29. Kohlhaw, G. B., Leary, T. R. & Umbarger, H. E. Alpha-isopropylmalate synthase from *Salmonella typhimurium*. Purification and properties. *J. Biol. Chem.* **244,** 2218–2225 (1969).

Supplementary Information accompanies the paper on *Nature*'s website (http://www.nature.com/nature).

Acknowledgements We thank L. Bao, B.-M. Dai, J. Yan, C. Werts, M. Picardeau and G. Baranton for suggestions and comments on our research strategy and manuscript preparation; C. Jin and G.-C. Liu of the Institute of Microbiology, Chinese Academy of Science, for help in the attempt at assaying the enzymatic activity of PafAH; Y. Liu and H.-G. Zhu for help in preparing the drawings; X. Mao and G. Cai for help in computer simulation; B.-Y. Hu and Y-X. Nie for help in bacterial culture preparation; and the members of CHGCS for support and encouragement. This work was supported by the National Natural Science Foundation of China, the Chinese National High Technology Development Program (863), the National Key Program for Basic Research (973) and the Sciences and Technology Commission of the People's Government of Shanghai Municipality. It was also supported by the Pôle Sino-Français en Sciences du Vivant et en Génomique and le Programme de Recherches Avancées Franco-Chinois PRA B00-05.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to G.-P.Z. (e-mail: gpzhao@sibs.ac.cn). The sequences have been submitted to NCBI under accession numbers GB: AE010300 and GB: AE010301 for the large (CI) and small (CII) chromosomes, respectively.

The methylated component of the *Neurospora crassa* genome

Eric U. Selker*, Nikolaos A. Tountas†‡, Sally H. Cross†‡, Brian S. Margolin*‡, Jonathan G. Murphy*, Adrian P. Bird† & Michael Freitag*

* *Department of Biology and Institute of Molecular Biology, University of Oregon, Eugene, Oregon 97403, USA*
† *Institute of Cell and Molecular Biology, University of Edinburgh, Edinburgh, EH9 3JR, UK*

Cytosine methylation is common, but not ubiquitous, in eukaryotes. Mammals[1] and the fungus *Neurospora crassa*[2,3] have about 2–3% of cytosines methylated. In mammals, methylation is almost exclusively in the under-represented CpG dinucleotides, and most CpGs are methylated[1] whereas in *Neurospora*, methylation is not preferentially in CpG dinucleotides and the bulk of the genome is unmethylated[4]. DNA methylation is essential in mammals[5] but is dispensable in *Neurospora*[3,6], making this simple eukaryote a favoured organism in which to study methylation. Recent studies indicate that DNA methylation in *Neurospora* depends on one DNA methyltransferase, DIM-2 (ref. 6), directed by a histone H3 methyltransferase, DIM-5 (ref. 7), but little is known about its cellular and evolutionary functions. As only four methylated sequences have been reported previously in *N. crassa*, we used methyl-binding-domain agarose chromatography[8] to isolate the methylated component of the genome. DNA sequence analysis shows that the methylated component of the genome consists almost exclusively of relics of transposons that were subject to repeat-induced point mutation—a genome defence system that mutates duplicated sequences[9].

To isolate the methylated component of the *N. crassa* genome, we cleaved genomic DNA with the 5-methylcytosine-sensitive restriction enzyme *Sau*3AI (recognition sequence GATC) so as to leave intact patches of methylated DNA, and then passed it over a methyl-CpG domain (MBD) column, which fractionates according to the degree of CpG methylation[8]. Bound DNA was eluted with increasing concentrations of salt, and fractions were analysed by Southern hybridizations, probing for an unmethylated sequence (*am*) and previously identified methylated regions ($\zeta - \eta$, Ψ_{63} and ribosomal DNA; Fig. 1). DNA complementary to the *am* probe eluted principally in pool four but trailed through to pool nine. In contrast, $\zeta - \eta$ sequences peaked later, in pool nine, suggesting that the MBD column successfully fractionated *Neurospora* DNA on the basis of methylation. Considering that the MBD does not bind methylated non-CpG sites[10], these findings suggest co-localization of methylated CpG and non-CpG (*Sau*3AI) sites. By this assay, the

