J.A., Liao, Z., Pepivani, I., Zheng, B., Eklund, A.C., Ney, P.A., Ng, J., McGoldrick, M., Mollenhauer, B., Bresnick, E.H. and Schlossmacher, M.G. (2008) GATA transcription factors directly regulate the Parkinson's disease-linked gene α-synuclein. *Proc. Natl. Acad. Sci. U. S. A.*, **105**, 10907–10912.

58. Gatev, E., Gladish, N., Mostafavi, S. and Kobor, M.S. (2020) CoMeBack: DNA methylation array data analysis for co-methylated regions. *Bioinformatics*, **36**, 2675–2683.

59. Gillis, J., Mistry, M. and Pavlidis, P. (2010) Gene function analysis in complex data sets using ErmineJ. *Nat. Protoc.*, **5**, 1148–1159.

60. Mellen, M., Pinar, A. and Heintz, N. (2017) 5-Hydroxymethylcytosine accumulation in post-mitotic neurons results in functional demethylation of expressed genes. *Proc. Natl. Acad. Sci. U. S. A.*, **114**, E7812–E7821.

61. Wijetunga, N.A., Johnston, A.D., Maekawa, R., Delahaye, F., Ula-hannan, N., Kim, K. and Greally, J.M. (2017) SMITE: an R/Bioconductor package that identifies network modules by integrating genomic and epigenomic information. *BMC Bioinform.*, **18**, 41.

62. Peng, X., Tehranian, R., Dietrich, P., Stefanis, L. and Perez, R. (2005) Alpha-synuclein activation of protein phosphatase 2A reduces tyrosine hydroxylase phosphorylation in dopaminergic cells. *J. Cell Sci.*, **118**, 3523–3530.

63. Beninger, R.J. (1983) The role of dopamine in locomotor activity and learning. *Brain Res.*, **287**, 173–196.

64. Oliveira, L.M.A., Falmoir-Lockhart, L.J., Borelho, M.G., Lin, K., Wales, P., Koch, J.C., Gerhardt, E., Taschenberger, H., Outeiro, T.F., Lingor, P. et al. (2015) Elevated α-synuclein caused by SNCA gene triplication impairs neuronal differentiation and maturation in Parkinson's patient-derived induced pluripotent stem cells. *Cell Death Dis.*, **6**, e1994.

65. Dong, X., Wang, Y. and Qin, Z. (2009) Molecular mechanisms of excitotoxicity and their relevance to pathogenesis of neurodegenerative diseases. *Acta Pharmacol. Sin.*, **30**, 379–387.

66. Stadler, F., Kolb, G., Rubush, L., Baker, S.P., Jones, E.G. and Akbarian, S. (2005) Histone methylation at gene promoters is associated with developmental regulation and region-specific expression of ionotropic and metabotropic glutamate receptors in human brain. *J. Neurochem.*, **94**, 324–336.

67. Pinheiro, P.S. and Mulle, C. (2008) Presynaptic glutamate receptors: physiological functions and mechanisms of action. *Nat. Rev. Neurosci.*, **9**, 423–436.

68. Johnson, K.A., Conn, P.J. and Niswender, C.M. (2009) Glutamate receptors as therapeutic targets for Parkinson's disease. *CNS Neurol. Disord. Drug Targets*, **8**, 475–491.

69. Cox, P.A., Davis, D.A., Mash, D.C., Metcalf, J.S. and Banack, S.A. (2016) Dietary exposure to an environmental toxin triggers neurofibrillary tangles and amyloid deposits in the brain. *Proc. Biol. Sci.*, **283**, 20152397.

70. Chen, S., Wang, Z., Ma, Y., Zhang, W., Lu, J., Liang, Y. and Zheng, X. (2018) Neuroprotective effects and mechanisms of tea bioactive components in neurodegenerative diseases. *Molecules*, **23**, 512.

71. Trudler, D., Sanz-Blasco, S., Eisele, Y.S., Ghatak, S., Bodhinathan, K., Akhtar, M.W., Lynch, W.P., Piña-Crespo, J.C., Talantova, M., Kelly, J.W. and Lipton, S.A. (2021) α-Synuclein oligomers induce glutamate release from astrocytes and excessive extrasynaptic NMDAR activity in neurons, thus contributing to synapse loss. *J. Neurosci.*, **41**, 2264–2273.

72. Lister, E., Mukamel, E., Nery, J., Urich, M., Puddifoot, C.A., Johnson, N.D., Lucero, J., Huang, Y., Dwork, A.J., Schultz, M.D. et al. (2013) Global epigenomic reconfiguration during mammalian brain development. *Science*, **341**, 1237905.

73. Farré, P., Jones, M.J., Meaney, M.J., Emberly, E., Turecki, G. and Kobor, M.S. (2015) Concordant and discordant DNA methylation signatures of aging in human blood and brain. *Epigenetics Chromatin*, **8**, 19.

