

A genome-wide analysis of brain DNA methylation identifies new candidate genes for sporadic amyotrophic lateral sclerosis

JULIA M. MORAHAN^{1,3}, BING YU^{2,3}, RONALD J. TRENT^{2,3} & ROGER PAMPHLETT¹

¹Department of Pathology and ²Central Clinical School, The Stacey Motor Neuron Disease Laboratory, The University of Sydney, and ³Department of Molecular and Clinical Genetics, Royal Alfred Hospital, Australia

Abstract

Genetic variants may underlie sporadic amyotrophic lateral sclerosis (SALS), but in only a few percent of patients have causative mutations been found. This is possibly because SALS is more often due to a variation in DNA methylation, an epigenetic phenomenon involved in gene silencing. Methylation across the whole genome was examined in brain DNA of 10 SALS patients and 10 neurologically-normal controls. Methylated DNA was immunoprecipitated and interrogated by Affymetrix GeneChip Human Tiling 2.0R Arrays. Methylation levels were compared between SALS patients and controls at each region of methylation across the genome. SALS patients had either hypo- or hyper-methylation at 38 methylation sites ($p \leq 0.01$). Of these, 23 were associated with genes and three with CpG islands. Pathway analysis showed that genes with different methylation in SALS were particularly involved in calcium homeostasis, neurotransmission and oxidative stress. In conclusion, a number of genes, either unsuspected in SALS or in potential cell death pathways, showed altered methylation in SALS brains. The possibility of epigenetic therapy for SALS should encourage confirmation of these initial results in a future larger whole-genome DNA methylation study.

Key words: Amyotrophic lateral sclerosis, whole genome analysis, DNA methylation, brain, epigenetics

Introduction

The cause of sporadic amyotrophic lateral sclerosis (SALS) in the great majority patients remains unknown. The sporadic form is clinically indistinguishable from the genetic form, which raises the possibility that both disorders are caused by genetic variations (1). However, searches for disease-causing mutations and susceptibility loci in SALS have been largely unsuccessful (for review, see Siddique 2008). A major stumbling block in looking for a genetic cause of SALS has been the inability to explain why a presumed germline genetic variant is not commonly found in patients or their family members.

One mechanism that can mimic genetic change, but where Mendelian transmission is lacking, is DNA methylation (2). This epigenetic phenomenon covalently adds methyl groups to the cytosines of CpG dinucleotides. CpGs are often clustered in regions of high density known as CpG islands, which can be found in promoters. Methylation in promoters is recognized as a fundamental regulator of gene

silencing, and the CpG islands of genes that are actively expressed are usually free of methylation (3). Once established, methylation status is usually mitotically heritable, but the methylation slate is wiped clean during embryonic development (4). This raises the possibility that abnormal methylation of a gene can arise during development, or in adulthood due to environmental influences (5) and not be passed on to the next generation. A similar model of genetic-epigenetic interaction is increasingly being considered in cancer (6).

In SALS, pathogenic variations in methylation (epimutations) could lead to reduced expression of genes necessary for motor neuron survival, or to inappropriate expression of genes that are normally kept silent. Methylation in SALS has previously been examined only in the candidate genes SOD1 and VEGF (7) and in members of the metallothionein gene family (8). In none of these genes was the methylation pattern found to be different in SALS.

Recent technological advances allow for methylation to be studied across the entire genome (9), removing the requirement for *a priori* knowledge of

candidate genes. ChIP-on-Chip technology identifies methylation by combining chromatin immunoprecipitation (ChIP) with whole-genome scanning (Chip) (9–11). An antibody against methylcytosine isolates methylated fragments from genomic DNA and the methylated fragments are interrogated using gene chips (9). In this study we used this technique to analyse the whole-genome distribution of methylation in patients with SALS and compared these to controls. We examined DNA from brain tissue because epigenetic alterations vary widely between tissue types. Epimutations can affect transcriptional control elements that are distant from promoters (2), so that in addition to promoter methylation we measured methylation within genes and at intergenic regions. We found a number of changes in the pattern of methylation in ALS brains.

Materials and methods

Study subjects

Frozen brain samples from 10 male SALS patients (mean age 63 years, $SD \pm 12$ years) and 10 male neurologically-normal controls (mean age 62 years, $SD \pm 12$ years) were obtained from the New South Wales and Victorian Tissue Resource Centres (for clinical details see Supplementary Table I). SALS patients had no family history of ALS. The clinical diagnosis of probable or definite ALS was made on revised El Escorial criteria and confirmed by neuropathological examination of the brain and spinal cord (12). Mean post mortem delay was 14 h $SD \pm 8$ h in SALS and 14 h $SD \pm 6$ h in control subjects. All subjects were of European descent. The study was approved by the Human Research Ethics Committee of the Sydney South-West Area Health Service.

ChIP-on-Chip analysis

Genomic DNA was extracted using the phenol chloroform method from 1 g of lateral frontal cortex in the middle frontal region (Brodmann area 46). The same anatomic region was sampled in SALS patients and controls. Chromatin immunoprecipitation followed by microarray hybridization was performed by Genpathway (San Diego, CA, USA). Methylated fragments were interrogated with GeneChip Human Tiling 2.0R Arrays (Affymetrix, Santa Clara, CA, USA). These arrays are based on repeat-masked genomes and so do not include features such as *Alu* repeats or transposons.

Data analysis

Two-sample comparisons were made with Tiling Analysis Software version 1.0.08 (Affymetrix). ‘Pre-

sent’ or ‘absent’ calls were generated by applying a >2.0 threshold to intensity values. Intervals were derived by combining methylated sites that were less than 300 bp apart and that ran for a total length of at least 180 bp. Raw data files were analysed using Genpathway software which coordinated the genomic annotations of intervals, intensity peak metrics and sample comparisons. Gene annotations were taken from NCBI Build 36. CpG islands were defined according to March 2006 UCSC Genome Browser criteria, i.e. GC content of segment $>50\%$, length >200 bp and ratio of CpG to C and G >0.6 .

A group analysis compared the proportion of methylated sites with either present or absent methylation, using Fisher’s exact tests. To detect individual differences that may have been concealed in the group analysis, we compared peak intensity values for each SALS patient with the average value of controls. SALS:control probe intensities >1.3 or <0.7 were considered significant based on correlated GenPathway qPCR values of representative genes. The DAVID database (13) was used to analyse gene ontology functional categories and MatInspector (14) identified transcriptional binding site motifs. Statistical analyses were performed with JMP 6 (Cary, NC, USA) and Microsoft Excel (Redmond, WA, USA). Corrections for multiple testing were not performed due to controversies about these in whole-genome studies (see Discussion) and the modest sample number.

Replicate samples run at different times showed that 72% of intervals overlapped in the two runs. Much of this variation is probably due to small differences in probe intensities that result in values falling just below the threshold. Adjusting the threshold from 2.0 to a lower value of 1.7 increased the reproducibility to 92%, indicating that thresholding is responsible for most of the test-test difference.

Results

Nomenclature

We defined an interval as a length of DNA containing methylated sites that exceed a set threshold (Figure 1). A methylation region on the DNA strand contains a single interval in at least one subject, or intervals that overlap between different subjects (Figure 2). A methylation region can be methylated or unmethylated in different subjects. A gene domain is an area within 10 kb upstream or downstream of an annotated gene or pseudogene and consists of the gene body (exons and introns) and the flanking areas (Figure 3). The intergenic area is outside the gene domains. Data have been deposited in GEO (accession number GSE12525).

