

## ARTICLE

DOI: [10.1038/s41467-018-05325-y](https://doi.org/10.1038/s41467-018-05325-y)

OPEN

# Epigenome-wide DNA methylation profiling in Progressive Supranuclear Palsy reveals major changes at *DLX1*

Axel Weber <sup>1</sup>, Sigrid C. Schwarz<sup>2,3</sup>, Jörg Tost <sup>4</sup>, Dietrich Trümbach<sup>5</sup>, Pia Winter<sup>1</sup>, Florence Busato<sup>4</sup>, Pawel Tacik<sup>6,7</sup>, Anita C. Windhorst<sup>8</sup>, Maud Fagny<sup>4</sup>, Thomas Arzberger<sup>3,9,10</sup>, Catriona McLean<sup>11</sup>, John C. van Swieten<sup>12</sup>, Johannes Schwarz<sup>2</sup>, Daniela Vogt Weisenhorn<sup>3,5,13</sup>, Wolfgang Wurst<sup>3,5,13,14</sup>, Till Adhikary<sup>15</sup>, Dennis W. Dickson <sup>6</sup>, Günter U. Höglinger <sup>2,3,14</sup> & Ulrich Müller <sup>1</sup>

Genetic, epigenetic, and environmental factors contribute to the multifactorial disorder progressive supranuclear palsy (PSP). Here, we study epigenetic changes by genome-wide analysis of DNA from postmortem tissue of forebrains of patients and controls and detect significant ( $P < 0.05$ ) methylation differences at 717 CpG sites in PSP vs. controls. Four-hundred fifty-one of these sites are associated with protein-coding genes. While differential methylation only affects a few sites in most genes, *DLX1* is hypermethylated at multiple sites. Expression of an antisense transcript of *DLX1*, *DLX1AS*, is reduced in PSP brains. The amount of DLX1 protein is increased in gray matter of PSP forebrains. Pathway analysis suggests that DLX1 influences *MAPT*-encoded Tau protein. In a cell system, overexpression of *DLX1* results in downregulation of *MAPT* while overexpression of *DLX1AS* causes upregulation of *MAPT*. Our observations suggest that altered *DLX1* methylation and expression contribute to pathogenesis of PSP by influencing *MAPT*.

<sup>1</sup>Institute of Human Genetics, Justus-Liebig-Universität, Gießen 35392, Germany. <sup>2</sup>Department of Neurology, Technische Universität München, Munich 81377, Germany. <sup>3</sup>German Center for Neurodegenerative Diseases (DZNE), Munich 81377, Germany. <sup>4</sup>Laboratory for Epigenetics and Environment, Centre National de Recherche en Génomique Humaine, CEA—Institut de Biologie Francois Jacob, Evry 91000, France. <sup>5</sup>Institute of Developmental Genetics, Helmholtz Center München, Munich 85764, Germany. <sup>6</sup>Department of Neuroscience, Mayo Clinic, Jacksonville, FL, 32224, USA. <sup>7</sup>Department of Neurodegenerative Diseases and Geriatric Psychiatry, University of Bonn Medical Center, Bonn 53127, Germany. <sup>8</sup>Institute of Medical Informatics, Justus-Liebig-Universität, Gießen 35392, Germany. <sup>9</sup>Department of Psychiatry, Ludwig-Maximilians-Universität, Munich 81377, Germany. <sup>10</sup>Center for Neuropathology and Prion Research, Ludwig-Maximilians-Universität, Munich 81377, Germany. <sup>11</sup>Alfred Anatomical Pathology and NNF, Victorian Brain Bank, Carlton, VIC, 3053, Australia. <sup>12</sup>Department of Neurology, Erasmus Medical Centre, Rotterdam 3000, The Netherlands. <sup>13</sup>Chair of Developmental Genetics, Technische Universität München-Weihenstephan, Neuherberg/Munich 85764, Germany. <sup>14</sup>Munich Cluster for Systems Neurology (SyNergy), Munich 81377, Germany. <sup>15</sup>Institute for Molecular Biology and Tumor Research, Center for Tumor Biology and Immunology, Philipps University, Marburg 35043, Germany. These authors contributed equally: Axel Weber, Sigrid C. Schwarz. These authors jointly supervised this work: Günter U. Höglinger, Ulrich Müller. Correspondence and requests for materials should be addressed to A.W. (email: [axel.weber@humangenetik.med.uni-giessen.de](mailto:axel.weber@humangenetik.med.uni-giessen.de)) or to G.U.H. (email: [Gunter.Hoeglinger@dzne.de](mailto:Gunter.Hoeglinger@dzne.de)) or to U.M. (email: [Ulrich.mueller@med.uni-giessen.de](mailto:Ulrich.mueller@med.uni-giessen.de))

**P**rogressive supranuclear palsy (PSP) is a progressive and fatal neurodegenerative disease with a prevalence of 5–/100,000<sup>1,2</sup>. Disease onset is usually beyond 60 years of age and the average survival time is 6–7 years after onset<sup>1</sup>. Symptoms include ocular motor dysfunction, postural instability, akinesia, cognitive dysfunction, and dysphagia, with the latter being the most frequent cause of death in PSP<sup>1,2</sup>.

PSP is neuropathologically defined by intracellular aggregation of the microtubule-associated protein Tau in neurofibrillary tangles and tufted astrocytes. The aggregates eventually cause neuronal cell death in the cerebral cortex, diencephalon, brainstem, and cerebellar nuclei<sup>3</sup>.

PSP is a “complex” disorder; genetic, environmental and epigenetic modifications contribute to disease. A variant of the gene *MAPT* is the major genetic risk factor in PSP<sup>4,5</sup>. Variants of the genes *STX6*, *EIF2AK3*, and *MOBP* also increase the risk of PSP<sup>5</sup>. Among environmental factors, advanced age is the best established risk factor<sup>6</sup>. Epigenetic modifications reported so far in PSP include aberrant DNA methylation at the *MAPT* locus<sup>7–9</sup> and miRNA dysregulation<sup>10,11</sup>.

In order to learn more about the possible relevance of epigenetic changes in PSP we set out to study epigenetic alterations at the DNA level in prefrontal lobe tissue of PSP patients. We describe significant DNA methylation differences between patients and controls at many CpG sites, amounting to 451 protein-coding genes. While methylation differences only affect one or a few sites at most genes, highly significant ( $\geq 5\%$ ) hypermethylation is found at multiple sites associated with the gene *DLX1*. Functional analyses of both *DLX1* and its antisense transcript *DLX1AS* are consistent with an important role of *DLX1* in the pathogenesis of PSP.

## Results

**Differentially methylated sites in PSP.** The genome-wide DNA methylation patterns of 94 PSP patients ( $72 \pm 5.3$  years; 57% male, 43% female) were compared to 71 controls ( $76 \pm 7.9$  years; 67% male, 33% female) without neurological or psychiatric diseases (Supplementary Data 1). We studied prefrontal lobe tissue since it is consistently pathologically damaged in PSP, but less so than other brain regions<sup>3</sup>. We estimated the amount of neuronal and non-neuronal cells in our samples, as described by Guintivano et al.<sup>12</sup>. The percentage of neurons in PSP patients (median 36.1% of cells) did not significantly differ from the proportion of neuronal cells in controls (median 38.0% of cells; Wilcoxon test,  $P = 0.31$ ) (Supplementary Fig. 1). Thus, we detected disease-specific alterations and minimized a potential bias by a massive change in regional cellular composition due to severe neurodegeneration and gliosis.

Methylation differences at CpG sites between patients and controls were analyzed on 450 K BeadChips<sup>13</sup> applying a linear regression model with age, sex, and non-neuronal cell content as covariates at a Benjamini-Hochberg<sup>14</sup> corrected level of significance of  $P < 0.05$ . Significant CpGs previously shown to be influenced by genetic variants (mQTLs) in adult prefrontal cortex<sup>15</sup> were not included in further analyses. However, only three CpGs were found to match this criterion, i.e., cg01378667, cg0325535, cg10318222 in the promoter region of *GABRA5*, all of which are associated in cis with rs7496866 located in the same region. An influence of presently unknown genetic variants on the methylation pattern cannot be excluded. It is also not possible to correct for environmental factors such as individual medications and/or accompanying neurological diseases that might affect DNA methylation in the forebrain.

After these corrections, significant methylation differences were detected at 717 sites (627 hyper-, 90 hypomethylated). Mean

differences of  $\geq 5\%$  were found at 38 of these sites (34 hyper-, 4 hypomethylated) (Fig. 1a and Supplementary Data 2). Of the hypomethylated sites, 70% were associated with protein-coding genes, 4% with genes for non-coding RNAs (miRNAs, lncRNAs, etc.), and 26% were located beyond 1.5 kb of genes. The respective percentages for hypermethylated genes were 62, 3.5, and 34.5% (Fig. 1b).

The percentage of hypomethylated CpG sites within gene bodies was 40% for hypo- and 28% for hypermethylated sites (Fig. 1c). Twelve percent of hypo- and 8% of hypermethylated CpG sites were located within 5'UTRs (Fig. 1c). Seven percent hypo- and 2% hypermethylated sites were within the first exons of the respective genes.

Among sites hypomethylated in patients 44% were CpG islands and among hypermethylated sites 25% represented CpG islands. Thirty-four percent CpG sites were isolated (“open sea”<sup>13</sup>) and 31% of hypermethylated sites were in “open sea” regions. The percentages of hypomethylated CpG sites in “shelves” (2–4 kb from CpG island) was 6%, and that in “shores” (up to 2 kb from a CpG island<sup>13</sup>) was 17%. The corresponding percentages for hypermethylated sites are 8 and 36% (Fig. 1d).

### Chromosomal location of differentially methylated genes.

Figure 2 depicts the 375 genes with significant methylation differences between patients and controls and highlights the ten genes with differences  $\geq 5\%$  (see also Supplementary Data 2). Eight of these genes are hyper- and two are hypomethylated.