74. Bernstein, C. and Bernstein, H. (2019) The Role of DNA Repair and the Epigenetic Markers Left after Repair in Neurologic Functions, Including Memory and Learning. In Mognato, M. (ed.), *DNA Repair - An Update*. IntechOpen, London, UK. https://doi.org/10.5772/intechopen.83477.

75. Huang, Y., Gu, L. and Li, G. (2018) H3K36me3-mediated mismatch repair preferentially protects actively transcribed genes from mutation. *J. Biol. Chem.*, **293**, 7811–7823.

76. Kontopoulos, E., Parvin, J.D. and Feaney, M.B. (2006) Alpha-synuclein acts in the nucleus to inhibit histone acetylation and promote neurotoxicity. *Hum. Mol. Genet.*, **15**, 3012–3023.

77. Desplats, P., Spencer, B., Coffee, E., Patel, P., Michael, S., Patrick, C., Adame, A., Rockenstein, E. and Masliah, E. (2011) Alpha-synuclein sequesters Dnmt1 from the nucleus: a novel mechanism for epigenetic alterations in Lewy body diseases. *J. Biol. Chem.*, **286**, 9031–9037.

78. Davidi, D., Schechter, M., Elhadi, S.A., Mataov, A., Nathanson, L. and Sharon, R. (2020) α-Synuclein translocates to the nucleus to activate retinoic-acid-dependent gene transcription. **23**, 100910.

79. Irizarry, R.A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., Cui, H., Gabo, K., Ronigione, M., Webster, M. *et al.* (2009) The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat. Genet.*, **41**, 178–186.

80. Scholz, D., Pöltl, D., Genewsky, A., Weng, M., Waldmann, T., Schildknecht, S. and Leist, M. (2011) Rapid, complete and large-scale generation of post-mitotic neurons from the human LUHMES cell line. *J. Neurochem.*, **119**, 957–971.

81. Lea, A.J., Vockley, C.M., Johnston, R.A., Del Carpio, C.A., Barreiro, L.B., Reddy, T.E. and Tung, J. (2018) Genome-wide quantification of the effects of DNA methylation on human gene regulation. **7**, e37513.

82. Lam, L.L., Emberly, E., Fraser, H.B., Neumann, S.M., Chen, E., Miller, G.E. and Kobor, M.S. (2012) Factors underlying variable DNA methylation in a human community cohort. *Proc. Natl. Acad. Sci. U. S. A.*, **109**, 17253–17260.

83. Reynolds, L.M., Taylor, J.R., Ding, J., Lohman, K., Johnson, C., Siscovick, D., Burke, G., Post, W., Shea, S., Jacobs, D.R., Jr. *et al.* (2014) Age-related variations in the methylome associated with gene expression in human monocytes and T cells. *Nat. Commun.*, **5**, 5366.

84. Chen, C., Zhang, C., Cheng, L., Reilly, J.L., Bishop, J.R., Sweeney, J.A., Chen, H., Gershon, E.S. and Liu, C. (2014) Correlation between DNA methylation and gene expression in the brains of patients with bipolar disorder and schizophrenia. *Bipolar Disord.*, **16**, 790–799.

85. Price, A.J., Collado-Torres, L., Ivanov, N.A., Xia, W., Burke, E.E., Shin, J.H., Tao, R., Ma, L., Jia, Y., Hyde, T.M. *et al.* (2019) Divergent neuronal DNA methylation patterns across human cortical development reveal critical periods and a unique role of CpH methylation. *Genome Biol.*, **20**, 196.

86. Kantor, B., Tagliafierro, L., Gu, J., Zamora, M.E., Ilich, E., Grenier, C., Huang, Z.Y., Murphy, S. and Chiba-Falek, O. (2018) Down-regulation of *SNCA* expression by targeted editing of DNA methylation: a potential strategy for precision therapy in PD. *Mol. Ther.*, **26**, 2638–2649.

87. Provençal, N., Arloth, J., Cattaneo, A., Anacker, C., Cattane, N., Wiechmann, E., Röh, S., Ködel, M., Klengel, T., Czamara, D. *et al.* (2020) Glucocorticoid exposure during hippocampal neurogenesis primes future stress response by inducing changes in DNA methylation. *Proc. Natl. Acad. Sci. U. S. A.*, **117**, 23280–23285.

88. Weber, M., Davies, J.J., Wittig, D., Oakeley, E.J., Haase, M., Lam, W.L. and Schübeler, D. (2005) Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet.*, **8**, 853–862.

89. de Boni, L., Gasparoni, G., Haubenreich, C., Tierling, S., Schmitt, I., Peitz, M., Koch, P., Walter, J., Wüllner, U. and Brüstle, O. (2018) DNA methylation alterations in iPSC- and hESC-derived neurons: potential implications for neurological disease modeling. *Clin. Epigenetics*, **10**, 13.

90. de Boni, L. and Wüllner, U. (2019) Epigenetic analysis in human neurons: considerations for disease modeling in PD. *Front. Neurosci.*, **13**, 276.