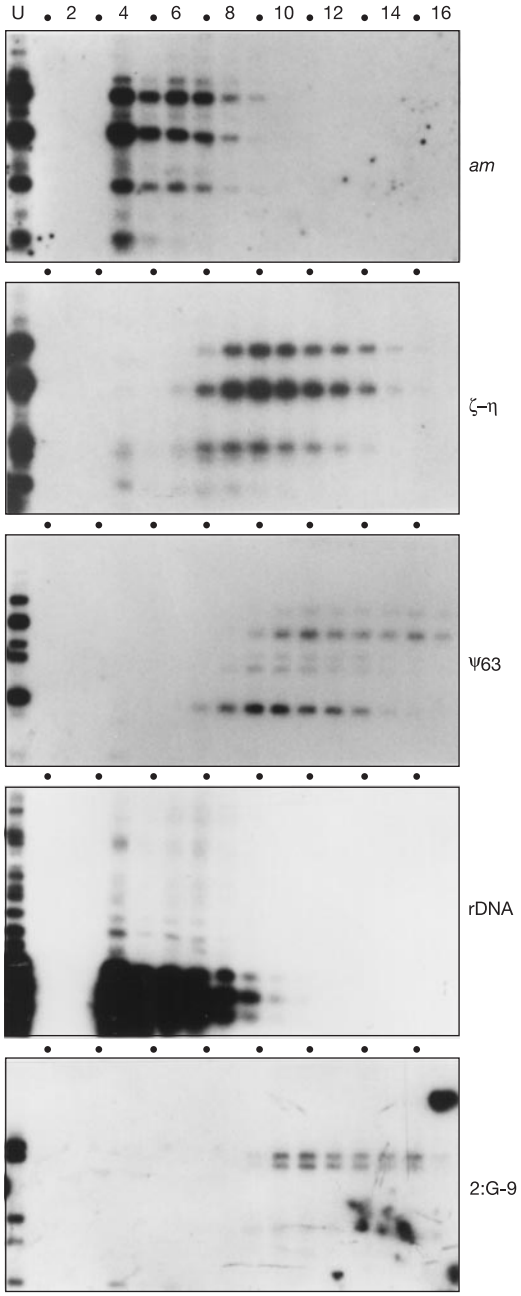

Figure 1 Fractionation of *Neurospora* DNA on a methylated-DNA-binding column. Samples (about 0.5 μg) of pooled pairs of fractions off the MBD column (1–16 represent column fractions 1–32) were fractionated by agarose gel electrophoresis, along with an unfractionated (U) sample of the *Sau*3AI-digested DNA, blotted to nylon membrane, and probed sequentially for known unmethylated (*am*) and methylated ($\zeta - \eta$, Ψ_{63} and rDNA) sequences, as well as a candidate methylated sequence from this study (2:G-9).

‡ Present addresses: Center for Cell Signaling, University of Virginia School of Medicine, Charlottesville, Virginia 22908, USA (N.A.T.); MRC Human Genetics Unit, Western General Hospital, Edinburgh, EH4 2XU, UK (S.H.C.); Department of Biochemistry and Biophysics, University of California San Francisco, California 94143, USA (B.S.M.).

© 2003 Nature Publishing Group

density of CpG methylation varied within and between sequences. For example, a significant amount of the $\zeta - \eta$ and Ψ_{63} regions of DNA were distributed between pools 4–13, presumably reflecting heterogeneous methylation (Fig. 1). Also, some Ψ_{63} fragments bound more tightly than nearly all the $\zeta - \eta$ DNA, whereas rDNA, which is known to include some methylated cytosines[2], eluted early, indicating that it has lighter methylation than the two relics of repeat-induced point mutation (RIP).

DNA from tightly bound fractions was cloned and used to probe Southern blots of the MBD-fractionated DNA and/or total genomic DNA digested with *Mbo*I and *Sau*3AI. Some clones (such as 2:G-9; Fig. 1) represented methylated chromosomal regions as expected, but many showed no evidence of methylation, indicating that the MBD column enriched for, but did not fully purify, the methylated sequences (data not shown). We therefore used differential colony hybridization with probes made from early or late column fractions to screen the library for clones that should represent methylated sequences (see Supplementary Fig. 1). DNA was isolated from about 100 of the resulting candidate methylated DNA clones and tested by probing against DNA from a wild type or from the DNA methyltransferase (DMTase) mutant *dim-2* (ref. 6) (Fig. 2). The fraction of high-molecular-mass signal in *Sau*3AI bands was highly variable, consistent with the idea that methylation levels in different chromosomal regions are variable. Results obtained with *dim-2* DNA support the conclusion that the DIM-2 DMTase is responsible for all DNA methylation in vegetative cells of *Neurospora*[6]. About 70% of the clones were found to contain sequences that are unambiguously methylated in the Oak Ridge (OR) wild-type strain (source of the DNA). We thus established a set of experimentally validated methylated DNA clones.