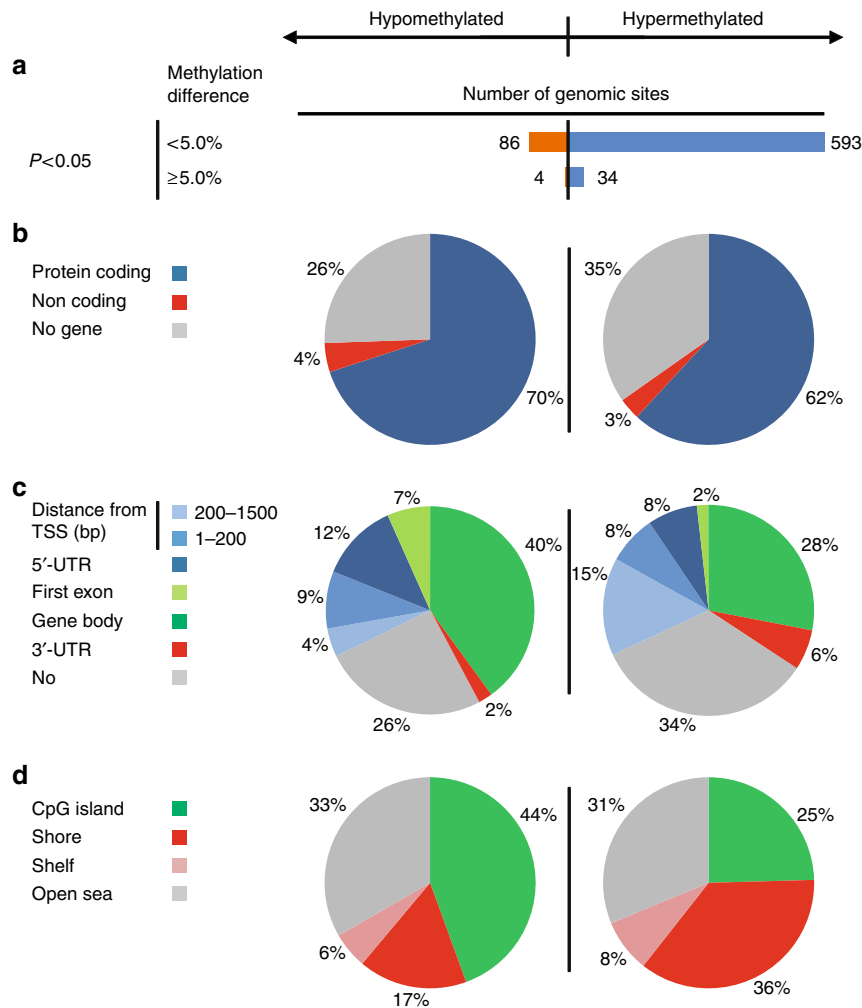
In order to check validity of the BeadChip-based primary results we analyzed the methylation status at selected loci by pyrosequencing of bisulfite-converted DNA of the same samples. This confirmed the methylation differences in a representative subset of six genes, i.e., *DLX1/DLX1AS*, *DLX2*, *METAP1D*, *SLC15A3*, *SLIT1*, and *TRRAP* (Fig. 3c and Supplementary Fig. 2).

**No significant differential methylation of *MAPT*.** Analysis of the region 17q21.31 that spans 1.6 Mb (43,470,000 to 45,061,000 on chromosome 17) and includes *MAPT* revealed nominally significant methylation differences at 11 sites. However, the findings were not significant after correction for multiple testing. The smallest  $P$ -value corrected for multiple testing within *MAPT* was 0.0578 at chr17:44026659 (Supplementary Data 3).

**Pronounced hypermethylation of *DLX1*.** Differential methylation of  $\geq 5\%$  was only detected at a few CpG sites in a small number of genes in PSP (Fig. 2 and Supplementary Data 2). Most pronounced hypermethylation was detected at a region of chromosome 2 that includes the gene *DLX1* (Distal-Less Homeobox 1). Many sites of *DLX1*, mainly within its 3'UTR, were hypermethylated by  $\geq 5\%$ , as shown in a representative heat-map of 11 sites (Fig. 3a).

The *DLX1* gene is composed of three exons<sup>16</sup>. Greatest methylation differences were found at a CpG island (i.e., a genomic region of  $> 200$  bp with a CG content of  $> 50\%$  and an observed/expected CpG ratio of  $> 60\%$ ) in the 3'UTR, spanning positions 172952810–172953160 [hg19] on chromosome 2 (Fig. 3b). Pyrosequencing confirmed hypermethylation of nine CpGs within the CpG island that is located in the 3'UTR of *DLX1* (Fig. 3c).

***DLX1* transcript.** We proceeded to test the level of transcription of *DLX1* by reverse transcription quantitative PCR (RT-qPCR). Expression of the *DLX1* sense transcript did not correlate with *DLX1* methylation and did not significantly differ in forebrains between patients and controls (Fig. 4a).



**Fig. 1** Epigenome-wide methylation analysis. **a–d** CpG sites hypo- and hypermethylated in  $n = 94$  PSP patients vs.  $n = 71$  controls are displayed on the left and right panels, respectively. **a** Number of genomic sites aberrantly methylated at a difference of  $< 5\%$  and  $\geq 5\%$  between patients and controls (Benjamini-Hochberg corrected  $P < 0.05$ ). **b** Location of aberrantly methylated sites within protein-coding genes, non-coding genes and intergenic regions. **c** Location of aberrantly methylated CpG sites in relation to defined regions of genes (TSS: transcription start site; 5' and 3' UTR: 5' and 3' untranslated region). **d** Association of aberrantly methylated sites with CpG islands (shores are  $< 2$  kb and shelves are 2–4 kb from a CpG island)

Several reports indicate the existence of a *Dlx1* antisense transcript (*Dlx1as*) in the mouse<sup>17–19</sup>. We found such a *DLX1* antisense (*DLX1AS*) transcript in cDNA from human brain. Based on the mouse sequence we predicted a homologous sequence in human DNA in silico. We used primer walking from the putative transcription start site 3' of exon 3 of the sense transcript to different predicted *DLX1AS*-exons. We found several alternatively spliced transcripts of *DLX1AS* by sequencing different PCR products and were able to extend longest transcripts beyond exon 3 of *DLX1* (Fig. 3b and Supplementary Fig. 3). The hypermethylated CpG sites are located in the region of exon 3 of the *DLX1AS* gene. A recently described enhancer region of *DLX1* overlaps with exon 1 of *DLX1AS*<sup>19,20</sup>. For detection of the antisense transcript a region encoded by *DLX1AS* exon 1 was used since this exon is part of all splice variants of the gene (Supplementary Fig. 3).

Transcription of *DLX1AS* was significantly reduced in patients ( $P < 0.001$ ) and expression values inversely correlated with the degree of methylation ( $P < 0.001$ , Fig. 4b). Delta Ct values, calculated according to Pfaffl et al.<sup>21</sup>, shown in Fig. 4b correspond to a 0.64-fold expression of *DLX1AS* in PSP as compared to controls.

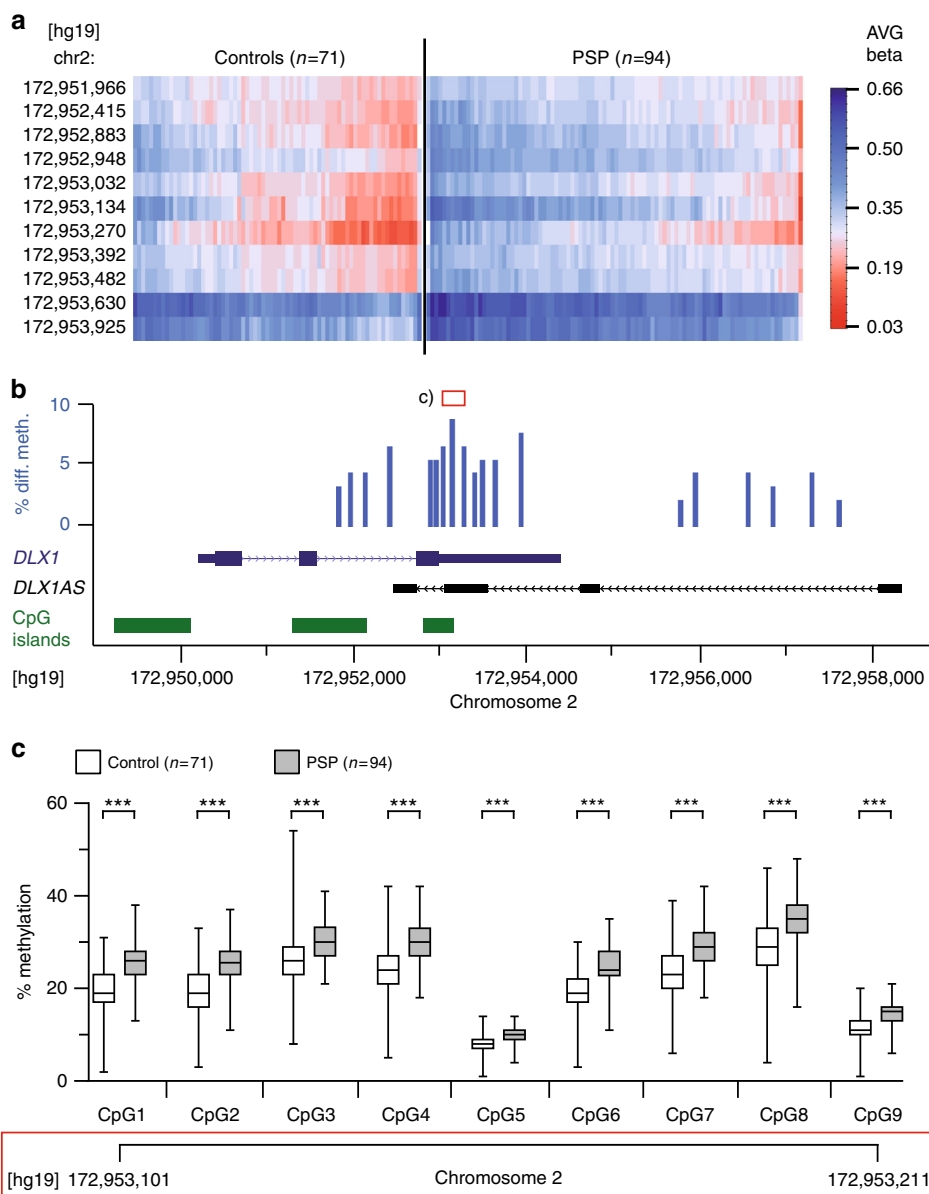
**Single-cell analysis in healthy human cortex.** *Dlx1* and *Dlx1as* are almost exclusively expressed in neuronal cells in the mouse<sup>22</sup> (Supplementary Fig. 4a). Using published RNA-sequencing data in single cells from healthy human cortex<sup>23</sup> we quantified the reads of *DLX1* and *DLX1AS* transcripts. Consistent with the mouse data, we found *DLX1* and *DLX1AS* expression mainly in neurons (Supplementary Fig. 4b, c). *DLX1* was expressed in 26.15%, *DLX1AS* in 15.39% and both were expressed in 2.31% of neurons (Supplementary Fig. 4d).