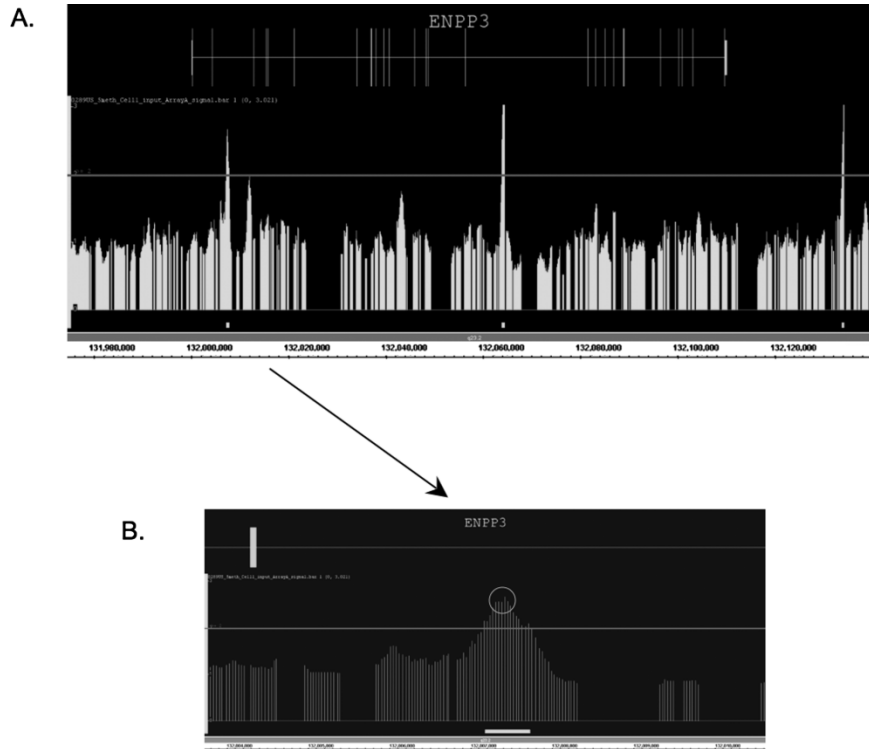


Figure 1. A. Affymetrix Integrated Genome Browser output showing exons (vertical lines) and introns (horizontal lines) of the gene *ENPP3* (in the upper part of the diagram). The *ENPP3* gene domain contains three intervals, two intronic and one in the right flanking region. The lower part of the diagram shows that intervals are made up of lengths of methylated probe intensities (vertical lines) above a threshold value (horizontal line). y-axis, probe intensity: x-axis, genomic location. Intervals are also indicated as small horizontal bars at the base of the diagram. B. Higher magnification of the left intronic interval. The peak intensity value (circle) is the highest probe value of the interval.

Distribution of intervals in different genomic regions

A total of 515,616 intervals were detected across the 20 subjects (Table I). These intervals gave rise to 64,424 methylation regions. Five percent of these were in CpG islands and 57% in gene domains (with or without CpG islands). The total number of

intervals per subject ranged from 14,105 to 37,765, with a similar variation in interval numbers in CpG islands and gene domains. Despite this wide range, the average number of intervals was similar in SALS ($25,384$; $SD \pm 7268$) and control ($26,178$; $SD \pm 6799$) brains. About half the intervals were within

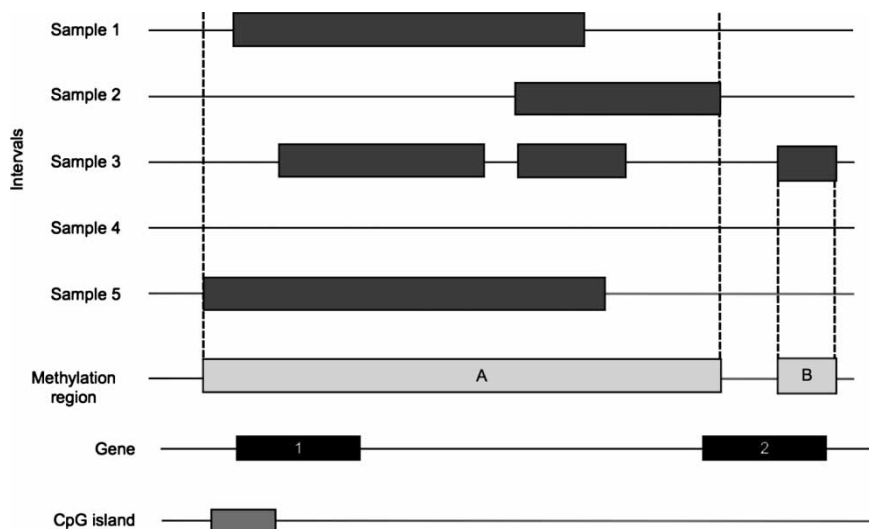


Figure 2. Data from five samples showing that intervals (upper darker bars) in the same genomic area can be combined into methylation regions (lighter bars). Samples with intervals in the methylation region are given a 'present' call. Samples may have more than one interval within a single methylation region (sample 3) or have no intervals (absent) within the methylation region (sample 4). Intervals that do not overlap within methylation region A (interval on far right) fall within another methylation region (B). Methylation regions can be associated with more than one genomic feature (e.g. methylation region A is associated with gene 1, gene 2 and a CpG island). Conversely, genes can be associated with more than one methylation region (e.g. gene 2 with methylation regions A and B).

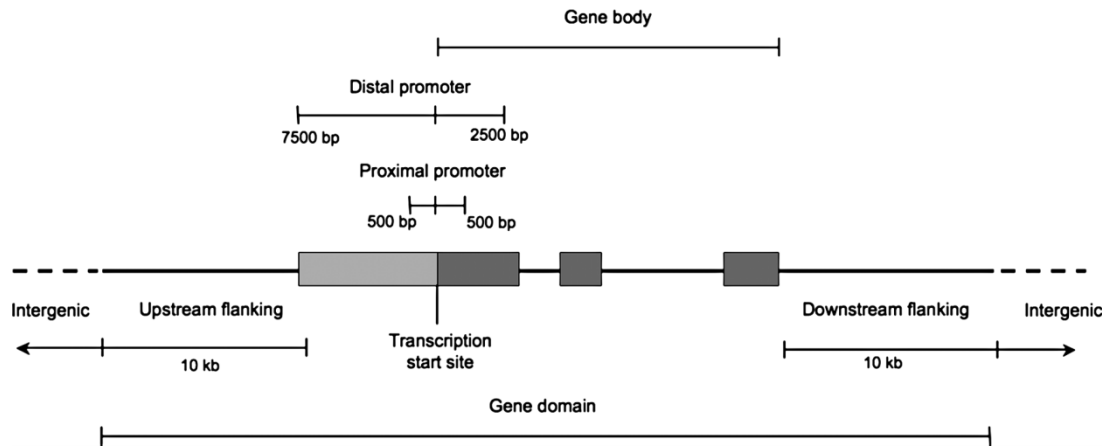


Figure 3. Terminology for genomic features. This genomic region contains a gene body with exons (dark bars) and introns, and a promoter (light bar).

gene domains and half in intergenic regions (Figure 4). Intervals in gene domains were mostly within the gene body, with smaller numbers flanking the gene or in the promoter (Figure 4). Fewer than 4% of intervals were in CpG islands. The proportion of intervals in each of the above genomic areas was not different in SALS and control brains (Table I).

Distribution of intervals across chromosomes

The total number of intervals in each chromosome was similar in SALS and control brains (data not shown). The proportion of intervals within gene domains ranged from 20% on chromosome Y to 73% on chromosome 19, and was not different between groups (Supplementary Figure 1). Promoter methylation ranged from 4% in chromosome 18 to 25% in chromosome 19, and was increased in controls at chromosomes 5 ($p=0.04$) and Y ($p=0.01$). CpG island methylation ranged from 1% in chromosome 3 to 18% in chromosome 19, and was increased in controls at chromosomes 3 ($p=0.02$), 7 ($p=0.04$), X ($p=0.02$) and Y ($p=0.01$).

Distribution of methylated genes and CpG islands

The average number of methylated gene domains in the whole group was 6957 (range 4596–8979), which corresponded to methylation at 17% of gene and pseudogene domains (Figure 5). The total number of methylated CpG islands was 55,859 (7% of CpG

islands in total). The number of gene domains that were methylated was similar in SALS and control brains (Supplementary Table II). In both SALS and control brains, 17% of gene domains and 7% of promoters were methylated. The proportion of promoters with methylated CpG islands was also similar in SALS (7%) and control (8%) brains.

Influence of age and post mortem delay on interval number

No correlation was found between total interval number and age ($r^2=0.09$) or post mortem delay ($r^2=0.002$). The number of methylated gene domains also did not correlate with age ($r^2=0.13$) or post mortem delay ($r^2=0.0001$). In addition, no correlations were found when SALS and control groups were analysed separately. Disease duration in SALS patients did not correlate with interval number ($r^2=0.03$) or the number of methylated gene domains ($r^2=0.03$).

Group comparisons

Present or absent methylation calls at methylation regions were compared between patients and controls using Fisher's exact tests. This identified 38 methylation regions at $p < 0.01$, with three of these in CpG islands and 23 in gene domains (Table II). A further 424 differently-methylated regions were identified at $p < 0.05$, 28 in CpG islands and 257

Table I. Distribution of intervals in different genomic regions in control and SALS brains.