Previous work revealed that trichostatin A (TSA), a potent inhibitor of histone deacetylases, causes hypomethylation of RIP (repeat-induced point mutation) copies of *am*, but does not affect methylation of rDNA and the inactivated transposon *Punt*^{RIP1} in the Ψ_{63} region[11]. We used our collection of methylated DNA clones to explore the range of the TSA effect and found only minor consequences (Fig. 3). TSA-induced hypomethylation probably reflects the sensitivity of the DIM-5 HMTase to the acetylation state of lysines in histone H3 (H. Tamaru and E.U.S., unpublished data), and the differential effects may reflect different activities of histone acetyltransferases and/or histone deacetylases in various chromosomal regions.

We also used the methylated DNA clones to characterize two mutants defective in methylation (*dim-1* and *dim-3* (refs 12 and 3, respectively)) and to investigate whether the identified methylated sequences are present and methylated in a relatively distant wild-type strain, Mauriceville (M)[13]. Unlike *dim-2*, neither *dim-1* nor *dim-3* caused complete loss of methylation; however, they caused different patterns of hypomethylation (Fig. 3). Probings of M-DNA revealed many polymorphisms, both in primary sequences and methylation levels, and although many of the sequences are repetitive in both OR and M strains, the levels and patterns of repetition are quite variable. Some of the strongest hybridizing sequences of the OR strain appear absent in the M strain and little or no DNA methylation was observed in the M strain for many sequences (such as 8:A6 and 8:B9). We conclude that regions of DNA methylation are not highly conserved.

The *Neurospora* genome has relatively little (approximately 8%) repetitive DNA[14], presumably because of the operation of RIP[9]. Notably, nearly 50% of the methylated clones hybridized to repetitive DNA (data not shown). We sequenced the ends of 50 methylated clones and calculated two 'RIP indices'[15], based on the observation that RIP preferentially mutates 5′CpA: 5′TpG to generate TpA dinucleotides. For 38 of 50 methylated DNA clones analysed, both indices were diagnostic for RIP. Four clones were positive for one index only and a further three showed evidence of RIP only after detailed analysis (see below). Only 8 of 50 sequences showed ratios typical of unmutated *Neurospora* genes, but three of these were too short to calculate the indices reliably. We conclude that the methylated component of the *Neurospora* genome consists almost exclusively of sequences mutated by RIP.

To gain insight into the natural targets of RIP, we used BLAST[16] to compare the methylated sequences to sequences in GenBank and *N. crassa* genome and proteome databases (including sequences not assembled; http://www-genome.wi.mit.edu/annotation/fungi/neurospora/). Notably, 32 of 46 sequences analysed resemble transposons including *Punt*^{RIP1} (Table 1, 5:H8 (ref. 15)), *gypsy*-like retrotransposon *dab-1* (11:G2 (ref. 17)), the long interspersed nucleotide element (LINE)-like element *Tad* (11:D3 (ref. 18)), and the *copia*-like element *Tcen* (2:C10, 2:H9, 5:D4, 8:F8 (ref. 19)). We also found relics of transposons not previously known in *Neuro-*

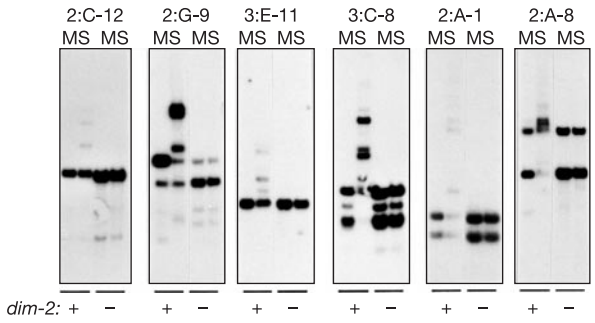

Figure 2 Verification of DNA methylation in genomic DNA regions. Clones representing the indicated potential methylated chromosomal regions were used as probes for Southern blots of *Mbo*I-digested (M) or *Sau*3AI-digested (S) genomic DNA prepared from wild-type (+) or *dim-2* (−) strains.