**DLX1 protein.** Expression of *DLX1* was analyzed at the protein level in total protein extracts of frontal lobes from PSP patients and controls.

Western blot analysis of eight PSP and eight control forebrains demonstrated that DLX1 protein levels do not differ between patients and controls in white matter (Fig. 4c and Supplementary Fig. 5). In gray matter, however, higher levels of DLX1 protein were detected in PSP as compared to controls (Fig. 4d and Supplementary Fig. 6).

The findings were confirmed by quantitative immunohistochemistry on histological sections of frontal lobes from an independent set of patients. Densitometry of DLX1-





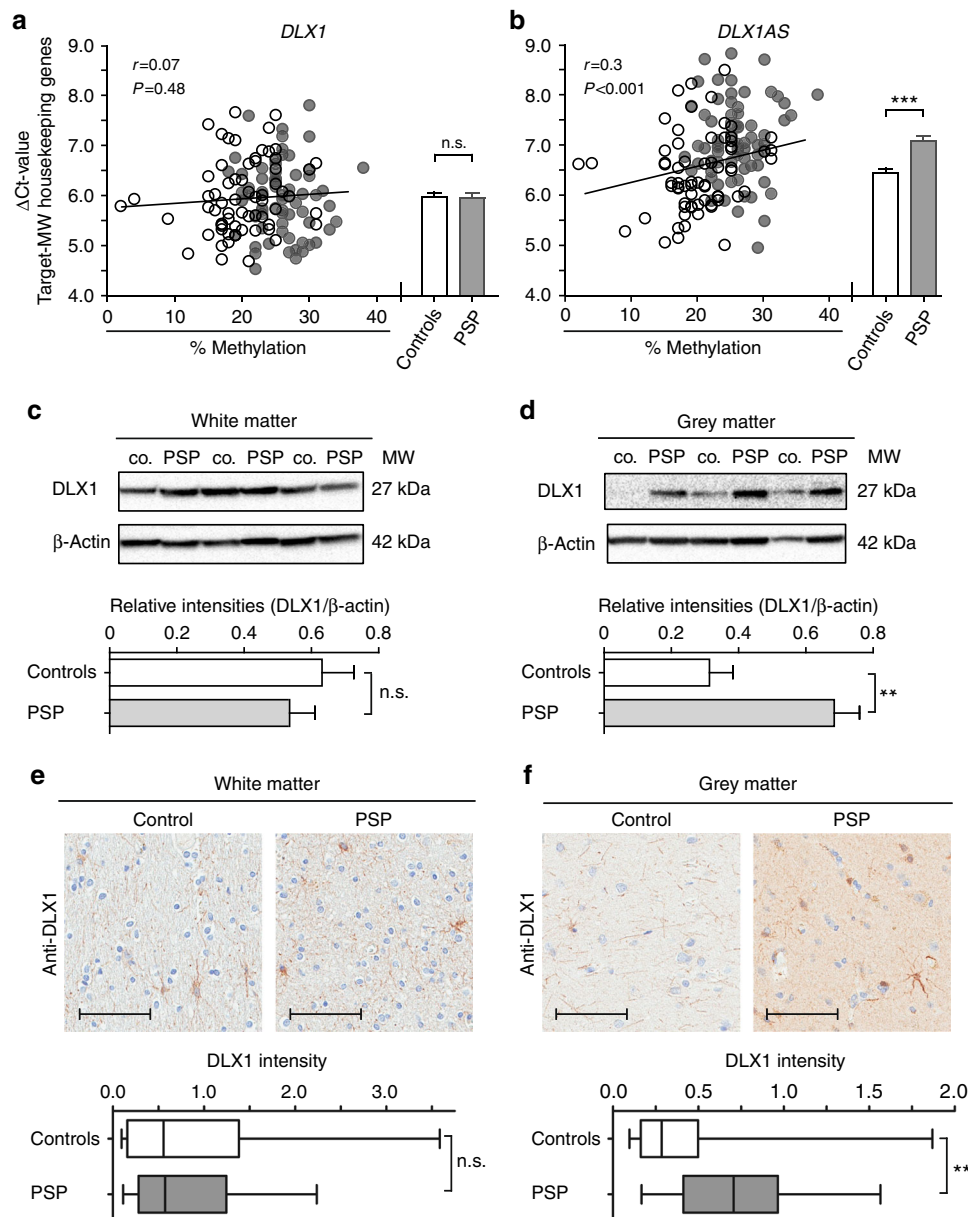
**Fig. 3** Methylation status of *DLX1*. **a** Heat-map showing degree of individual methylation at various genomic sites within *DLX1* on chromosome 2 (chr2) in forebrains of  $n = 94$  PSP patients vs.  $n = 71$  controls. Average (AVG) beta indicates the color-coded methylation value (1.00 equals 100%, Genome Studio Software Version 2011.1, Illumina, San Diego, CA). **b** *DLX1* is composed of three alternatively spliced exons (dark blue). *DLX1* antisense transcript (*DLX1AS*) is encoded by at least four alternatively spliced exons (black). The location of CpG islands relative to *DLX1* and *DLX1AS* is shown according to the UCSC genome browser data (green). The percentage difference in methylation in PSP as compared to controls at various sites within *DLX1* and *DLX1AS* is shown as bar chart (blue). **c** Pyrosequencing confirmed the differential methylation at nine CpGs within the CpG island of the 3' UTR of *DLX1* [red boxes in **b** and **c** indicate corresponding genomic regions; \*\*\*  $P < 0.001$ , Welch's corrected unpaired  $t$ -test]. The line in the middle of the box and whisker graph represents the median (50th percentile). The box extends from the 25th to 75th percentile. The whiskers extend from the lowest to the highest value

**Knock-down of *DLX1* and *DLX1AS* in human striatal NPC.** In order to test whether a putative down-stream effect of *DLX1*/*DLX1AS* on *MAPT* also occurs in neural precursor cells (NPC) derived from human fetal striatum (strNPC) we transfected these cells with siRNAs that target *DLX1* and *DLX1AS*. As shown in Fig. 5c knock-down of *DLX1* resulted in significant upregulation of *MAPT* and of *DLX1AS*. Conversely, knock-down of *DLX1AS* resulted in downregulation of *MAPT* and in upregulation of *DLX1*.

We proceeded to test whether *DLX1* affects Tau-dependent viability of strNPCs using the ATP firefly luciferase assay<sup>28</sup>. strNPCs overexpressing either 3R- or 4R-Tau were co-transfected with siRNAs directed against either *DLX1* or *DLX1AS*. As shown

in Fig. 5d, siRNA-mediated knock-down of *DLX1* decreased cellular viability, in particular of cells overexpressing 4R-Tau. Conversely, knock-down of *DLX1AS* significantly increased survival of strNPCs (Fig. 5d).

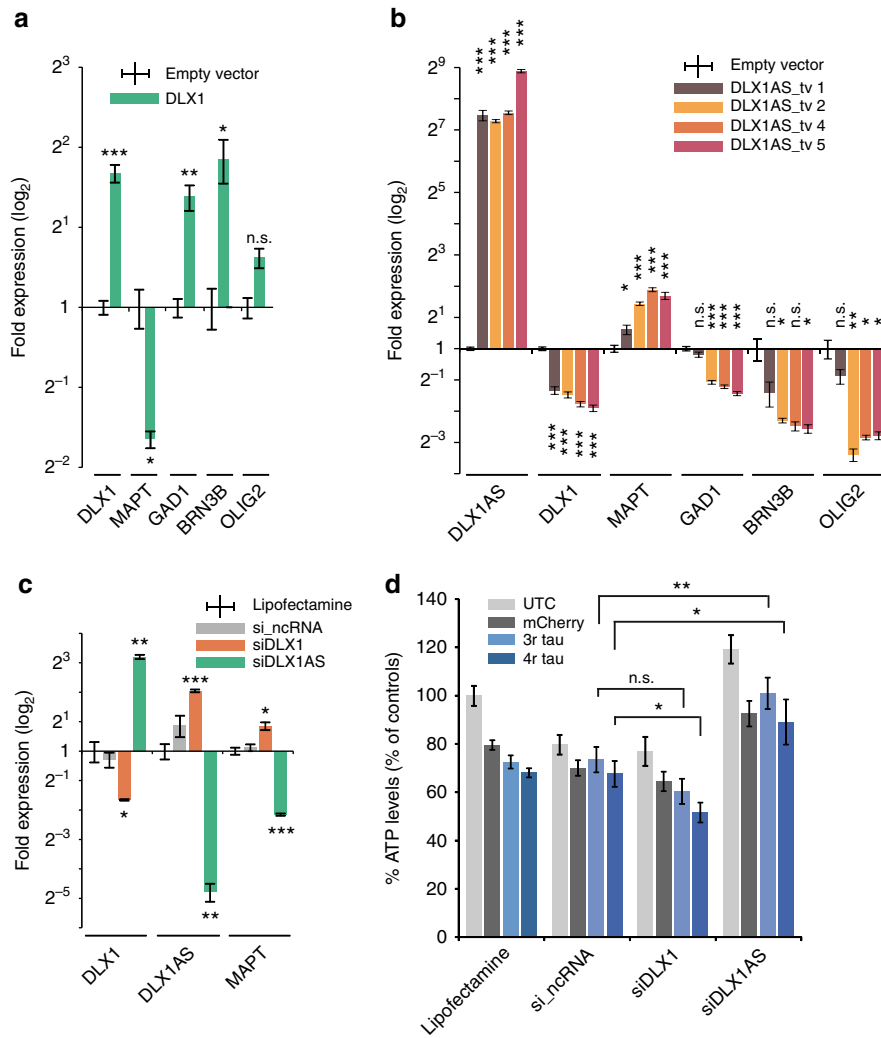
**Functional analysis of differentially methylated genes.** We performed in silico functional analyses of 375 different annotated genes that are represented by 451 (out of a total of 717) differentially methylated CpG sites on the Illumina 450 kb chip (Supplementary Data 2). Applying the Pathway Studio software and Fisher's exact test we searched for enrichment of the Gene Ontology (GO) category "biological process" (Supplementary Data 4).  $P$ -values were corrected for multiple testing according to