91. Lunnon, K., Hannon, E., Smith, R.G., Dempster, E., Wong, C., Burrage, J., Troakes, C., Al-Sarraj, S., Kepa, A., Schalkwyk, L. *et al.* (2016) Variation in 5-hydroxymethylcytosine across human cortex and cerebellum. *Genome Biol.*, **17**, 27.

92. Reyes, J.F., Olsson, T.T., Lamberts, J.T., Devine, M.J., Kunath, T. and Brundin, P. (2015) A cell culture model for monitoring alpha-synuclein cell-to-cell transfer. *Neurobiol. Dis.*, **77**, 266–275.

93. Zeng, X., Geng, W.S., Jia, J.J., Chen, L. and Zhang, P.P. (2018) Cellular and molecular basis of neurodegeneration in Parkinson disease. *Front. Aging Neurosci.*, **10**, 109.

94. Flierl, A., Oliveria, L.M.A., Falomir-Lockhart, L.J., Mak, S.K., Hesley, J., Soldner, F., Arndt-Jovin, D.J., Jaenisch, R., Langston, J.W., Jovin, T.M. *et al.* (2014) Higher vulnerability and stress sensitivity of neuronal precursor cells carrying an alpha-synuclein gene triplication. *PLoS One*, **9**, e112413.

95. Paiva, I., Jain, G., Lázaro, D.F., Jerčic, K.G., Hentrich, T., Kerimoglu, C., Pinho, R., Szego, E.M., Burkhardt, S., Capece, V. *et al.* (2018) Alpha-synuclein deregulates the expression of COL4A2 and impairs ER-Golgi function. *Neurobiol. Dis.*, **119**, 121–135.

96. Zufferey, R., Dull, T., Mandel, R.J., Bukovsky, A., Quiroz, D., Naldini, L. and Trono, D. (1998) Self-inactivating lentivirus vector for safe and efficient *in vivo* gene delivery. *J. Virol.*, **72**, 9873–9880.

97. Lizée, G., Aerts, J.L., Gonzales, M.I., Chinnasamy, N., Morgan, R.A. and Topalian, S.L. (2003) Real-time quantitative reverse transcriptase-polymerase chain reaction as a method for determining lentiviral vector titers and measuring transgene expression. *Hum. Gene Ther.*, **14**, 497–507.

98. Pidsley, R., Wong, C.C.Y., Volta, M., Lunnon, K., Mill, J. and Schalkwyk, L.C. (2013) A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics*, **14**, 293.

99. Pidsley, R., Zotenko, E., Peters, T.J., Lawrence, M.G., Risbridger, G.P., Molloy, P., Van Djik, S., Muhlhausler, B., Stirzaker, C. and Clark, S.J. (2016) Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol.*, **17**, 208.

100. Leek, J.T., Johnson, W.E., Parker, H.S., Jaffe, A.E. and Storey, J.D. (2012) The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*, **28**, 882–883.

101. Ritchie, R.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W. and Smyth, G.K. (2015) Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.*, **43**, e47.

102. Gu, H., Smith, Z.D., Bock, C., Boyle, P., Gnirke, A. and Meissner, A. (2011) Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nat. Protoc.*, **6**, 468–481.

103. Xi, Y. and Li, W. (2009) BSMAP: whole genome bisulfite sequence MAPping program. *BMC Bioinform.*, **10**, 232.

104. Hebestreit, K., Dugas, M. and Klein, H. (2013) Detection of significantly differentially methylated regions in targeted bisulfite sequencing data. *Bioinformatics*, **29**, 1647–1653.

105. Halder, R., Hennion, M., Vidal, R.O., Shomroni, O., Rahman, R., Rajput, A., Centeno, T.P., van Bebber, F., Capece, V., GarciaVizcaino, J.C. *et al.* (2016) DNA methylation changes in plasticity genes accompany the formation and maintenance of memory. *Nat. Neurosci.*, **19**, 102–110.

106. Liao, Y., Smyth, G.K. and Shi, W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30**, 923–930.

107. Love, M.I., Huber, W. and Ander, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.

108. Patro, R., Duggal, G., Love, M.I., Irizarry, R.A. and Kingsford, C. (2017) Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods*, **14**, 417–419.

109. Love, M.I., Soneson, C., Hickey, P.F., Johnson, L.K., Pierce, N.T., Shepherd, L., Morgan, M. and Patro, R. (2020) Tximeta: reference sequence checksums for provenance identification in RNA-seq. *PLoS Comput. Biol.*, **16**, e1007664.

110. Zhu, A., Srivastava, A., Ibrahim, J.G., Patro, R. and Love, M.I. (2019) Nonparametric expression analysis using inferential replicate counts. *Nucleic Acids Res.*, **47**, e105.

111. Langmead, B. and Salzberg, S. (2012) Fast gapped-read alignment with bowtie 2. *Nat. Methods*, **9**, 357–359.

112. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W. and Liu, X.S. (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.