Interval position	All subjects no.	Control no.	Control % of group total	SALS no.	SALS % of group total
Whole genome	515,616	261,780	–	253,836	–
Gene domain	248,516	126,429	48.3	122,087	48.1
Outside of gene domain	267,100	135,351	51.7	131,749	51.9
Promoter	39,905	20,315	7.8	19,590	7.7
Transcription start site	4349	2267	0.9	2082	0.8
CpG islands	18,004	9377	3.6	8627	3.4

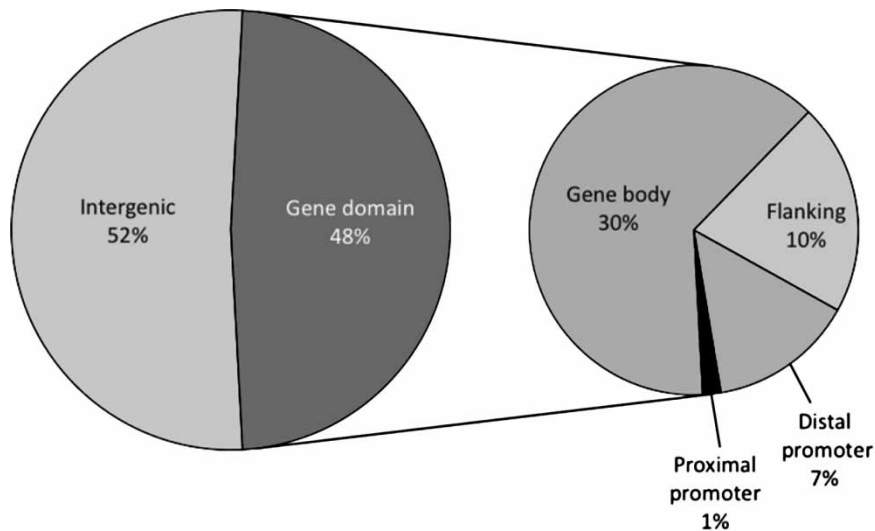


Figure 4. Distribution of methylated sites in brain DNA. About half of all methylation was in gene domains. Within gene domains, about one third of methylation was within the gene body; 8% of methylated regions were in the promoter. A 'distal promoter' is between 7500 bp upstream and 2500 bp downstream, and a 'proximal promoter' within 500 bp, of the transcriptional start site.

in 287 gene domains (data not shown). Genes with significant variation in methylation levels were reviewed as possible candidate genes for SALS, based on NCBI descriptions and the current literature on the suspected pathogenesis of ALS. The four genes with the most significant methylation differences ($p < 0.005$) were *ATRN*, *MSRA*, *PRDM16* and *SGCZ*. Of these, only *MSRA* was considered as a potential candidate due to its role in the oxidative stress response (15). The full list of suggested candidate genes is shown in Table III. In 24 of these genes, the methylation region contained a predicted transcription factor binding site.

Individual comparisons

When the peak probe intensity values of methylation regions in individual SALS patients were compared to the average intensity values of the controls, 47 methylation regions were found to be either methylated or unmethylated across at least five SALS patients (Table IV). These included 16 methylation regions in gene domains and 11 in CpG islands. Methylation regions located both in CpG islands and gene domains were found in *CPNE7*, *RPH3AL*, *LOC100131606*, *LOC100127980* and *LOC728117*. The *RPH3AL* methylation region, located upstream of the gene, was unmethylated in all SALS patients. In addition, a CpG island located within *RPH3AL* was unmethylated in eight of the 10 SALS patients. Another methylation region, located on chromosome 8, was unmethylated in all SALS brains but this contained no genes or CpG islands. Of the above genes, *RPH3AL* and *NSF* were considered possible candidate genes for ALS and both contained predicted transcriptional binding sites in their methylation regions.

Analysis of biochemical pathways

The DAVID database contained 83% of the genes that were identified in the $p < 0.05$ group analysis. Of these, 194 (81%) had a Gene Ontology annotation. The three biological processes with the most significant enrichment were calcium ion transport ($p = 0.00002$), the extracellular matrix ($p = 0.00005$) and cellular development ($p = 0.00006$). The 10 most statistically significant functional pathways are listed in Supplementary Table III.

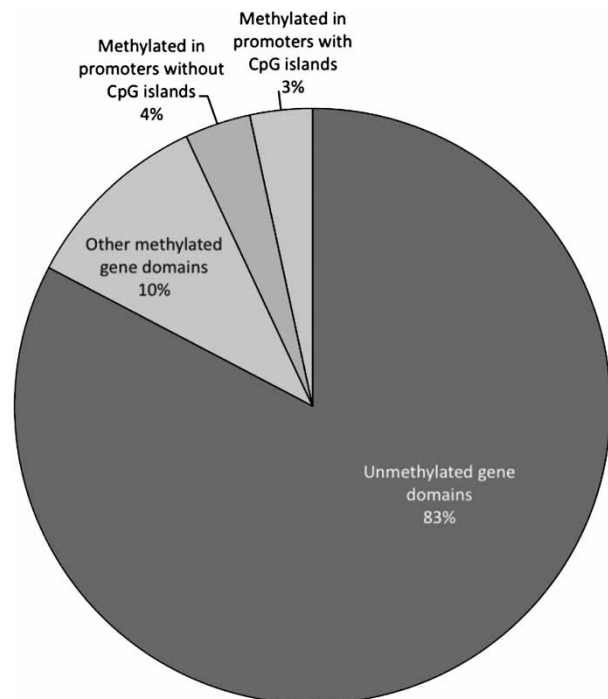


Figure 5. Methylation status of gene domains in the brain. Seventeen percent of genes have methylation regions within 10 kb of the gene body. Most gene domains are unmethylated. 'Other methylated gene domains' comprise the gene body and flanking areas, excluding the promoter.

Table II. Differently-methylated regions in SALS at $p \leq 0.01$. Ranked by increasing p -values.

Chr	Start position	Length (bp)	SALS present	SALS absent	Control present	Control absent	Gene domain	CpG island	p
20	3,557,635	494	0	10	8	2	<i>ATRN</i>		0.000
8	10,294,646	960	3	7	10	0	<i>MSRA*</i>		0.002
1	3,175,886	1187	1	9	8	2	<i>PRDM16</i>		0.003
3	148,133,280	698	9	1	2	8			0.003
8	14,603,469	715	2	8	9	1	<i>SGCZ</i>		0.003
11	84,554,430	732	2	8	9	1			0.003
2	22,447,428	629	0	10	6	4			0.005
2	36,232,341	1566	0	10	6	4			0.005
3	28,863,611	644	6	4	0	10			0.005
3	131,825,508	866	10	0	4	6	<i>COL6A6*</i>		0.005
6	163,187,405	661	0	10	6	4	<i>PACRG*</i>		0.005
7	127,949,472	258	0	10	6	4			0.005
12	30787946	1019	4	6	10	0	<i>CAPRN2</i>		0.005
13	112,850,942	821	0	10	6	4	<i>F10, PROZ</i>	Yes	0.005
15	69,686,865	451	6	4	0	10	<i>THSD4</i>		0.005
19	876,694	584	0	10	6	4	<i>KISS1R, ARI-D3A</i>	Yes	0.005
21	39,020,840	711	6	4	0	10			0.005
21	45,684,675	892	0	10	6	4	<i>COL18A1*</i>	Yes	0.005
1	37,427,784	736	3	7	9	1			0.010
1	58,485,198	567	9	1	3	7	<i>DAB1*</i>		0.010
4	15,008,264	808	1	9	7	3	<i>LOC100129903, C1QTNF7</i>		0.010
5	10,517,151	516	9	1	3	7	<i>ROPN1L</i>		0.010
5	169,083,100	390	7	3	1	9	<i>DOCK2</i>		0.010
6	31,362,546	1238	1	9	7	3			0.010
7	1,554,825	408	9	1	3	7	<i>MAFK, TME-M184A</i>		0.010
7	23,430,668	467	1	9	7	3	<i>IGF2BP3*</i>		0.010
7	26,617,652	426	9	1	3	7			0.010
9	109,432,020	661	7	3	1	9			0.010
10	30,409,099	687	3	7	9	1	<i>LOC729663</i>		0.010
11	98,750,711	234	3	7	9	1	<i>CNTN5*</i>		0.010
11	116,035,359	793	9	1	3	7			0.010
12	39,992,907	436	3	7	9	1			0.010
12	89,619,237	532	1	9	7	3			0.010
15	56,642,786	452	9	1	3	7	<i>LIPC</i>		0.010
16	9,242,720	438	7	3	1	9			0.010
16	81,263,410	461	9	1	3	7	<i>CDH13*</i>		0.010
17	7,107,962	350	3	7	9	1	<i>C17orf81, CLDN7</i>		0.010
22	22,009,656	610	9	1	3	7	<i>LOC649264</i>		0.010

Present/absent: methylation status.*genes identified as potential candidates; Chr: chromosome; bp: base pairs.

Methylation in known ALS candidate genes

A number of possible candidate genes for SALS have been suggested (for review, see Pasinelli and Brown, 2006). None of these candidate gene domains or their associated CpG islands showed different methylation proportions in SALS brains. In particular, no changes were seen in the methylation patterns of *SOD1*, *TARDBP*, *VEGF*, *ANG* nor any of the neurofilament genes. The methylation status of all the identified SALS candidate genes is shown in Supplementary Table IV.