**Fig. 4** *DLX1* expression. **a** No correlation between expression of *DLX1* and degree of methylation in human forebrains (pyrosequencing value at CpG [hg19] chr2:172,953,097) [Pearson’s correlation analysis including both PSP patients ( $n = 69$ , gray dots) and controls ( $n = 67$ , white dots)]. Expression of *DLX1* did not differ between patients and controls (Welch’s corrected unpaired  $t$ -test, n.s. = not significant, bar plot with mean and SEM). **b** Significant correlation between expression of *DLX1AS* and the degree of methylation. Expression of *DLX1AS* is significantly reduced in patients as compared to controls ( $***P < 0.001$ , Welch’s corrected unpaired  $t$ -test, bar plot with mean and SEM). **c, d** No difference between the amount of *DLX1* protein in white matter of frontal lobe of PSP patients and controls (co.). **c** Significant increase of *DLX1* protein in frontal lobe gray matter of PSP as compared to controls. **d** ( $n = 8$  per group,  $**P < 0.01$ , Welch’s corrected unpaired  $t$ -test, bar plot with mean and SEM).  $\beta$ -Actin was used as loading control. **e** No difference in immunoreactivity of *DLX1* in white matter of gyrus frontalis between PSP and controls. **f** Significant increase of *DLX1* protein in frontal lobe gray matter of PSP patients as compared to controls ( $n = 24$  PSP,  $n = 9$  controls,  $**P < 0.01$ , Mann-Whitney Test). Scale bar: 100  $\mu$ m. The line in the middle of the box and whisker graph represents the median (50th percentile). The box extends from the 25th to 75th percentile. The whiskers extend from the lowest to the highest value

Benjamini and Hochberg<sup>14</sup>. The top 20 highly significantly enriched categories include important functions pertinent to *DLX1* and *DLX2*, i.e., “anatomical structure development” (GO ID 48856), “regulation of signaling” (GO ID 23051), “cell fate commitment” (GO ID 45165), “positive regulation of transcription” (GO ID 45893) and the “cell surface receptor signaling pathway” (GO ID 7166). All 18 genes assigned to the term “Wnt signaling pathway” (GO ID 16055), which show a significantly

corrected  $P$ -value of  $6.57 \times 10^{-04}$  are part of the “cell surface receptor signaling pathway” (Supplementary Data 4). Particularly many neuronal functions and pathways are distributed across the list of significantly enriched categories, i.e., “neuron fate commitment” (GO ID 48663), “cerebral cortex GABAergic interneuron fate commitment” (GO ID 28193), and “negative regulation of neurogenesis” (GO ID 50768) (all of which include *DLX1* and *DLX2*, Supplementary Data 4).



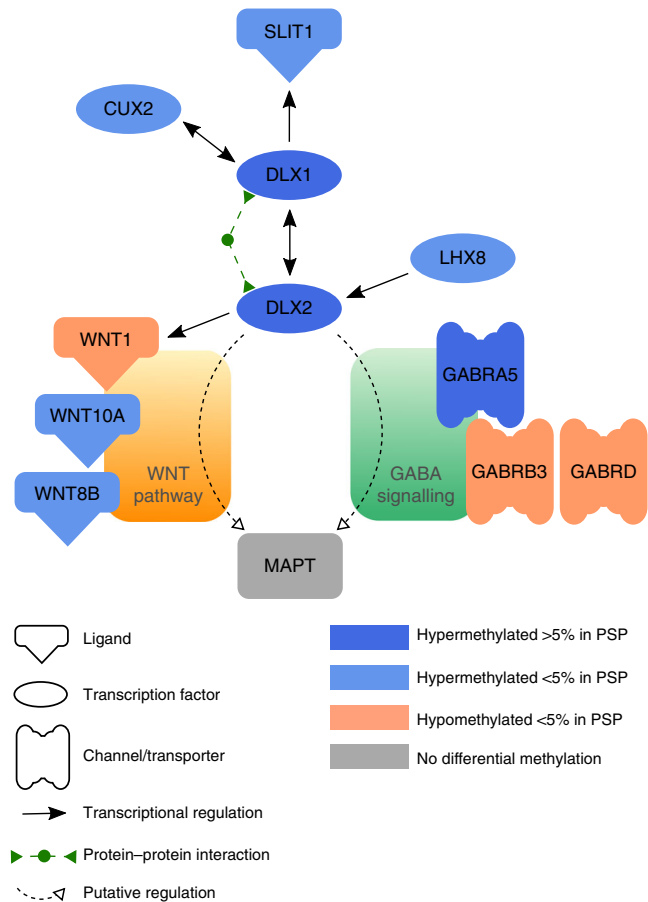
**Fig. 5** Overexpression and siRNA-mediated knock-down of *DLX1* and *DLX1AS*. **a** Overexpression of *DLX1* in Ntera2 cells results in upregulation of the *DLX1*-target genes *GAD1*, *BRN3B*, and *OLIG2* and in downregulation of *MAPT* (Student’s *t*-test). **b** Overexpression of *DLX1AS* in SH-EP cells using four different transcript variants (tv) (see Methods and Supplementary Fig. 3) results in downregulation of *DLX1* and its target genes and in upregulation of *MAPT* (Student’s *t*-test). **c** Specific siRNA-mediated knock-down of *DLX1* (siDLX1) in human fetal striatal neuronal precursor cells (strNPCs) results in upregulation of *MAPT* as compared to non-specific siRNAs (si\_ncRNA). Conversely, knock-down of *DLX1AS* (siDLX1AS) causes upregulation of *DLX1* and downregulation of *MAPT* (Student’s *t*-test). **d** ATP assay showing decreased viability of strNPCs overexpressing 4R-Tau protein after siRNA-mediated knock-down of *DLX1*. *DLX1* knock-down does not significantly reduce viability of un-transfected cells (UTC) or of cells overexpressing either 3R-Tau or the control protein mCherry. Knock-down of *DLX1AS* increases viability in cells overexpressing 3R- or 4R-Tau (two-tailed, unpaired Student’s *t*-test, *P*-values: \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001, n.s. = not significant)

**Pathway analysis of differentially methylated genes.** Pathway analysis revealed interdependence of several genes found to be differentially methylated in PSP (Fig. 6 and Supplementary Data 5). Based on literature mining we propose two main possible pathways linking *DLX1* and *MAPT*:

(1) Activation of *MAPT*-encoded Tau protein via the Wnt signaling pathway: This notion is supported by the finding that *DLX2* bound to Necdin activates the *WNT1* promoter<sup>29</sup>. In PSP patients, *DLX1* and *DLX2* are highly ( $\geq 5\%$ ) and the WNT ligand family members *WNT10A*, *WNT8b* are distinctly ( $> 3\%$ ) hypermethylated. Several additional differentially methylated genes (methylation differences  $> 1\%$ ) are members of the WNT signaling pathway as well (Supplementary Data 4). Furthermore, WNT signaling appears to affect Tau phosphorylation in Alzheimer’s disease<sup>30–32</sup>.

(2) Tau phosphorylation via GABA(A) receptors: *DLX1*, *DLX2*, and GABA(A) receptors (encoded by the differentially methylated genes *GABRA5*, *GABRB3*, and *GABRD*) are members of the GABAergic interneuron-related network in humans<sup>33</sup>. Within this network *DLX1/DLX2* regulate GABA synthesis<sup>24</sup>. Expression changes of *DLX1/DLX2* can alter activation of GABA(A). GABA(A) receptors in turn play an important role in Tau phosphorylation<sup>34</sup>. This observation is consistent with the notion that Tau may be affected via *DLX1/DLX2* – GABA(A) in PSP.

Interestingly, we also detected two specific *DLX1*-binding sites in the *MAPT* promoter. One is located 1972-bp upstream of the TSS. The previously not described *DLX1*-binding motif (CAT-AATTAAAAT) was detected using the DiAlign TF program (Genomatix)<sup>35</sup> by utilizing an optimized matrix similarity. It was found in the *MAPT* promoter of human and rhesus monkey. At a



**Fig. 6** Pathway analysis. The pathway proposed was deduced from in silico literature mining for functional interactions of the differentially methylated genes. The network was consolidated by verification of each interaction in the published literature. Hypermethylated genes are depicted in blue and hypomethylated genes are given in orange. Note that DLX1/DLX2 may influence *MAPT* either via the WNT (brown) or via the GABA signaling (green) pathway (for details see text)

slightly lowered threshold of the optimized matrix similarity of 2% a second DLX1-binding site (TCTAATTTAAGA) was identified 550 bp upstream of the human TSS of *MAPT*. It is evolutionarily more conserved than the site 1972-bp upstream of the TSS. Apart from primates (humans and rhesus monkey) it is also found in other mammalian species such as cow and horse (Supplementary Fig. 9).

## Discussion

This epigenome-wide association study in brains of PSP patients and controls interrogated > 485,000 CpG sites representing 99% of RefSeq genes. Significant differential DNA methylation between patients and controls was found at 717 CpG sites. Four-hundred fifty-one of these sites correspond to 375 annotated genes. Most methylation differences in PSP were subtle (< 2% at  $P < 0.05$ ), but greater differences ( $\geq 5\%$ ) were observed at 38 sites representing ten genes. The degree of differential methylation in PSP was smaller than in cancer, where large differences in methylation are typically found<sup>36</sup> but comparable to other complex disorders such as multiple sclerosis<sup>37,38</sup> or Alzheimer's disease<sup>39–41</sup>.

Genes involved in neuronal development and function were significantly overrepresented among the genes differentially methylated in PSP (Supplementary Data 4 and Supplementary

Data 5). These changes might have cumulative effects on disease origin and progression.