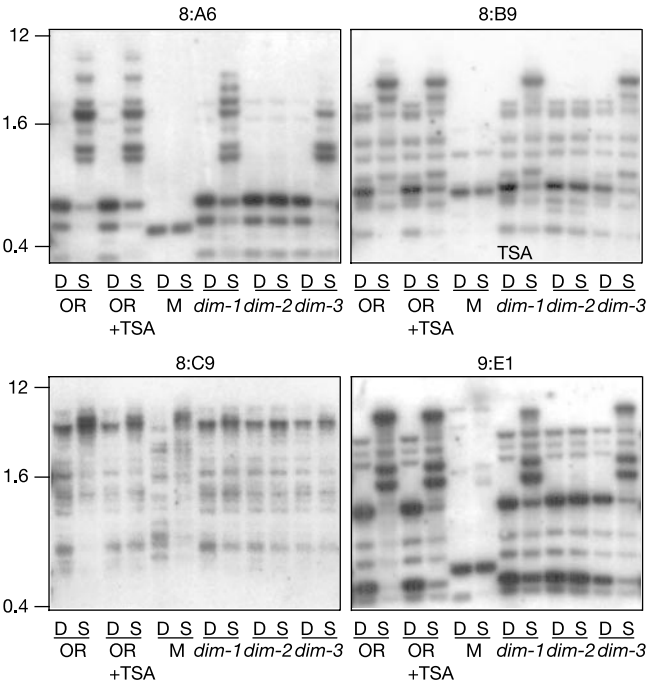

Figure 3 Comparison of DNA methylation in Oak Ridge (OR), Mauriceville (M) and *dim* strains of *N. crassa*. Approximately 1 μg of DNA from OR (FGSC 2489), M (FGSC 2225), *dim-1* (N2517), *dim-2* (N1104) or *dim-3* (N1089) were digested with *Dpn*II (D) or *Sau*3AI (S) and used for Southern hybridizations with probes from clones generated in this study. The OR strain was grown in the presence (+) or absence (−) of 1 μg ml^{-1} TSA as previously described[11]. Positions of size standards (kb) are shown on left. Additional examples are presented in Supplementary Fig. 2.

© 2003 Nature Publishing Group

spora, including the *gypsy*-like *Lolligag* (2:A11, 3:C11, Table 1) and DNA-type transposons of the *hAT*[20] and *Tc1/mariner*[21] superfamily (Table 1; see also Supplementary Fig. 3). No relics of true *mariner* transposons occur in the *Neurospora* genome sequence[22]. Several methylated fragments could not be identified unequivocally as relics of transposons but represented RIP-mutated repeated sequences (for example, 2:A1, 3:E11, 4:D12, 5:F8, 9:E1, 11:E5). We also isolated four methylated segments consisting of RIP-mutated rDNA (7:D8, 7:E5, 7:G9, 8:A10, Table 1) whose origin is uncertain. Finally, three of the four *Tcen* homologues and five additional methylated fragments from centromeric regions provide evidence of DNA methylation in centromeric regions of *Neurospora*.

What caused the methylation of the rare methylated regions with non-diagnostic RIP indices? Previous studies showed that some relics of RIP can induce methylation *de novo* and that methylation can extend far from a sequence serving as a methylation signal[23]. Thus, methylation could result from flanking signals. Alternatively, sequences with non-diagnostic RIP indices might nevertheless be products of RIP. Sequence comparisons between the methylated sequences and related sequences in the genome revealed instances of the latter possibility. Two independently isolated, heavily methylated and (A + T)-rich fragments of a sequence related to the *Minos* transposon in *Drosophila* (8:A6 and 8:B1, Table 1) provide an example. Comparison with a homologous sequence (contig 3.422) revealed 21 and 68 mutations characteristic of RIP, respectively. Similarly, methylated clone 5:A7, which was not predicted to be a product of RIP on the basis of RIP indices (Table 1), contains a RIP-mutated 259-base-pair (bp) segment of the *Dodo2* transposon relic.