Gene expression

Gene expression in SALS has been examined in eight whole-genome studies (17–24). Genes reported to have different expression patterns in

SALS that showed altered methylation in the present study were *ALK*, *ATP9B*, *C4BPB*, *GRB14*, *NSF*, *PCP4* and *RYR3* (Supplementary Table V). However, the decreases or increases of gene expression did not correlate consistently with the presence or absence of methylation. None of the above genes had methylation at their promoters, although some methylation regions contained predicted transcriptional binding sites.

Discussion

DNA methylation has been a relatively understudied area in neurodegeneration, even though epimutations have been linked to a number of pathological processes, in particular carcinogenesis. Technical limitations have until recently restricted methylation studies to individual candidate genes

Table III. Candidate genes from the SALS/control group analysis. Ranked by functional pathway.

Functional pathway	Gene	Site	SALS present	SALS absent	Control present	Control absent	SALS methylation	<i>p</i>	
Calcium dynamics	<i>CACNA1B</i>	Intronic*	4	6	0	10	Hyper	0.043	
	<i>CACNA1C</i>	Exonic	4	6	0	10	Hyper	0.043	
	<i>CACNA1E</i>	Intronic	2	8	8	2	Hypo	0.012	
	<i>FBLN5</i>	Flanking	4	6	0	10	Hyper	0.043	
	<i>HPCAL1</i>	Intronic	0	10	5	5	Hypo	0.016	
	<i>ITPRIP</i>	Gene body	0	10	4	6	Hypo	0.043	
	<i>PCP4</i>	Flanking§	4	6	0	10	Hyper	0.043	
	<i>RAMP1</i>	Intronic§	0	10	4	6	Hypo	0.043	
	<i>RYR3</i>	Intronic§	4	6	9	1	Hypo	0.029	
	<i>SLC24A3</i>	Gene body	4	6	0	10	Hyper	0.043	
Excitotoxicity	<i>ACCN1</i>	Intronic	6	4	10	0	Hypo	0.043	
	<i>SLC17A7</i>	Flanking§	0	10	5	5	Hypo	0.016	
Oxidative stress	<i>MSRA</i>	Intronic§	3	7	10	0	Hypo	0.002	
	<i>PON2</i>	Intronic§	4	6	0	10	Hyper	0.043	
	<i>PTGS1</i>	Gene body§	2	8	8	2	Hypo	0.012	
Neuronal exocytosis	<i>NRXN1</i>	Intronic	0	10	4	6	Hypo	0.043	
	<i>SYTL3</i>	Intronic§	6	4	1	9	Hyper	0.029	
Brain development	<i>ALK</i>	Intronic	8	2	3	7	Hyper	0.035	
	<i>ALK</i>	Intronic§	2	8	7	3	Hypo	0.035	
	<i>BSG</i>	Gene body*§	0	10	5	5	Hypo	0.016	
	<i>CDH13</i>	Intronic	9	1	3	7	Hyper	0.010	
	<i>CDH13</i>	Intronic	10	0	6	4	Hyper	0.043	
	<i>CDH4</i>	Intronic	1	9	6	4	Hypo	0.029	
	<i>CDH4</i>	Intronic	2	8	7	3	Hypo	0.035	
	<i>CNTN5</i>	Intronic§	3	7	9	1	Hypo	0.010	
	<i>DOK5</i>	Intronic§	6	4	10	0	Hypo	0.043	
	<i>METRN1</i>	Flanking	0	10	4	6	Hypo	0.043	
	<i>NRP1</i>	Intronic§	1	9	6	4	Hypo	0.029	
	<i>PCDH21</i>	Intronic	0	10	4	6	Hypo	0.043	
	<i>ROBO1</i>	Intronic§	5	5	0	10	Hyper	0.016	
	<i>SIM1</i>	Intronic§	4	6	9	1	Hypo	0.029	
	Collagen	<i>COL23A1</i>	Intronic	4	6	0	10	Hyper	0.043
		<i>COL4A1</i>	Gene body§	4	6	0	10	Hyper	0.043
<i>COL6A3</i>		Intronic	3	7	8	2	Hypo	0.035	
<i>COL6A6</i>		Intronic§	10	0	4	6	Hyper	0.005	
Neurological conditions	<i>CA8</i>	Intronic§	4	6	9	1	Hypo	0.029	
	<i>CIITA</i>	Flanking	5	5	0	10	Hyper	0.016	
	<i>DAB1</i>	Intronic	9	1	3	7	Hyper	0.010	
	<i>DMD</i>	Intronic§	0	10	4	6	Hypo	0.043	
	<i>LRRK1</i>	Intronic	10	0	6	4	Hyper	0.043	
	<i>PACRG</i>	Intronic§	0	10	6	4	Hypo	0.005	
	<i>SNTB1</i>	Intronic§	4	6	9	1	Hypo	0.029	
	<i>SNTG2</i>	Intronic§	5	5	0	10	Hyper	0.016	
Ubiquitin	<i>UBE2E3</i>	Flanking§	1	9	6	4	Hypo	0.029	
DNA repair	<i>PLA2G4C</i>	Gene body	0	10	4	6	Hypo	0.043	
Survival/apoptosis	<i>CERKL</i>	Intronic§	1	9	6	4	Hypo	0.029	
	<i>STAT3</i>	Gene body	10	0	6	4	Hyper	0.043	
Neurotrophic factor	<i>GFRA1</i>	Intronic	0	10	4	6	Hypo	0.043	
	<i>GFRA2</i>	Intronic	0	10	4	6	Hypo	0.043	
	<i>IGF2BP3</i>	Intronic	1	9	7	3	Hypo	0.010	
Xenobiotic protection	<i>AS3MT</i>	Intronic	7	3	2	8	Hyper	0.035	
	<i>PAPSS2</i>	Intronic	3	7	8	2	Hypo	0.035	

*Methylation region in CpG island; §: Methylation region contains a predicted transcription factor binding site; Gene body: methylation region covers both exons and introns. Genes that appear twice contain two methylation regions.

(7,8). Epigenome analyses now allow the examination of both the distribution of methylation in the brain and comparisons between disease and control phenotypes.

Of interest was the wide variation in the total number of intervals, with some brains showing over twice the number of others. This variation was

consistent across all genomic regions and was similar in SALS and control brains. However, the average number of intervals did not differ between SALS and control subjects. Our finding of large inter-individual differences in interval number is surprising, given a previous report of a consistent methylation pattern between brain samples (25). Methylation status can

Table IV. Comparison of methylation regions in individual SALS brains. Methylation regions where at least five SALS patients had differences (either methylated or unmethylated) are shown. Ranked by decreasing number of patients with differences.

Chr	Start position	Length (bp)	SALS methylated	SALS unmethylated	Gene domain	CpG island	SALS methylation
8	1,732,991	1660	0	10			Hypo
17	210,280	907	0	10	<i>RPH3AL*</i>		Hypo
17	122,587	925	0	8	<i>RPH3AL*</i>	Yes	Hypo
17	76,331,968	1091	0	7	<i>KLAA1303</i>		Hypo
13	111,717,247	1053	7	0		Yes	Hyper
8	145,475,709	780	0	6	<i>BOPI, HSF1</i>		Hypo
9	114,882,049	246	0	6			Hypo
10	134,704,773	414	0	6			Hypo
16	22,955,808	833	0	6			Hypo
16	88,181,317	1104	0	6	<i>CPNE7*</i>	Yes	Hypo
19	41,457,506	688	0	6		Yes	Hypo
19	41,467,665	591	0	6		Yes	Hypo
1	224,876,724	430	6	0	<i>ITPKB</i>		Hyper
10	792,978	814	6	0			Hyper
1	3,738,455	653	0	5	<i>KLAA0562</i>		Hypo
3	278,451	809	0	5	<i>CHL1*</i>		Hypo
4	16,544,148	209	0	5			Hypo
4	149,680,380	229	0	5			Hypo
4	188,117,410	14,408	0	5		Yes	Hypo
4	190,521,648	774	0	5		Yes	Hypo
5	4,919,316	548	0	5		Yes	Hypo
5	17,629,808	235	0	5	<i>LOC402207, LOC729719, LOC729724, LOC729731, LOC391761, LOC729735</i>		Hypo
7	2,041,299	1517	0	5	<i>MAD1L1</i>		Hypo
9	45,332,510	439	0	5		Yes	Hypo
12	9,531,025	285	0	5	<i>OYOS</i>		Hypo
12	94,554,802	513	0	5			Hypo
12	105,425,357	639	0	5	<i>POLR3B</i>		Hypo
14	45,672,710	907	0	5			Hypo
17	42,024,723	407	0	5	<i>NSF*</i>		Hypo
19	41,470,201	715	0	5		Yes	Hypo
19	41,475,265	714	0	5		Yes	Hypo
19	41,477,787	715	0	5		Yes	Hypo
19	41,480,322	688	0	5		Yes	Hypo
19	41,485,389	516	0	5	<i>LOC100131606</i>	Yes	Hypo
19	41,487,904	516	0	5	<i>LOC100131606, LOC100127980</i>	Yes	Hypo
21	45,877,357	1465	0	5	<i>LOC728117</i>	Yes	Hypo
22	48,047,211	1186	0	5			Hypo
X	54,936,684	235	5	0			Hyper
3	11,958,423	2399	5	0			Hyper
7	157,410,089	790	5	0	<i>PTPRN2</i>		Hyper
11	122,780,554	274	5	0			Hyper
11	125,601,114	688	5	0	<i>FAM118B</i>		Hyper
15	67,574,403	273	5	0	<i>LOC145837</i>		Hyper
18	74,734,887	207	5	0			Hyper
18	75,227,877	803	5	0	<i>ATP9B</i>		Hyper
21	43,007,510	630	5	0	<i>PDE9A</i>		Hyper
22	34,303,699	447	5	0			Hyper

*Genes identified as potential candidates; Chr: chromosome; bp: base pairs.

change during ageing (26), but numbers of intervals did not correlate with age in our subjects. The samples in our study were taken from one area of the brain so regional differences are unlikely to account for the variation (25). The previous study that looked at 1505 potential methylation regions in CpG islands (25) (compared to our 64,424 regions across the genome) used a different technique of methylation analysis, so

direct comparison is difficult. Our findings suggest that differences in brain methylation status between individuals are more marked than previously thought.