Two prior studies performed targeted, hypothesis-driven epigenetic analyses in DNA extracted from brain tissue and found differential methylation of *MAPT* on chromosome 17q21.31 in PSP patients<sup>7,8</sup>. Another prior study found differential methylation in the region of *MAPT* in peripheral blood DNA of PSP patients<sup>9</sup>. This methylation difference was associated with the H1 haplotype of *MAPT*, which is overrepresented in PSP<sup>1,5</sup>. In our epigenome-wide analysis, nine CpG sites within *MAPT* yielded significantly different methylation values in PSP vs. controls, which, however, did not hold up to correction for multiple testing (Supplementary Data 3). Thus, our findings are not at odds with these prior reports, but represent a more conservative interpretation.

Among the 375 genes identified in our study, differential methylation was most pronounced at *DLX1* that was hypermethylated at multiple CpG sites including a CpG island. Its methylation status differed by up to >10% between PSP and controls. The other genes showed altered methylation levels at one or a few CpG sites only, and methylation differences were mostly low (< 2%). The observed hypermethylation of *DLX1* in its 3' region did not alter *DLX1* transcription. However, the amount of the newly detected *DLX1AS* transcript that overlaps with parts of exon 3 and with the 3' UTR of *DLX1*, was significantly reduced to 0.64-fold in PSP as compared to controls. *DLX1AS* is a likely negative regulator of *DLX1* translation by dimer formation between the sense and antisense RNA<sup>42</sup>. Consistently, we observed increased expression of DLX1 protein in gray matter of forebrains in PSP on western blots and by immunohistochemistry. Interestingly, an enhancer has been assigned to the 5' region of *DLX1AS*<sup>18</sup>. This enhancer appears to be part of *DLX1AS*. As such it might be involved in modification of the amount of DLX1 protein in different tissues (e.g., gray vs. white matter) by controlling the amount of *DLX1/DLX1AS* RNA dimer formation. Alternatively, instead of forming a RNA/RNA dimer, *DLX1AS* might also form a complex with various proteins that in turn regulate expression of *DLX1*. This mode of action has been shown for *DLX6AS/EVF2*<sup>43,44</sup>.

*DLX1* encodes a homeobox-containing transcription factor. It is mainly expressed in GABAergic inhibitory interneurons. In concert with other *DLX* genes it impacts interneuron development and function<sup>45–47</sup>. This is consistent with our observation of mainly neuronal expression of *DLX1* in single cells in humans, confirming previously described neuronal specificity of *Dlx1* expression in mice<sup>22</sup>. Similarly, *Dlx1as* and *DLX1AS* are also predominantly expressed in neurons in both mouse and human. Interestingly, individual neurons of healthy human cerebral cortex expressed either *DLX1* or *DLX1AS*, but only very few neurons expressed both *DLX1* and *DLX1AS*, suggesting that expression of *DLX1* and *DLX1AS* might be regulated in an opposite manner (Supplementary Fig 4). This interpretation would be consistent with our cell culture experiments showing increased *DLX1AS* expression upon *DLX1* silencing and vice versa.

The finding of differential methylation of various genes within a putative *DLX1* pathway (Fig. 6) further strengthens the notion of an important role of *DLX1* in the etiology of PSP. While methylation differences were comparatively subtle at most of these genes, together they might be sufficient to alter the pathway in a functionally relevant manner in PSP. During normal ageing, the methylation pattern of cells may change dramatically, without having apparent functional effects in most cases<sup>48,49</sup>. In some instances, however, functionally related genes may become differentially methylated and predispose to disease via their interaction. Of particular interest is our finding that the putative *DLX1* pathway extends to *MAPT*. Via this pathway overexpressed *DLX1*



might modify protein Tau by increasing its phosphorylation, which in turn contributes to tangle formation<sup>1</sup>. In vitro experiments support dependency of *MAPT* expression on the concentration of *DLX1/DLX1AS* transcripts. Currently, it remains unclear whether *DLX1* influences *MAPT* expression directly or indirectly. Given that we could not conclusively show direct binding of *DLX1* to the *MAPT* promoter, the effect of *DLX1* on *MAPT* appears to be more indirect (via the Wnt or the GABAergic interneuron-related network pathway). Furthermore, a high-throughput ChIP-sequencing study with enrichment profiles for a large set of factors<sup>50</sup> also used an antibody against *DLX1* but did not find enrichment at or near the *MAPT* locus. Differential methylation of the various genes related to neuronal development and function might also be a secondary event rather than the primary cause of disease in PSP. It has been shown, that *DLX1* is essential for survival of adult interneurons in the neocortex<sup>46</sup> and in mature retinal cells<sup>51</sup>. Thus, differential regulation of *DLX1* and interacting genes involved in neuronal development, differentiation and plasticity, might refer to endogenous repair mechanisms in neocortical areas affected by PSP<sup>52</sup>. Consistent with these findings we found that siRNA-mediated knock-down of *DLX1* decreases neuronal viability in strNPCs overexpressing either 3R-Tau or 4R-Tau. Knock-down of *DLX1AS* rescued cells from death independent of overexpression of Tau.

In conclusion, we found significant methylation differences in DNA from forebrains of PSP patients as compared to controls. Differential methylation affected a high percentage of genes involved in neuronal differentiation and function. Methylation differences were particularly high at *DLX1/DLX1AS* that are expressed in neurons almost exclusively. *DLX1* protein was increased in gray matter of patients' forebrains and was shown to negatively affect *MAPT* expression in vitro. *DLX1* may thus contribute to disease by disturbing the normal function of neurons and might serve as a novel molecular target for the development of disease-modifying therapies.

## Methods

**Brain tissue.** The use of human brain tissue for this project was approved by the ethics committees of the University of Giessen and of the Technical University of Munich. Brain samples were from patients who had given informed consent before death.

For the epigenome-wide DNA methylation analysis, postmortem prefrontal lobe tissue of PSP patients ( $N = 94$ ,  $72 \pm 5.3$  years of age,  $N = 54$  (57%) male) and controls without neurological or psychiatric diseases ( $N = 71$ ,  $76 \pm 7.9$  years of age,  $N = 46$  (67%) male) was obtained from the CurePSP brain bank, Mayo Clinic, Jacksonville, Florida and the Victorian Brain Bank, Carlton, Australia. Detailed information on clinical and neuropathological findings is given in Supplementary Data 1. Of the PSP patients, 83 were homozygous and 11 were heterozygous for allele H1 of *MAPT*, none were homozygous for the H2 allele. Among the controls, 54 were homozygous for the H1 allele, 14 were heterozygous and 3 were homozygous for the H2 allele.

For western blot analysis, postmortem gyrus frontalis superior tissue of PSP patients ( $N = 8$ ,  $71.9 \pm 6.3$  years of age,  $N = 5$  (63%) male) and controls without neurological or psychiatric diseases ( $N = 8$ ,  $74.6 \pm 7.0$  years of age,  $N = 2$  (25%) male) was obtained from the Ludwig-Maximilians-Universität Munich and the Hospital Clínic de Barcelona brain banks.

For immunohistochemistry, postmortem gyrus frontalis superior tissue of PSP patients ( $N = 24$ ,  $72.6 \pm 4.8$  years of age,  $N = 18$  (75%) male) and controls without neurological or psychiatric diseases ( $N = 9$ ,  $70.6 \pm 9.7$  years of age,  $N = 4$  (44%) male) was obtained from the CurePSP brain bank, Mayo Clinic, Jacksonville, Florida.

**Nucleic acids and PCR.** DNA and RNA were extracted from tissue samples by standard procedures. DNA was prepared using the tissue extraction kit from Qiagen (Hilden, Germany) and RNA was extracted using the RNeasy Lipid Tissue-Kit (Qiagen). Genomic DNA was bisulfite-converted using the EpiTect Bisulfite Kit (Qiagen). For RT-qPCR, total RNA was converted into cDNA using the QuantiTect-Reverse-Transcription kit (Qiagen).

For investigation of the brain tissue samples, all primers used for standard PCR and RT-qPCR were designed using OLIGO Primer analysis software (Vers. 6.41; Molecular Biology Insights, Inc., Colorado Springs, CO, USA). Primers used for qPCR are given in Supplementary Table 1. Fluorescence data were compiled using

the CFX384 qPCR system from Bio-Rad (Hercules, CA, USA). Housekeeping genes *EIF4A2* and *CYCI* were co-analyzed for normalization of RNA content in each brain sample<sup>53</sup>.

Quantification of expression differences (fold-expression) was calculated using the REST-Software-Package-2009 as described elsewhere<sup>54</sup>.

Expression of the housekeeping genes *B2M*, *EIF4A2*, and *CYCI* was analyzed for normalization in cell culture experiments with SH-EP und Ntera2 cells. Stability values, i.e., coefficient of variation (CV-value) and gene stability value (*M*-value) for each housekeeping gene were  $< 0.25$  for CV- and  $< 0.5$  for the *M*-value. Fold expression was calculated with the integrated qBase module of the Bio-Rad CFX Manager 3.1 (3.1.1517.0823) software. Significance of expression changes was evaluated by unpaired two-tailed Student's *t*-test.

We have used DNase treated total-RNA in all RT-qPCR experiments. This excluded co-amplification of DNA that might have contaminated RNA preparations. Exclusion of DNA co-amplification was particularly important for analysis of exon1-derived RNA of *DLX1AS*. Exon1 RNA of *DLX1AS* was analyzed since this exon was present in all alternative transcripts.

RT-qPCR analyses were performed on RNAs extracted from strNPCs using SYBR Green Select qPCR Supermix (#4472954, Life Technologies, Carlsbad, USA), 5 ng cDNA synthesized from total RNA, 0.2  $\mu$ M forward and reverse primers (primer sequences are given in Supplementary Table 1). qPCR analysis was performed on a Step One Plus instrument (ThermoFisher Scientific, Carlsbad, USA). Initial incubation at 50 °C for 2 min and an additional 2 min at 95 °C was followed by 40 cycles of 15 s at 95 °C and 60 s at 60 °C. Threshold cycle (CT) values were set within the exponential phase of the PCR. Data were normalized to five housekeeping genes: *TBP*, *GPBP1*, *PPIA*, *PSMCI*, *UBQLN2* and the comparative normalized relative quantity (CNRQ) was used to calculate fold expression (qBase<sup>®</sup>, Biogazelle, Belgium). Gene regulation was statistically evaluated by two-tailed Student's *t*-test on the assumption of equal variances.