To test directly for methylation signals within the 5:A7 region, we introduced 5:A7 segments into *Neurospora* and checked for *de novo* methylation. We targeted single copies of the full 736-bp fragment, the 451-bp *Dodo2* fragment or the adjacent 447-bp sequence, without extraneous (for example, vector) sequences, to the *his-3* or *am* loci and tested for induction of methylation (Supplementary Fig. 4). All fragments induced substantial methylation, indicating that flanking methylation signals are not required. The non-*Dodo2* segment of 5:A7 represents a rare example of *de novo* methylation in

Table 1 ***Neurospora crassa*** **methylated DNA clones**

Clone	Accession*	Coordinates	Linkage group†	Methylation (%)‡	Repeats§	$\frac{TpA}{ApT}$	$\frac{(CpA+TpG)}{(ApC+GpT)}$	Comments‖
2:A1	AY227784	–	–	~50	~ 2	1.41	0.51	Best match near *Cen-III*
2:A10	AABX01000180	35973–36621	III (*acr-2*)	~50	~ 3	1.28	0.75	Fo *Folyt1*; near *Cen-III*; Nc *Listless*
2:A11	AABX01000358	148–2458	VII (*wc-1*)	~50	~ 8	1.40	0.36	Mg MGLR-3; Nc *Lolligag*
2:B3	AABX01000083	8526–9112	I (*hsp30*)	~60	1	1.26	0.91	Putative RIP-mutated kinase gene (5:D1)
2:C9	AABX01000208	141257–141617	IV (*aod-1*)	~90	~10	1.14	1.05	Nc *Punt* linked to Nc *Punt*^{RIP1}
2:C10	AABX01000326	36860–37830	VI (*cax*)	~90	~30	1.47	0.42	Nc *Tcen*; near TEL VI L
2:C11	AY227785	–	–	~50	~15	1.27	0.64	Match adjacent to *Ant1*; Nc *Dodo2*
2:C12	AABX01000364	11925–13203	VII (*arg-10*)	~30	~12	1.38	0.63	Mg Occan; Nc *Punt3*
2:E8	AABX01000505	20022–21060	–	~50	~12	1.42	0.35	Mg Occan; Nc *Punt3*
2:G9	AABX01000238	15504–16798	IV (*nit-4*)	>90	~ 5	1.33	0.66	Hs *Tigger*; Nc *Nogo*
2:G12	AABX01000227	20529–21108	IV (*pyr-1*)	~80	~ 5	1.33	0.74	Hs *Tigger*; Nc *Nogo*
2:H9	AABX01000304	27426–28777	V (*ilv-2*)	Trace	~30	1.13	0.90	Nc *Tcen*; near *Cen-V*
3:C11	AL670001	10535–12971	II (*aro-1*)	~70	~ 8	1.47	0.32	Mg MGLR-3; Nc *Lolligag* (3:C8)
3:E11	AABX01000368	44939–45532	VII (*met-7*)	~40	1	1.45	0.89	Adjacent to repeats
4:D12	AY227786	–	–	~30	~10	1.23	0.45	Related to RIP-mutated sequences
5:A7	AABX01000101 AABX01000060	12231–12478 1–477	I (*met-6*)	~50	~25	0.75	1.15	An *Ant1*; links two contigs; Nc *Dodo2*
5:B8	AL355932	41108–43653	II (*gpd-1*)	>90	~ 5	1.42	0.59	Hs *Tigger*; Nc *Nogo* (8:E12)
5:C6	AABX01000442	2357–3354	–	~ 5	1	0.76	1.06	Adjacent to Mg *pth11* homologue
5:C9	AABX01000651	755–1022:1–268	–	~40	~20	1.36	0.61	Adjacent to An *Ant1*-like transposon
5:D1	AABX01000083	8059–9112	I (*hsp30*)	~40	1	1.40	0.93	Near *Cen-I* (2:B3)
5:D2	AABX01000425	2242–3293	–	~50	~30	1.08	0.94	Resembles Nc *En/Spm* flanks; *SNF2*-like
5:D4	AABX01000270	3332–4249	V (*ilv-2*)	~20	~30	1.35	0.54	Nc *Tcen*; near *Cen-V*
5:D6	AABX01000110	3898–4202	I (*ad-5*)	Trace	~25	1.08	0.95	An *Ant1*; Nc *Dodo2*
5:F8	AABX01000051	4623–6648	I (*cyt-1*)	~50	~ 9	1.47	0.50	Repeat region
5:H8	AABX01000241	26017–27007	IV (*aod-1*)	~40	~10	1.12	0.91	Identical to Nc *Punt*^{RIP1}
6:A3	AABX01000207	1412–3547	IV (*con-10*)	~50	~40	1.26	0.67	Adjacent to Aa *Quetzal*; Nc *Dodo1* (8:F5)
6:B8	AABX01000028	82503–86325	I (*vma-11*)	~ 5	~12	1.43	0.56	Mg Occan; Nc *Punt3*
6:D5	AABX01000436	61687–63770	–	30¶	~30	1.31	0.86	Within repeats near *Cen*
7:D8	AY227787	–	V L	~40	~ 150	0.61	0.89	rDNA#
7:E5	AY227788	–	V L	~10	~ 150	1.00	0.86	rDNA#
7:G9	AY227789	–	V L	~15	~ 150	0.83	1.05	rDNA#
8:A6	AABX01000266	36323–37403	V (*spe-2*)	~90	~12	0.86	1.27	Dh *Minos*; Nc *Dodo1* (8:B1)
8:A10	AY227790	–	V L	~20	~ 150	0.92	1.15	rDNA#
8:B1	AABX01000266	34422–37522	V (*spe-2*)	>90	~12	0.74	1.14	Dh *Minos*; Nc *Dodo1*(8:A6)
8:B9	AABX01000207	61975–62675	IV (*con-10*)	~70	~11	0.82	0.96	Mg Pot3; Nc *Punt2*
8:C9	AABX01000206	136650–141576	IV (*pho-5*)	~20	~25	1.39	0.52	An *Ant1*; Nc *Dodo2*
8:E12	AL355932	41108–43659	II (*gpd-1*)	~30	~ 5	1.42	0.60	Hs *Tigger*; Nc *Nogo* (5:B8)
8:F3	AL670001	20545–22136	II (*aro-1*)	>90	~ 8	1.44	0.54	*gypsy*-type relic
8:F5	AABX01000207	1409–3547	IV (*con-10*)	~40	~40	1.25	0.67	Adjacent to Aa *Quetzal*; Nc *Dodo1* (6:A3)
8:F8	AABX01000436	38371–40249		~50	~30	1.56	0.52	Rice helicase homologue; adjacent Nc *Tcen*
8:F10	AY251480	450914–454187	III (*mip-1*)	~50	1	1.47	0.37	–
8:G3	AL669988	70281–71810	II (*ro-3*)	>90	~ 9	1.21	0.73	*Nh*-like transposase; Nc *Dodo3*
9:E1	AABX01000427	71841–73691	–	~90	~ 4	1.57	0.53	RIP-mutated repeat
11:D3	AABX01000610	2064–2752	–	30¶	~60	1.39	0.30	Nc *Tad*; near *Cen*
11:E5	AABX01000204	110785–116088	IV (*ro-1*)	~60	~ 8	1.28	0.72	RIP-mutated repeat
11:G2	AABX01000297	35403–36604	V (*spe-3*)	~60	~20	1.78	0.47	Nc *dab-1*