Some chromosomes, in particular chromosome Y, showed varying methylation between SALS and control brains at CpG islands and gene domains. This is of interest since most series have shown a male predominance in SALS. An analysis of a mixed

gender cohort would be required for further characterization of gender differences in brain methylation.

The methylation rate of CpG islands we found was similar to that in the Human Epigenome Project (27) as well as a recent MeDIP study of methylation in other tissues (28). About half the intervals we detected were within 10 kb of genes, the remaining half being at intergenic sites. Previous epigenetic studies in human brain have investigated methylation patterns near genes only (25,29–32) rather than across the whole genome. This study indicates that a large proportion of methylation in the human brain is at intergenic sites, where it may regulate genes over large distances (27).

Several biochemical pathways were identified where methylation differences could be relevant to the pathogenesis of SALS, including calcium ion transport. The current literature implicates several mechanisms in the development of SALS and we consider the following groups of genes to be of particular interest as candidate genes for SALS.

Calcium dynamics

Among the genes with differing methylation levels in SALS the most significant enrichment was for those related to calcium ion transport. Dysregulation of calcium homeostasis is thought to play a central role in neuronal loss in SALS (33). Identified genes included those for calcium channels (*CACNA1B*, *CACNA1C*, and *CACNA1E*) and a sodium-calcium exchanger (*SLC24A3*). Of interest is *CACNA1B*, which is involved in neurotransmitter release (34) and was hypermethylated in patients compared to controls. Another attractive candidate is *HPCAL1*, a calcium sensor for neuronal signalling (35) that showed relative hypomethylation in patients. A previous study showed the gene for the calcium channel *RYR3* is down-regulated in the motor cortex of SALS patients (17) and our study found *RYR3* was hypomethylated at a binding site for homeodomain transcription factors, including the repressor TGIF2. The lack of methylation at this site may allow repressors to bind and down-regulate *RYR3* in SALS. *PCP4* is a neuronally-expressed calcium binding protein that has previously been shown to be down-regulated in SALS motor cortex (22). Our study identified a hypermethylated region upstream to *PCP4* in the SALS patients. This region contained a homeodomain transcriptional binding site that may be blocked in the brain and contribute to the down-regulation seen in this disease. *RAMP1* is involved in the trafficking of calcitonin receptors (36) and was hypomethylated at a transcriptional binding site in SALS. Other candidate genes involved in calcium dynamics included *FBLN5* and *ITPR1P*.

Excitotoxicity

Glutamate excitotoxicity has long been suspected to underlie cell death in SALS (37) and differences in glutamate handling in motor neurons are thought to account for the site-specific damage seen in this disease (37). *SLC17A7* is a sodium-dependent phosphate transporter that functions in glutamate transport at synaptic vesicles (38). A region downstream to *SLC17A7* that contained brain-specific transcriptional binding sites was hypomethylated in all SALS patients. Excitotoxicity also involves the influx of sodium ions (39), and one of our candidates, *ACCNI*, is a sodium ion transporter linked to excitotoxicity in Parkinson's disease (40). This gene was hypomethylated in SALS.

Oxidative stress

Oxidative stress is implicated in SALS and three of our differently-methylated genes protect against oxidative damage. *PON2* has been the subject of SNP association studies in SALS (41) and was hypermethylated at a transcriptional binding site in SALS. *MSRA* repairs oxidative damage and is highly expressed in nervous tissue (15). *PTGS1*, known commonly as cyclo-oxygenase, protects cells from oxidative damage and has been shown to play a role in neuroinflammation (39). Both the latter genes were hypomethylated at transcriptional binding sites.

Neuronal exocytosis

Neurexins are cell surface proteins that mediate the assembly of presynaptic terminals and couple calcium channels to synaptic vesicle exocytosis (42). *NRXN1* was unmethylated in all SALS brains while controls showed variable methylation. *SYTL3* interacts with *NRXN1* and is likely to play a part in synaptic vesicle exocytosis (43). *SYTL3* was hypermethylated at an intronic homeodomain binding site in SALS. *NSF*, essential to vesicle fusion at synaptic terminals (44), is down-regulated in SALS (22) and was hypomethylated at a predicted repressor binding site. *RPH3AL*, the human ortholog of a rat GTPase involved in calcium-regulated exocytosis (45), was hypomethylated at two sites, one located upstream of the gene that was unmethylated in all SALS patients. The other site was a CpG island located within *RPH3AL* that was unmethylated in eight of the 10 SALS patients. These two sites contain transcriptional binding motifs for homeodomain proteins, neurogenesis and apoptosis.

Other neuromuscular conditions

DMD, the mutated gene responsible for Duchenne and Becker muscular dystrophies (46), was hypomethylated at a predicted repressor binding site in

SALS. Two syntrophin genes, *SNTB1* (47) and *SNTG2* (48), are associated with dystrophin in cells and showed SALS hypo- and hyper-methylation, respectively. *LRRK1* and *PACRG* are associated with Parkinson's disease. Other identified genes included *CA8*, *CIITA* and *DAB1*, which have been associated with neuromuscular conditions in humans or animal models.

Brain development

This group of candidate genes is involved in brain development and neural network formation, and includes the cell adhesion molecules *PCDH21*, *CDH4*, *CDH13* and *CHL1* (49). Of interest is *NRP1*, a receptor for vascular endothelial growth factor (50), which itself is implicated in SALS (51). *NRP1* was hypomethylated in SALS at a transcriptional binding site.

Neurotrophic factors, ubiquitin, apoptosis, xenobiotics, DNA repair, collagen and microRNA

Impairment of neurotrophic factors may lead to selective motor neuron death in ALS (52). GDNF is up-regulated in SALS (18) and we found the GDNF receptors *GFRA1* and *GFRA2* were hypomethylated in all SALS patients. The transcriptional repressor *IGF2BP3*, which acts on the neurotrophic factor IGF2, was hypomethylated in SALS. Other candidates were related to ubiquitination (*UBE2E3*), neuronal survival and apoptosis (*CERKL* and *STAT3*), xenobiotic detoxification (*PAPSS2* and *AS3MT*), DNA repair (*PLA2G4C*) and the collagen genes (*COL23A1*, *COL4A1*, *COL6A3* and *COL6A6*). One neuronally-specific microRNA, *miR-124-A3*, was located upstream to a hypomethylated region in SALS.

Expression studies

Genes reported to have different expression in SALS were investigated to see if expression could be linked to changes in methylation. Genes with differences in both their methylation and expression patterns were *ALK*, *RYR3*, *GRB14*, *PCP4*, *C4BPB*, *NSF* and *ATP9B*, some of which have been discussed above. *ATP9B* is down-regulated in SALS (17) and was hypermethylated in SALS at binding sites for factors that regulate the cellular calcium response and neuronal excitability, and are involved in synaptic transmission and neuronal apoptosis. Although the methylation differences in these seven genes were outside promoters, the proximity of the methylation regions to transcriptional binding sites may allow these regions to influence expression.

A limitation of this study was the modest numbers of disease and control brains analysed.

The large number of methylated sites detected means that a considerable number of differences are likely to be due to statistical chance. We did not apply corrections for multiple testing, given the small numbers of subjects and the controversy in applying these corrections to whole-genome studies (53). The current high experimental cost of these whole-genome studies and the scarcity of human brain tissue (both diseased and normal) will limit this approach to modest numbers for the foreseeable future. However, accurate age-, sex- and ethnicity-matching in this study reduced the sources of variation to a minimum. In addition, the use of pathway analysis identified groups of genes in which methylation differences may well have biological significance.