**Array analysis.** The Infinium 450k array of Illumina Inc. (San Diego, CA) was used for methylation analysis. This array analyzes more than 485,000 CpG sites distributed over the entire genome. Two-hundred nanograms of bisulfite-converted DNA were hybridized to the arrays according to the manufacturer's instructions. In order to exclude batch effects, we alternated hybridization of PSP and control samples. Arrays were scanned on an Illumina iScan platform at the Life & Brain Center (Bonn/Germany). Array data were evaluated, preprocessed and normalized using the ChAMP pipeline (V.2.8.1)<sup>55,56</sup> with minor modifications. Specifically, raw array data were uploaded to the ChAMP pipeline using the minfi option<sup>57</sup>. Relative proportions of neuronal and non-neuronal cells in each sample were estimated based on the raw data applying the compositeCellType = "DLPFC" option of estimateCellCounts function of the R Bioconductor minfi package<sup>58</sup>. Probes with less than three measured beads or a detection *P*-value  $> 0.01$  as well as probes interrogating CpGs that fall on or near to a SNP were removed based on recommendations by Zhou et al.<sup>59</sup>. Similarly, probes aligning to multiple locations as defined by Nordlund et al.<sup>60</sup> were removed from the analysis. All samples had more than 99.3% of valid probes, therefore, no samples were removed from the analysis based on the quality of the data. The sex of each sample was verified using an in-house R script based on the DNA methylation profile of the X and Y chromosomes. One single control sample was not in concordance with the assigned sex and was excluded from further analysis. Differences between probes due to Infi and InfiI probe usage as well as between samples were normalized using BMIQ<sup>61</sup> on the 435,803 remaining probes. Differential methylation was determined using a modified version of the champ.dmp function with age, sex, and estimated proportion of non-neuronal cells as covariates in a linear regression analysis based on the limma package<sup>62</sup> adjusting *P*-values for multiple testing with a Benjamini-Hochberg correction<sup>14</sup>. Differentially methylated CpGs were considered significant at adjusted *P*-values  $< 0.05$ . To account for potential confounding owing to genetic variation, significant probes were filtered for significant mQTLs in adult prefrontal cortex using previously published data, which removed three CpGs all of which associated with *GABRA5* (cg01378667, cg03325535, cg10318222) as defined by a SNP in close proximity to the gene (rs7496866)<sup>15</sup>.

We computed the exact statistical power for each of the 485,577 probe sets on the microarray based on the data obtained in our sample by applying the function pwr.t2n.test of the pwr package of R statistical software. The parameters of this function are the number of patients and controls, i.e.,  $n = 94$  and  $n = 71$ , respectively, as well as the significance level of 0.05 and Cohen's *d* (effect size). The effect size is defined by the difference between the means of the group divided by the pooled standard deviations of the two groups. Given these parameters, a power of at least 80% is calculated for 14,553 out of the total of 485,577 probes. Of the 717 CpG sites, which were significantly differentially methylated in PSP as compared to controls 664 (92.6%) showed a power of greater than 80% (Supplementary Data 6).

The circle plot was built with hg19 as reference genome using the OmicCircos R-Package Version 1.4.0<sup>63</sup> for R version 3.1.3<sup>64</sup>.

**Pyrosequencing.** Pyrosequencing was done on a PyroMark-Q24 using PyroMark-Q24 tools and reagents (Qiagen) according to the manufacturer's protocol. Primers for pyrosequencing (PCR product and sequencing primer) were designed using PyroMark-Assaydesign-2.0.1.15 (Qiagen) software. Primer sequences are shown in Supplementary Table 1. Data were analyzed with the PyroMark-Q24 software.

**Western blot.** White and gray matter from the Gyrus frontalis superior of post-mortem tissue was extracted in ice-cold N-PER™ lysis buffer (#87792, ThermoFisher Scientific, Waltham, MA, USA) using the Dounce homogenizer with 10–20 strokes on ice. Proteins were obtained after removal of cellular debris by centrifugation at 10,000 rpm for 10 min at 4 °C. Thereafter, 1 × HALT™ protease inhibitor (#78442, ThermoFisher Scientific, Waltham, MA, USA) was added to each supernatant. Protein concentration was determined by Bicinchoninic acid assay (#23225, ThermoFisher Scientific, Waltham, MA, USA).

Equal amounts of proteins were separated by sodium dodecyl sulfate polyacrylamide gel electrophoresis with 4–20% Criterion™, TGX Stain-free™ Protein gels (#5678094, Bio-Rad, Munich, Germany). Molecular weight markers used were # 830537 and # 830552 (Hessisch Oldendorf, Germany). Thereafter the protein staining was activated by UV light using the ChemiDoc-XRS system (Bio-Rad, Munich, Germany). After transfer to immuno-blot polyvinylidene difluoride (PVDF) membrane (#162-0177, Bio-Rad, Munich, Germany) proteins were fixed on the membrane with 3.75% paraformaldehyde (PFA) in 1 × PBS at room temperature (RT) for 15 min. After 3 × 10 min washing steps in 1 × PBS, membranes were incubated with primary DLX1 antibody (dilution 1:1000, ab126054, polyclonal rabbit AB, Abcam, CA, UK) for 48 h at 4 °C. (Human Protein Atlas entry: [<https://www.proteinatlas.org/ENSG00000114335-DLX1/tissue>]). Positive bands were detected using HRP-conjugated secondary antibody (dilution 1:5000, ab P0448, polyclonal goat anti rabbit immunoglobulins, affinity isolated, Agilent, CA, USA) and chemolumina substrate (#170-5061, Clarity™ Western ECL Blotting Substrate, Bio-Rad, Munich, Germany). Blots were stripped and hybridized with  $\beta$ -actin antibody overnight (dilution 1:2000, ab 3700, clone 8H10D10, mouse monoclonal  $\beta$ -actin antibody, Cell Signaling Technology, CA, UK). Positive bands were detected by an HRP-conjugated secondary antibody (dilution 1:5000, ab P0447 polyclonal goat anti mouse immunoglobulins, affinity isolated) and chemoluminescent substrate (see above). Bands were scanned using the ChemiDoc-XRS system and Image Lab software V.5.1 was applied to measurement of optical densities (Bio-Rad). Optical densities of anti-DLX1 antibody positive bands were compared to TGX labeled total protein and the housekeeping protein  $\beta$ -actin. DLX1 antibody was validated in positive and negative controls, e.g., tissue specific human neuronal progenitor cells derived from striatum or protein lysates from the human pancreas.

**Aperio methodology.** DLX1 immunopositive structures on tissue sections were quantitatively measured applying Aperio technology (Leica Microsystems Inc., Buffalo Grove, IL). Immunostained frontal cortical sections were scanned at 20 × magnification on the ScanScope® AT2 (Leica Biosystems, Wetzlar, Germany). The tissue sections were annotated using ImageScope (Leica Biosystems) to trace both gray matter and neighboring white matter along the strait of the gyrus, specifically avoiding the depth of the sulcus and rise of the gyrus. Annotated slide files were analyzed with a custom-designed color deconvolution algorithm using the eSlideManager (Leica Biosystems).

**Cell culture experiments.** We used Ntera2, SH-EP cells, and strNPCs in our studies. Ntera2 cells (pluripotent human embryonal carcinoma cells with characteristics of neuroepithelial precursor cells) were applied to overexpression experiments of DLX1 since intrinsic DLX1 mRNA expression is low. Ntera2 cells display neuronal characteristics (e.g., differentiation into neurons upon induction with retinoic acid), grow well under standard cell culture conditions and can be easily transfected.

SH-EP cells, a human neuroblastoma-derived cell line, were used in experiments involving *DLX1AS* overexpression. Endogenous expression of *DLX1* is increased as compared to other cell lines tested. We studied human striatal neuronal precursor cells (strNPCs) for siRNA experiments since they are commonly used as proxies of neuronal cells. They are not transformed and thus lack the malignant phenotype of other cell lines.

Human cell lines SH-EP were grown in RPMI Medium (GibCo) and Ntera-2 (NT2) in DMEM Medium (GibCo), containing 10% fetal calf serum (SIGMA), Glutamine [2 mM], Penicillin-G [100 U/ml] and Streptomycin [100  $\mu$ g/ml] at 37 °C and 5% CO<sub>2</sub>. (SH-EP and NT2 Cells were provided by the Neuroblastoma Study Group Cologne, Germany).

A full-length cDNA encoding *DLX1* (NM\_178120.4) and cDNAs of *DLX1AS* transcript variants (tv) 2 (KU179669), 4 (KU179671), 5 (KU197672) and partial transcript variant 1 (5' part of KU179668 [hg19]chr2:172,958,068–172,958,335 joined with chr2:172,954,669–172,954,850) were PCR-amplified from fetal brain mRNA, sequenced and cloned into a pcDNA-3.1-TOPO vector.

Cells were transfected with these constructs or empty vector using Viromer-Red transfection reagent (Lipocalyx, Germany) and selected in medium containing G418. Cells were grown in this medium for 14–21 days. They were harvested at a confluency of 70–80%.

siRNA-mediated knock-down experiments were performed in fetal striatum-derived human neuronal progenitor cells (strNPCs)<sup>65</sup> using siPOOL-DLX1 siRNA, siPOOL-DLX1AS and siPOOL non-coding (nc)siRNA (siTOOLS BIOTECH, Planegg, Germany). Experiments were repeated three times. Transfection of siRNA-Pools was done with Lipofectamine® RNAiMAX reagent (#13778150, ThermoFisher Scientific, Carlsbad, USA) according to the manufacturer's protocol

at a final concentration of 5 nM. Cells were harvested and lysates prepared after 6 days of culture.

ATP Assay was performed according to standard procedures<sup>27</sup>, strNPCs were transfected with siPOOL-DLX1, siPOOL-DLX1AS or siPOOL-ntsiRNA on day one of culture. After 48 h Lentiviruses overexpressing either mCherry, MAPT3R or MAPT4R<sup>66,67</sup> were added to the cells. Cells were harvested 6 days after onset of culture and ATP assays were performed using ViaLight™ plus kit (LT07-221, Lonza, Germany).

**Literature mining and pathway analysis.** Using Network Builder of the Pathway Studio software (Elsevier) version 12.0.1.5 we generated a literature-derived network that was based on text mining for direct interactions between significantly differentially methylated input genes. The input dataset comprised genes adjacent to the 717 genomic positions (CpG sites) found to significantly differ in methylation between PSP patients and controls. We only analyzed those genes by Network Builder that are known to be expressed in the telencephalon and in interneurons. The GeneRanker program (Genomatix) was used for analysis of association of differentially methylated genes in tissues (Supplementary Data 5)<sup>67</sup>. The literature-derived network was curated and extended by manual literature searches as well as by in silico prediction of transcription factor binding sites applying the MatInspector program (Genomatix).

The complete set of differentially methylated genes was further explored using the Pathway Studio software for enrichment in the category “biological process” of the Gene Ontology by applying Fisher's exact test. *P*-values were corrected for multiple testing by Benjamini and Hochberg<sup>14</sup>.

**Single-cell analysis.** We downloaded data on single cells from GEO and processed them according to the recommendations detailed in [[GSE67835](https://doi.org/10.1038/gsc67835)]<sup>22</sup>.

We applied the program Prinseq (-min\_len 30) to remove very short non-specific reads. Prinseq also trimmed both ends of the reads in order to eliminate 5' duplicates (-trim\_left 10) and to remove low quality 3' ends (-trim\_qual\_right 25). Furthermore Prinseq filtered reads of low complexity (-lc\_method entropy /-lc\_threshold 65). The program FASTQC was used to identify sequences that are overrepresented (adapter) in order to exclude them from further analysis. We used the Prinseq tool to remove orphan pairs less than 30 bp in length followed by removal of nextera adapters using Trim Galore (--stringency 1).

Reads were aligned to the hg19 genome with STAR using the following options (-outFilterType BySJout/--outFilterMultimapNmax 20/--alignSJoverhangMin 8/--alignSJBoverhangMin 1 /--outFilterMismatchNmax 999/--outFilterMismatchNoverLmax 0.04/--alignIntronMin 20 /--alignIntronMax 1000000 /--alignMatesGapMax 1000000 /--outSAMstrandField intronMotif).

Aligned reads were converted to counts for each gene using HTSeq (-m intersection-nonempty /-s no). The human Ensembl General Feature Format (GTF) annotation file (version 2013–09) necessary for HTSeq was extended by *DLX1AS* splice variants represented by the following Genbank identifier KU179668.1, KU179669.1, KU179670.1, KU179671.1, and KU179672.1. For all analyses Genome\_build hg19 was used. Finally, counts were converted to FPKM (Fragments per kilobase of transcript sequence per million mapped fragments) values.

Out of 466 available single-cell datasets, we selected 251 datasets that represent specific cortical cell types (astrocytes, microglia, oligodendrocytes, endothelia, neurons, oligodendrocyte precursor cells, see [<https://www.ncbi.nlm.nih.gov/Traces/study/?acc=SRP057196>]). Hybrid cells and fetal quiescent cells were not considered.

**In silico promoter analysis.** All sequences analyzed are from the promoter sequence retrieval database EIDorado 12–2016 (Genomatix) that is based on NCBI build GRCh38. The following Genomatix / Entrez Gene identifiers for *MAPT* were used: GXP\_10577 / 4137 (human), GXP\_7225047 / 574327 (rhesus\_monkey), GXP\_5397266 / 100054638 (horse), GXP\_3861378 / 281296 (cow). Promoter sequences of *MAPT* from four mammalian species were aligned using the DiAlign TF program of the Genomatix software suite GEMS Launcher in order to evaluate overall promoter similarity and to identify DLX1-binding sites. The corresponding position weight matrices V\$DLX1.01 and V\$DLX1.02 were applied to promoter analysis according to the Matrix Family Library Version 10.0 (October 2016). Binding sites were considered conserved if promoter sequences could be aligned in the region of the DLX1-binding site using the DiAlign TF program.

**Statistics.** Statistical tests were performed within the R computing environment<sup>64</sup> [<http://www.r-project.org>]. In order to calculate Pearson's product moment correlation the R function cor.test was used by setting alternative = “two-sided” and method = “pearson”. We applied the function t.test to two sample *t*-tests with Welch's correction and the arguments paired = “FALSE”, alternative = “two-sided” and var.equal = “FALSE”. Mann–Whitney tests were performed by using the R function wilcox.test in conjunction with paired = “FALSE” and alternative = “two-sided”.

Test of normality according to Kolmogorov–Smirnov was applied separately to CpG-control and CpG-PSP samples for one specific gene from pyrosequencing analysis. If one of both tests was significant the corresponding dataset was defined

as not normally distributed. If only one of all CpG samples for a specific gene are not normally distributed the unpaired *t*-test with Welch's correction was performed otherwise the Mann–Whitney test. In case of the densitometry of DLX1-immunoreactivity experiment the Mann–Whitney test was used because the corresponding datasets were not normal distributed according to Kolmogorov–Smirnov.

In general,  $P < 0.05$  was considered significant. For bar plots the mean  $\pm$  SEM (standard error of the mean) is given unless indicated otherwise.

Whenever multiple tests were computed e.g., for enrichment analyses of gene occurrences in pathways these *P*-values were corrected according to Benjamini–Hochberg or according to other appropriate methods, which are given in the corresponding subsections for array and pathway analyses.

**Data availability.** Normalized and raw BeadChipArray data have been deposited in NCBI–Gene Expression Omnibus (GEO) with the accession code [GSE75704](#).

DLX1AS transcript variants have been uploaded to NCBI under accession numbers KU179668, KU179669, KU179670, KU179671, and KU179672.

All other data are available within the paper and its associated supplementary material or upon reasonable request from the corresponding authors.