This study represents a solid starting point for further epigenetic analyses in SALS. Given the potential for large numbers of rarely-occurring methylation differences to interact to produce the disease, newer whole-genome technologies for methylation analysis would be of value to confirm our results. One such technique is chromatin immunoprecipitation combined with deep sequencing (54) which would allow further characterization of the regions of interest. This would enable DNA methylation to be compared with both single nucleotide and copy number variation in a single study, in addition to the detection of these variations in only a small percentage of the tissue. Since the normal variation in brain DNA methylation is not known, tissue from at least 100 SALS patients and 100 controls should be examined. In addition, because DNA methylation probably varies in different parts of the CNS, samples should be examined from a region maximally affected by the disease (the spinal cord), a region affected by the disease but with less cell loss (the frontal cortex), and a region not known to be affected by the disease (e.g. the occipital cortex). The numbers of patients (and in particular controls) needed, and the expense of such a study, will require an international collaborative effort. This will be feasible if tissue banks can be found that have frozen brain and spinal cord samples to allow extraction of high-quality genomic DNA from sufficient numbers of SALS and age- and gender-matched controls.

In conclusion, this study is the first to examine methylation in normal human brain tissue across the whole genome and represents a resource for future studies. It has revealed a number of differences in SALS patients. Although no altered methylation pattern was common to all SALS patients, multiple rare changes in methylation may have an important role in the pathogenesis of SALS. Novel candidates detected by this genome-wide methylation study include genes involved in calcium homeostasis, excitotoxicity and oxidative stress, all of which have been implicated in SALS.

Methylation remains an exciting field in SALS research both in terms of explaining pathogenesis and identifying interacting environmental factors. Of particular relevance to patients with SALS, finding an epigenetic cause for this disease could open up new avenues for therapy (3).

Acknowledgements

This work was supported by the Aimee Stacey Memorial Fund and the Australian Brain Foundation. JMM is a Bill Gole Fellow of the MND Research Institute of Australia. Tissue samples were provided by the Using our Brains donor programme and the New South Wales, Victorian, and South Australian Tissue Resource Centres, which are part of the Australian Brain Bank Network. The staff of Genpathway assisted with data analysis. Stephen Kum Jew provided technical support. We thank all donors and their families for providing tissue for this project.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

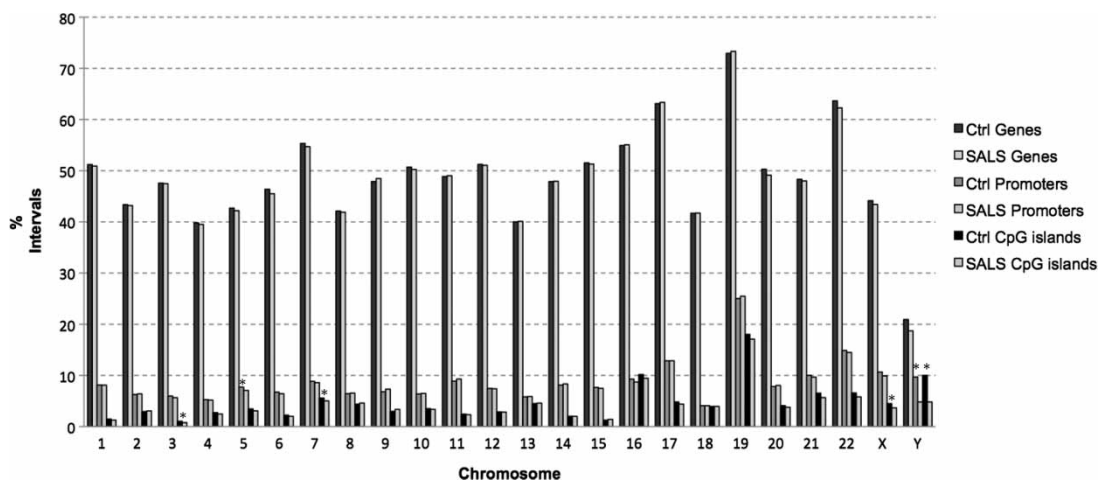
- Logroscino G, Traynor BJ, Hardiman O, Chio A, Couratier P, Mitchell JD, et al. Descriptive epidemiology of amyotrophic lateral sclerosis: new evidence and unsolved issues. *J Neurol Neurosurg Psychiatry*. 2008;79:6–11.
- Martin DI, Ward R, Suter CM. Germline epimutation: a basis for epigenetic disease in humans. *Ann N Y Acad Sci*. 2005;1054:68–77.
- Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. *Nature*. 2004;429:457–63.
- Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev*. 2002;16:6–21.
- Feinberg AP. Phenotypic plasticity and the epigenetics of human disease. *Nature*. 2007;447:433–40.
- Esteller M. Epigenetics in cancer. *N Engl J Med*. 2008;358:1148–59.
- Oates N, Pamphlett R. An epigenetic analysis of SOD1 and VEGF in ALS. *Amyotroph Lateral Scler*. 2007;8:83–6.
- Morahan JM, Yu B, Trent RJ, Pamphlett R. Are metallothionein genes silenced in ALS? *Toxicol Lett*. 2007;168:83–7.
- Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, et al. Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. *Cell*. 2006;126:1189–201.
- Keshet I, Schlesinger Y, Farkash S, Rand E, Hecht M, Segal E, et al. Evidence for an instructive mechanism of de novo methylation in cancer cells. *Nat Genet*. 2006;38:149–53.
- Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, et al. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat Genet*. 2005;37:853–62.
- Brooks BR, Miller RG, Swash M, Munsat TL. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Other Motor Neuron Disord*. 2000;1:293–9.
- Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol*. 2003;4:3.
- Quandt K, Frech K, Karas H, Wingender E, Werner T. MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res*. 1995;23:4878–84.
- Moskovitz J, Jenkins NA, Gilbert DJ, Copeland NG, Jursky F, Weissbach H, et al. Chromosomal localization of the mammalian peptide-methionine sulphoxide reductase gene and its differential expression in various tissues. *Proc Natl Acad Sci U S A*. 1996;93:3205–8.
- Pasinelli P, Brown RH. Molecular biology of amyotrophic lateral sclerosis: insights from genetics. *Nat Rev Neurosci*. 2006;7:710–23.
- Wang XS, Simmons Z, Liu W, Boyer PJ, Connor JR. Differential expression of genes in amyotrophic lateral sclerosis revealed by profiling the post mortem cortex. *Amyotroph Lateral Scler*. 2006;7:201–10.
- Tanaka F, Niwa J, Ishigaki S, Katsuno M, Waza M, Yamamoto M, et al. Gene expression profiling toward understanding of ALS pathogenesis. *Ann N Y Acad Sci*. 2006;1086:1–10.
- Ishigaki S, Niwa J, Ando Y, Yoshihara T, Sawada K, Doyu M, et al. Differentially expressed genes in sporadic amyotrophic lateral sclerosis spinal cords: screening by molecular indexing and subsequent cDNA microarray analysis. *FEBS Lett*. 2002;531:354–8.
- Malaspina A, Kaushik N, de Bellerocche J. Differential expression of 14 genes in amyotrophic lateral sclerosis spinal cord detected using gridded cDNA arrays. *J Neurochem*. 2001;77:132–45.
- Jiang Y-M, Yamamoto M, Kobayashi Y, Yoshihara T, Liang Y, Terao S, et al. Gene expression profile of spinal motor neurons in sporadic amyotrophic lateral sclerosis. *Ann Neurol*. 2005;57:236–51.
- Lederer CW, Torrisi A, Pantelidou M, Santama N, Cavallaro S. Pathways and genes differentially expressed in the motor cortex of patients with sporadic amyotrophic lateral sclerosis. *BMC Genomics*. 2007;8:26.
- Dangond F, Hwang D, Camelo S, Pasinelli P, Frosch MP, Stephanopoulos G, et al. Molecular signature of late-stage human ALS revealed by expression profiling of post mortem spinal cord grey matter. *Physiol Genomics*. 2004;16:229–39.
- Offen D, Barhum Y, Melamed E, Embacher N, Schindler C, Ransmayr G. Spinal cord mRNA profile in patients with ALS: comparison with transgenic mice expressing the human SOD1 mutant. *J Mol Neurosci*. 2008 Jul 24 [Epub ahead of print].
- Ladd-Acosta C, Pevsner J, Sabunciyan S, Yolken RH, Webster MJ, Dinkins T, et al. DNA methylation signatures within the human brain. *Am J Hum Genet*. 2007;81:1304–15.
- Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, et al. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci U S A*. 2005;102:10604–9.
- Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, et al. DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat Genet*. 2006;38:1378–85.
- Rakyan VK, Down TA, Thorne NP, Flicek P, Kulesha E, Graf S, et al. An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). *Genome Res*. 2008;18:1518–29.
- Rollins RA, Haghghi F, Edwards JR, Das R, Zhang MQ, Ju J, et al. Large-scale structure of genomic methylation patterns. *Genome Res*. 2006;16:157–63.