Received: 13 July 2017 Accepted: 25 June 2018

Published online: 26 July 2018

## References

- Boxer, A. L. et al. New diagnostics and therapeutics for progressive supranuclear palsy. *Lancet Neurol.* **16**, 552–563 (2017).
- Höglinger, G. U. et al. Clinical diagnosis of progressive supranuclear palsy—The Movement Disorder Society Criteria. *Mov. Disord.* **32**, 853–864 (2017).
- Kovacs, G. G. Neuropathology of tauopathies: principles and practice. *Neuropathol. Appl. Neurobiol.* **41**, 3–23 (2015).
- Baker, M. et al. Association of an extended haplotype in the tau gene with progressive supranuclear palsy. *Hum. Mol. Genet.* **8**, 711–715 (1999).
- Höglinger, G. U. et al. Identification of common variants influencing risk of the tauopathy progressive supranuclear palsy. *Nat. Genet.* **43**, 699–705 (2011).
- Coyle-Gilchrist, I. T. et al. Prevalence, characteristics, and survival of frontotemporal lobar degeneration syndromes. *Neurology* **86**, 1736–1743 (2016).
- Huin, V. et al. The MAPT gene is differentially methylated in the progressive supranuclear palsy brain. *Mov. Disord.* **31**, 1883–1890 (2016).
- Allen, M. et al. Gene expression, methylation and neuropathology correlations at progressive supranuclear palsy risk loci. *Acta Neuropathol.* **132**, 197–211 (2016).
- Li, Y. et al. An epigenetic signature in peripheral blood associated with the haplotype on 17q21.31, a risk factor for neurodegenerative tauopathy. *PLoS Genet.* **10**, e1004211 (2014).
- Smith, P. Y. et al. MicroRNA-132 loss is associated with tau exon 10 inclusion in progressive supranuclear palsy. *Hum. Mol. Genet.* **20**, 4016–4024 (2011).
- Tatura, R. et al. microRNA profiling: Increased expression of miR-147a and miR-518e in progressive supranuclear palsy (PSP). *Neurogenetics* **17**, 165–171 (2016).
- Guintivano, J., Aryee, M. J. & Kaminsky, Z. A. A cell epigenotype specific model for the correction of brain cellular heterogeneity bias and its application to age, brain region and major depression. *Epigenetics* **3**, 290–302 (2013).
- Sandoval, J. et al. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics* **6**, 692–702 (2011).
- Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **1995**, 57, 289–300 (1995).
- Hannon, E. et al. Methylation QTLs in the developing brain and their enrichment in schizophrenia risk loci. *Nat. Neurosci.* **19**, 48–54 (2016).
- Mc McGuiness, T. et al. Sequence organization and transcription of the *Dlx-1* and *Dlx-2* loci. *Genomics* **35**, 473–485 (1996).
- Dinger, M. E. et al. Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Res.* **18**, 1433–1445 (2008).
- Kraus, P. et al. Making sense of DLX1 antisense RNA. *Dev. Biol.* **376**, 224–235 (2013).
- Barry, G. et al. Long non-coding RNA expression during aging in the human subependymal zone. *Front. Neurol.* **6**, 1–10 (2015).
- Andersson, R. et al. An atlas of active enhancers across human cell types and tissues. *Nature* **507**, 455–461 (2014).
- Pfaffl, M. W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucl. Acids Res.* **29**, e45 (2001).
- Zhang, Y. et al. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J. Neurosci.* **34**, 11929–11947 (2014).
- Darmanis, S. et al. A survey of human brain transcriptome diversity at the single cell level. *Proc. Natl. Acad. Sci. USA* **112**, 7285–7290 (2015).
- Le, T. N. et al. GABAergic interneuron differentiation in the basal forebrain is mediated through direct regulation of glutamic acid decarboxylase isoforms by *Dlx* homeobox transcription factors. *J. Neurosci.* **37**, 8816–8829 (2017).
- Zhang, Q. et al. Regulation of *Brn3b* by *DLX1* and *DLX2* is required for retinal ganglion cell differentiation in the vertebrate retina. *Development* **144**, 1698–1711 (2017).
- Givens, M. L. et al. Developmental regulation of gonadotropin-releasing hormone gene expression by the *MSX* and *DLX* homeodomain protein families. *J. Biol. Chem.* **19**, 19156–19165 (2005).
- Petryniak, M. A., Potter, G. B., Rowitch, D. H. & Rubenstein, J. L. *Dlx1* and *Dlx2* control neuronal versus oligodendroglial cell fate acquisition in the developing forebrain. *Neuron* **55**, 417–433 (2007).
- Garewal, H. S., Ahmann, F. R., Schifman, R. B. & Celniker, A. ATP assay: ability to distinguish cytostatic from cytotoxic anticancer drug effects. *J. Natl. Cancer Inst.* **77**, 1039–1045 (1986).
- Kuwajima, T., Nishimura, I. & Yoshikawa, K. *Necdin* promotes GABAergic neuron differentiation in cooperation with *Dlx* homeodomain proteins. *J. Neurosci.* **26**, 5383–5392 (2006).
- Mudher, A. & Lovestone, S. Alzheimer's disease—do tauists and baptists finally shake hands? *Trends Neurosci.* **25**, 22–26 (2002).
- Dobrowski et al. Presenilin deficiency or lysosomal inhibition enhances Wnt signaling through relocalization of GSK3 to the late-endosomal compartment. *Cell Rep.* **2**, 1316–1328 (2012).
- Zhang, L., Bahety, P. & Ee, P. L. Wnt co-receptor *LRP5/6* overexpression confers protection against hydrogen peroxide-induced neurotoxicity and reduces tau phosphorylation in SH-SY5Y cells. *Neurochem. Int.* **87**, 13–21 (2015).
- Al-Jaberi, N., Lindsay, S., Sarma, S., Bayatti, N. & Clowry, G. J. The early fetal development of human neocortical GABAergic interneurons. *Cereb. Cortex* **25**, 631–645 (2015).
- Nykänen, N. P., Kysenius, K., Sakha, P., Tammela, P. & Huttunen, H.  $\gamma$ -Aminobutyric acid type A (GABA<sub>A</sub>) receptor activation modulates tau phosphorylation. *J. Biol. Chem.* **287**, 6743–6752 (2012).
- Cartharius, K. et al. MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics* **21**, 2933–2942 (2005).
- Kulis, M. et al. Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. *Nat. Genet.* **44**, 1236–1242 (2012).
- Huynh, J. L. & Casaccia, P. Epigenetic mechanisms in multiple sclerosis: implications for pathogenesis and treatment. *Lancet Neurol.* **12**, 195–206 (2013).
- Huynh, J. L. et al. Epigenome-wide differences in pathology-free regions of multiple sclerosis-affected brains. *Nat. Neurosci.* **17**, 121–130 (2014).
- Lord, J. & Cruchaga, C. The epigenetic landscape of Alzheimer's disease. *Nat. Neurosci.* **17**, 1138–1140 (2014).
- Lunnon, K. et al. Methyloic profiling implicates cortical deregulation of *ANK1* in Alzheimer's disease. *Nat. Neurosci.* **17**, 1164–1170 (2014).
- De Jager, P. L. et al. Alzheimer's disease: early alterations in brain DNA methylation at *ANK1*, *BIN1*, *RHBDF2* and other loci. *Nat. Neurosci.* **17**, 1156–1163 (2014).
- Pelechano, V. & Steinmetz, L. M. Gene regulation by antisense transcription. *Nat. Rev. Genet.* **14**, 880–893 (2013).
- Berghoff, E. G. et al. *Evf2* (*Dlx6as*) lncRNA regulates ultraconserved enhancer methylation and the differential transcriptional control of adjacent genes. *Development* **140**, 4407–4416 (2013).
- Cajigas, I. et al. *Evf2* lncRNA/*BRG1*/*DLX1* interactions reveal RNA-dependent inhibition of chromatin remodeling. *Development* **142**, 2641–2652 (2015).
- Panganiban, G., & Rubenstein, J. L. Developmental functions of the distal-less/*Dlx* homeobox genes. *Development* **129**, 4371–4386 (2002).
- Cobos, I. et al. Mice lacking *Dlx1* show subtype-specific loss of interneurons, reduced inhibition and epilepsy. *Nat. Neurosci.* **8**, 1050–1068 (2005).
- Stühmer, T. et al. Ectopic expression of *Dlx* genes induce glutamic acid decarboxylase and *Dlx* expression. *Development* **129**, 245–252 (2002).
- Jung, M. & Pfeifer, G. P. Aging and DNA methylation. *BMC Biol.* **13**, 7 (2015).
- Jones, M. J., Goodman, S. J. & Kobor, M. S. DNA methylation and healthy human aging. *Aging Cell* **14**, 924–932 (2015).
- Yan, J. et al. Transcription factor binding in human cells occurs in dense clusters formed around cohesion anchor sites. *Cell* **154**, 801–813 (2013).
- de Melo, J. et al. *Dlx1* and *Dlx2* function is necessary for terminal differentiation and survival of late-born retinal ganglion cells in the developing mouse retina. *Development* **132**, 311–322 (2005).
- Halliday, G. M., Macdonald, V. & Henderson, J. M. A comparison of degeneration in motor thalamus and cortex between progressive supranuclear palsy and Parkinson's disease. *Brain* **128**, 2272–2280 (2005).

53. Penna, I. et al. Selection of candidate housekeeping genes for normalization in human postmortem brain samples. *Int. J. Mol. Sci.* **12**, 5461–5470 (2011).
54. Pfaffl, M. W., Horgan, G. W. & Dempfle, L. Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucl. Acids Res.* **30**, e36 (2002).
55. Morris, T. J. et al. Champ: 450 k chip analysis methylation pipeline. *Bioinformatics* **30**, 428–430 (2017).
56. Tian, Y. et al. ChAMP: updated methylation analysis pipeline for Illumina BeadChips. *Bioinformatics* **33**, 3982–3984 (2017).
57. Aryee, M. J. Minfi: A flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363–1369 (2014).
58. Fortin, J. P., Triche, T. J. Jr. & Hansen, K. D. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. *Bioinformatics* **33**, 558–560 (2017).
59. Zhou, W., Laird, P. W. & Shen, H. Comprehensive characterization, annotation and innovative use of Infinium DNA Methylation BeadChip probes. *Nucl. Acids Res.* **45**, e22 (2017).
60. Nordlund, J. et al. Genome-wide signatures of differential DNA methylation in pediatric acute lymphoblastic leukemia. *Genome Biol.* **14**, r105 (2013).
61. Teschendorf, A. E. et al. A beta-mixture quantile normalization method for correcting probe design bias in illumina infinium 450 k DNA methylation data. *Bioinformatics* **29**, 189–196 (2013).
62. Smyth, G. K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3**, Article3 (2004).
63. Hu, Y. et al. OmicCircos: A simple-to-use R package for the circular visualization of multidimensional omics data. *Cancer Inform.* **13**, 13–20 (2014).
64. R Development Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, Vienna, Austria, 2011).
65. Cutts, J., Brookhouser, N. & Brafman, D. A. in *Stem Cell Heterogeneity. Methods in Molecular Biology*, Vol. 1516 (ed Turksen, K.) (Humana Press, New York) 121–144 (2016).
66. Bruch, J. et al. Perk activation mitigates Tau pathology in vitro and in vivo. *EMBO Mol. Med.* **9**, 371–384 (2017).
67. Berriz, F. P., King, O. D., Bryand, B., Sander, C. & Roth, F. P. Characterizing gene sets with FuncAssociate. *Bioinformatics* **19**, 2502–2504 (2003).

## Acknowledgements

We gratefully acknowledge technical support of Claudia Röhrsheim, Michaela Weis, and Stefanie Teschler (RT-qPCR, pyrosequencing), Monica Castanedes-Casey (immunohistochemistry), and Drs. Melissa Murray and Amanda Serie (Aperio technology). We thank Drs. Eilis Hannon and Jonathan Mill (University of Exeter Medical School) for the mQTL data of the adult prefrontal cortex. We also thank Drs. Per Hoffmann and Stefan Herms for advice on data analysis. Dr. Hong Xu is acknowledged for his help with DLX1 siRNA and tau overexpression experiments. We are indebted to Dr. Ellen Gelpi and Dr. Laura Molina at the Neurological Tissue Bank of the Biobank-Hospital Clinic-IDIBAPS, Barcelona, for data and sample procurement. We thank Dr. Hagen Scherbr (Institute of Computational Biology, Helmholtz Center Munich, Germany) for the support regarding statistical questions. This work was supported by a grant (#518-14)

from CurePSP (to G.U.H., and U.M.), a German Federal Ministry of Education and Research (BMBF)—and French National Research Agency (ANR)—funded project on “Epigenomics of Parkinson’s disease” (O1KU1403A/ ANR-13-EPIG-0003-05 EpiPD to G.U.H., J.T. and U.M.), the BMBF-funded project “HitTau” (01EK1605A to G.U.H.), the Deutsche Forschungsgemeinschaft (DFG, HO2402/6-2 & Munich Cluster for Systems Neurology SyNergy) (to G.U.H.), the NOMIS foundation (FTLD project to G.U.H.), the German Science Foundation Collaborative Research Center (CRC) 870 and the Helmholtz Portfolio Theme “Supercomputing and Modeling for the Human Brain” (SMHB) (to W.W., D.T. and D.V.W.), and the Bayerisches Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst within Bavarian Research Network “Human Induced Pluripotent Stem Cells” (ForIPS) (to G.U.H., S.C.S., W.W. and D.V.W.).

## Author contributions

A.W. planned the methylation study, performed sequencing and RNA expression experiments, analyzed the data and prepared the figures. S.C.S. performed the western blot and siRNA experiments. J.T. and F.B. performed the methylation data analysis. P.W. studied DLX1AS RNA transcripts. P.T. performed the immunohistochemistry experiments and J.S. established NPCs. D.T., D.V.-W., and W.W. contributed the gene ontology and pathway analysis. D.T. performed power analysis. A.C.W. prepared the circle plot. T.A., E.G., C.M., and J.v.S. contributed brain tissue. D.W.D. contributed brain tissue and supervised the immunohistological experiments. T. Ad. performed chromatin immunoprecipitation. M.F. analyzed mQTL data. G.U.H. and U.M. conceived of and supervised the studies, coordinated assembly of the data and wrote the paper with input from all authors.

## Additional information

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-018-05325-y>.

**Competing interests:** The authors declare no competing interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Publisher’s note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018