30. Schilling E, Rehli M. Global, comparative analysis of tissue-specific promoter CpG methylation. *Genomics*. 2007;90:314–23.
31. Siegmund KD, Connor CM, Campan M, Long TI, Weisenberger DJ, Binizskiewicz D, et al. DNA methylation in the human cerebral cortex is dynamically regulated throughout the life span and involves differentiated neurons. *PLoS ONE*. 2007;2:895.
32. Illingworth R, Kerr A, Desousa D, Jorgensen H, Ellis P, Stalker J, et al. A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS Biol*. 2008;6:22.
33. Meldrum B, Garthwaite J. Excitatory amino acid neurotoxicity and neurodegenerative disease. *Trends Pharmacol Sci*. 1990;11:379–87.
34. Williams ME, Brust PF, Feldman DH, Patthi S, Simerson S, Maroufi A, et al. Structure and functional expression of an omega-conotoxin-sensitive human N-type calcium channel. *Science*. 1992;257:389–95.
35. Spilker C, Gundelfinger ED, Braunewell KH. Evidence for different functional properties of the neuronal calcium sensor proteins VILIP-1 and VILIP-3: from subcellular localization to cellular function. *Biochim Biophys Acta*. 2002;1600:118–27.
36. Foord SM, Marshall FH. RAMPs: accessory proteins for seven transmembrane domain receptors. *Trends Pharmacol Sci*. 1999;20:184–7.
37. Williams TL, Day NC, Ince PG, Kamboj RK, Shaw PJ. Calcium-permeable alpha-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid receptors: a molecular determinant of selective vulnerability in amyotrophic lateral sclerosis. *Ann Neurol*. 1997;42:200–7.
38. Takamori S, Rhee JS, Rosenmund C, Jahn R. Identification of a vesicular glutamate transporter that defines a glutamatergic phenotype in neurons. *Nature*. 2000;407:189–94.
39. Choi DW. Ionic dependence of glutamate neurotoxicity. *J Neurosci*. 1987;7:369–79.
40. Joch M, Ase AR, Chen CX, MacDonald PA, Kontogianna M, Corera AT, et al. Parkin-mediated monoubiquitination of the PDZ protein PICK1 regulates the activity of acid-sensing ion channels. *Mol Biol Cell*. 2007;18:3105–18.
41. Cronin S, Greenway MJ, Prehn JH, Hardiman O. Paraoxonase promoter and intronic variants modify risk of sporadic amyotrophic lateral sclerosis. *J Neurol Neurosurg Psychiatry*. 2007;78:984–6.
42. Missler M, Zhang W, Rohlmann A, Kattenstroth G, Hammer RE, Gottmann K, et al. Alpha-neurexins couple Ca²⁺ channels to synaptic vesicle exocytosis. *Nature*. 2003;423:939–48.
43. Fukuda M, Mikoshiba K. Synaptotagmin-like protein 1-3: a novel family of C-terminal-type tandem C2 proteins. *Biochem Biophys Res Commun*. 2001;281:1226–33.
44. Hoyle J, Phelan JP, Bermingham N, Fisher EM. Localization of human and mouse N-ethylmaleimide-sensitive factor (NSF) gene: a two-domain member of the AAA family that is involved in membrane fusion. *Mamm Genome*. 1996;7:850–2.
45. Haynes LP, Evans GJ, Morgan A, Burgoyne RD. A direct inhibitory role for the Rab3-specific effector, Noc2, in Ca²⁺-regulated exocytosis in neuroendocrine cells. *J Biol Chem*. 2001;276:9726–32.
46. Worton RG, Thompson MW. Genetics of Duchenne muscular dystrophy. *Annu Rev Genet*. 1988;22:601–29.
47. Ahn AH, Kunkel LM. Syntrophin binds to an alternatively spliced exon of dystrophin. *J Cell Biol*. 1995;128:363–71.
48. Piluso G, Mirabella M, Ricci E, Belsito A, Abbondanza C, Servidei S, et al. Gamma1- and gamma2-syntrophins, two novel dystrophin-binding proteins localized in neuronal cells. *J Biol Chem*. 2000;275:15851–60.
49. Shapiro L, Love J, Colman DR. Adhesion molecules in the nervous system: structural insights into function and diversity. *Annu Rev Neurosci*. 2007;30:451–74.
50. Sulpice E, Plouet J, Berge M, Allanic D, Tobelem G, Merkulova-Rainon T. Neuropilin-1 and neuropilin-2 act as coreceptors, potentiating proangiogenic activity. *Blood*. 2008;111:2036–45.
51. Lambrechts D, Storkebaum E, Morimoto M, Del-Favero J, Desmet F, Marklund SL, et al. VEGF is a modifier of amyotrophic lateral sclerosis in mice and humans and protects motor neurons against ischaemic death. *Nat Genet*. 2003;35:507–13.
52. Ekester E. Neurotrophic factors and amyotrophic lateral sclerosis. *Neurodegener Dis*. 2004;1:88–100.
53. Rice TK, Schork NJ, Rao DC. Methods for handling multiple testing. *Adv Genet*. 2008;60:293–308.
54. Robertson G, Hirst M, Bainbridge M, Bilenky M, Zhao Y, Zeng T, et al. Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat Methods*. 2007;4:651–7.

Supplementary Table I. Clinical details of SALS patients and controls.

Group	Region first affected	Disease duration (years)*	Cause of death	Age at death (years)	PM delay (h)
SALS	Lower limb	3	ALS	54	9
SALS	Upper limb	2	ALS	47	13
SALS	Lower limb	2	ALS	70	12
SALS	Bulbar	1	ALS	53	18
SALS	Lower limb	1	ALS	54	19
SALS	Lower limb	2	ALS	68	5
SALS	Lower limb	10	ALS	84	23
SALS	Bulbar	2	ALS	62	28
SALS	Lower limb	1	ALS	77	6
SALS	Lower limb	1	ALS	56	8
Control	-	-	IHD	55	7
Control	-	-	IHD	44	13
Control	-	-	IHD	69	16
Control	-	-	IHD	54	29
Control	-	-	Cardiomyopathy	53	16
Control	-	-	IHD	69	13
Control	-	-	Colon carcinoma	85	9
Control	-	-	IHD	60	13
Control	-	-	Suicide	74	16
Control	-	-	IHD	59	12

SALS patients are age-matched with sequential controls. *The difference between the date of diagnosis of SALS and the date of death, rounded to the nearest year. PM: post mortem; IHD: ischaemic heart disease.



Supplementary Figure 1. Interval frequencies in different genomic areas by chromosome. Interval frequencies across different chromosomes were similar in SALS and control subjects, apart from those labelled with asterisks (significantly different at $p < 0.05$). Genes: gene domains; Ctrl: control; % Intervals: per cent intervals in each group associated with each genomic feature (e.g. the first bar showing 52% of intervals in controls were associated with genes). The most significant differences were in the Y chromosome.

Supplementary Table II. Distribution of genomic regions by methylation status.

Genomic region	Total number	All subjects (average)		Control (average)		SALS (average)	
		Number methylated	% methylated	Number methylated	% methylated	Number methylated	% methylated
Gene and pseudogene domains	40,834	6957	17.0	7065	17.3	6849	16.7
Promoters	40,834	2793	6.8	2844	6.9	2742	6.7
Promoters with CpG islands	17,947	1346	7.4	1382	7.7	1310	7.3
Promoters without CpG islands	22,887	1447	6.3	1462	6.4	1432	6.3

Supplementary Table III. Gene ontology functional categories that were over-represented in differently-methylated gene domains in SALS. Ranked by increasing *p*-values.

Gene ontology functional annotation	Count	<i>p</i>
Calcium ion binding	28	0.00002
Extracellular matrix	9	0.00005
Multicellular development	48	0.00006
Anatomical structure development	45	0.00007
Multicellular process	66	0.00014
Basement membrane	7	0.00017
Cell adhesion	22	0.00021
Biological adhesion	22	0.00021
Developmental process	58	0.00036
System development	37	0.00037

Supplementary Table IV. Methylation status of SALS previously-identified candidate genes (listed alphabetically).

Gene	SALS present	SALS absent	Control present	Control absent	Fisher's exact <i>p</i>	Methylation site	Within CpG	Other CpG islands in the gene domain
<i>ABCA1</i>	6	4	6	4	0.675	Downstream	No	Promoter, Intron
	2	8	3	7	0.5	Downstream	No	
	0	10	2	8	0.237	Intron	No	
	0	10	1	9	0.5	Intron	No	
	9	1	6	4	0.152	Exon	No	
	3	7	3	7	0.686	Intron	No	
	3	7	2	8	0.5	Exon	No	
	7	3	7	3	0.686	Intron	No	
	1	9	1	9	0.763	Intron	No	
	1	9	0	10	0.5	Intron	No	
	4	6	6	4	0.328	Intron	No	
	9	1	9	1	0.763	Intron	No	
	1	9	0	10	0.5	Intron	No	
	0	10	0	10	1	–	–	
<i>ABCD1</i>	5	5	6	4	0.5	Upstream	No	Promoter
	3	7	2	8	0.5	Upstream	No	
<i>ALS2</i>	0	10	1	9	0.5	Upstream	No	
	0	10	1	9	0.5	Intron	No	Promoter, Upstream
<i>ANG</i>	1	9	0	10	0.5	Intron	No	Promoter, Exon
<i>AR</i>	0	10	1	9	0.5	Intron	No	
	10	0	10	0	1	Upstream	Yes	Promoter, Exon, Exon
<i>ARHGEF10</i>	0	10	1	9	0.5	Intron	No	
	8	2	9	1	0.5	Intron	Yes	
	9	1	9	1	0.763	Intron	No	
	10	0	10	0	1	Intron	No	
	6	4	7	3	0.5	Intron	Yes	
	1	9	0	10	0.5	Intron	No	
	10	0	9	1	0.5	Intron	Yes	
	2	8	3	7	0.5	Intron	No	
	6	4	6	4	0.675	Intron	No	
	1	9	0	10	0.5	Intron	No	
	8	2	8	2	0.709	Intron	No	
	10	0	10	0	1	Intron	No	
	8	2	7	3	0.5	Intron	No	
	2	8	1	9	0.5	Intron	No	
	4	6	4	6	0.675	Exon	No	
	6	4	5	5	0.5	Intron	No	
	2	8	2	8	0.709	Intron	No	
2	8	1	9	0.5	Intron	No		
1	9	0	10	0.5	Downstream	No		
<i>ARSA</i>	0	10	0	10	1	–	–	Promoter
<i>BSCL2</i>	0	10	0	10	1	–	–	Upstream, Downstream, Intron
<i>CHMP2B</i>	0	10	0	10	1	–	–	Promoter
<i>COXC</i>	0	10	0	10	1	–	–	Promoter
<i>DCTN1</i>	0	10	0	10	1	–	–	None

<i>DNM2</i>	1	9	0	10	0.5	Intron	No	Promoter
	1	9	0	10	0.5	Intron	No	
	1	9	1	9	0.763	Downstream	No	
<i>EGR2</i>	1	9	0	10	0.5	Upstream	No	Promoter, Exon Downstream
<i>GAN</i>	2	8	3	7	0.5	Promoter	Yes	–
<i>GARS</i>	0	10	0	10	1	–	–	Promoter
<i>GDAP1</i>	0	10	0	10	1	–	–	Promoter
<i>G7B1</i>	0	10	0	10	1	–	–	Exon
<i>HEXA</i>	0	10	0	10	1	–	–	Exon
<i>HEXB</i>	0	10	0	10	1	–	–	Promoter
<i>HMBS</i>	0	10	0	10	1	–	–	Promoter, Downstream, Downstream
<i>HSP22</i>	5	5	4	6	0.5	Upstream	No	–
<i>HSP27</i>	0	10	0	10	1	–	–	Promoter
<i>HSP60</i>	0	10	0	10	1	–	–	Promoter
<i>IGHMBP2</i>	10	0	10	0	1	Intron	No	Promoter
<i>IKBKAP</i>	0	10	3	7	0.105	Intron	No	Promoter
<i>K1F1B</i>	0	10	0	10	1	–	–	Promoter
<i>KIF21A</i>	7	3	8	2	0.5	Intron	No	Exon
	4	6	7	3	0.185	Intron	No	
<i>KIF5A</i>	1	9	3	7	0.291	Downstream	No	Promoter, Upstream, Downstream
<i>L1CAM</i>	0	10	0	10	1	–	–	Upstream
<i>LITAF</i>	2	8	2	8	0.709	Downstream	No	Promoter
	0	10	1	9	0.5	Intron	No	
<i>LMNA</i>	9	1	7	3	0.291	Downstream	No	Promoter, Exon
<i>MFN2</i>	0	10	0	10	1	–	–	Promoter
<i>MtRNAS</i>	0	10	0	10	1	–	–	Promoter
<i>MTMR2</i>	6	4	8	2	0.314	Intron	No	Promoter
	10	0	10	0	1	Intron	No	
	6	4	4	6	0.328	Intron	No	
<i>NDRG1</i>	10	0	10	0	1	Intron	No	–
	0	10	2	8	0.237	Promoter	Yes	
	1	9	0	10	0.5	Upstream	No	
	4	6	5	5	0.5	Upstream	No	
<i>NEFH</i>	0	10	0	10	1	–	–	Promoter, Upstream
<i>NEFL</i>	2	8	5	5	0.175	Upstream	No	Promoter, Downstream
<i>NGFR</i>	0	10	0	10	1	–	–	Promoter
<i>NIPA1</i>	0	10	0	10	1	–	–	Promoter, Downstream
<i>P0</i>	0	10	1	9	0.5	Upstream	No	Upstream, Upstream
<i>PHYH</i>	0	10	0	10	1	–	–	Promoter, Upstream
<i>PLP1</i>	0	10	0	10	1	–	–	–
<i>PMP22</i>	10	0	10	0	1	Intron	No	Promoter, Exon
<i>PRPH</i>	0	10	1	9	0.5	Downstream	No	Promoter
								Downstream, Downstream, Exon, Intron,
<i>PRX</i>	1	9	1	9	0.763	Downstream	No	Upstream
	7	3	3	7	0.089	Exon	No	
	9	1	8	2	0.5	Intron	Yes	
	9	1	10	0	0.5	Intron	No	
<i>RAB7A</i>	0	10	0	10	1	–	–	Promoter, Intron
<i>SBF2</i>	10	0	10	0	1	Intron	No	Promoter, Upstream, Upstream
	10	0	9	1	0.5	Intron	No	
<i>SETX</i>	3	7	3	7	0.686	Intron	No	Promoter
	1	9	5	5	0.07	Intron	No	
	3	7	3	7	0.686	Exon	No	
<i>SH3TC2</i>	1	9	1	9	0.763	Intron	No	–
	9	1	10	0	0.5	Intron	No	
<i>SLC12A6</i>	0	10	0	10	1	–	–	Promoter
<i>SMN1</i>	0	10	0	10	1	–	–	Promoter
<i>SMN2</i>	3	7	1	9	0.291	Intron	No	Promoter
<i>SOD1</i>	0	10	0	10	1	–	–	Promoter
<i>SOX10</i>	0	10	0	10	1	–	–	Promoter
<i>SPAST</i>	0	10	0	10	1	–	–	Promoter, Downstream
<i>SPG20</i>	0	10	0	10	1	–	–	Promoter, Downstream
<i>SPG21</i>	0	10	0	10	1	–	–	Promoter
<i>SPG3A</i>	0	10	0	10	1	–	–	Promoter
								Promoter, Intron, Intron, Intron, Intron,
<i>SPG7</i>	1	9	4	6	0.152	Exon	No	Downstream
	2	8	3	7	0.5	Downstream	Yes	
<i>SPTLC1</i>	9	1	7	3	0.291	Intron	Yes	Promoter
	0	10	3	7	0.105	Intron	No	
	0	10	1	9	0.5	Intron	No	
<i>TARDBP</i>	0	10	0	10	1	–	–	Promoter

<i>VAPB</i>	10	0	10	0	1	Intron	No	Promoter
<i>VEGF</i>	0	10	0	10	1	–	–	Promoter
<i>YARS</i>	2	8	5	5	0.175	Intron	No	Promoter

*Gene symbol according to NCBI version 36; §: CpG islands as defined by the UCSC Genome Browser Build March 2006; Upstream, downstream: flanking regions; dashes: indicate the gene domain was unmethylated in all subjects; present/absent: methylation status.

Supplementary Table V. Genes with methylation and previously-identified expression differences in SALS.

Gene	Analysis	SALS methylation	Methylation site	Expression	Reference
<i>ALK</i>	Group	Methylated	Intronic	Down	Dangond et al. 2003
<i>ALK</i>	Group	Unmethylated	Intronic§	Down	Dangond et al. 2003
<i>ATP9B</i>	Individual	Methylated	Intronic§	Down	Wang et al. 2006
<i>C4BPB</i>	Group	Methylated	Intronic	Down	Dangond et al. 2003
<i>GRB14</i>	Group	Unmethylated	Downstream§	Down	Wang et al. 2006
<i>NSF</i>	Individual	Unmethylated	Intronic§	Down	Lederer et al. 2007
<i>PCP4</i>	Group	Methylated	Upstream§	Down	Lederer et al. 2007
<i>RYR3</i>	Group	Unmethylated	Intronic§	Down	Wang et al. 2006

§Methylation region contains a predicted transcription factor binding site.