*GenBank accession numbers for contigs (assembly 3 of WICGR; http://www-genome.wi.mit.edu/annotation/fungi/neurospora/). Accession numbers beginning with 'AL' refer to MIPS sequences (http://www.mips.biochem.mpg.de/proj/neurospora/). Fragments 2:A1, 2:C11, 4:D12, 7:D8, 7:E5, 7:G9 and 8:A10 had no perfect matches in Assembly 3.

†Linkage group and closest genetic marker (in parentheses).

‡Degree of DNA methylation estimated from Southern hybridizations (data not shown).

§The number of transposon relic repeats was estimated from database searches. Matches with a cutoff expected value of 1×10^{-4} were counted. Given the variability of RIP this may be a conservative estimate.

‖Best matches (for example, to transposons) are shown (independent overlapping clones are indicated in parentheses). Aa, *Anopheles albimanus*; An, *Aspergillus niger*; Dh, *Drosophila hydei*; Fo, *Fusarium oxysporum*; Hs, *Homo sapiens*; Mg, *Magnaporthe grisea*; Nc, *Neurospora crassa*; Nh, *Nectria haematococca*. Newly described *Neurospora* transposon relics include *Listless*, *Lolligag*, *Nogo*, *Punt2*, *Punt3* and *Dodo1* to *Dodo3*.

¶Heterogenous methylation (that is, some bands were unchanged, whereas others were >90% methylated).

#RIP-mutated copies of *Neurospora* 25S/28S rDNA, perhaps from the rDNA cluster on linkage group V L (similar or identical sequences in clones 7:E5, 7:D8, b2-1, b2-2, 7:D5, 8:A10).

© 2003 Nature Publishing Group

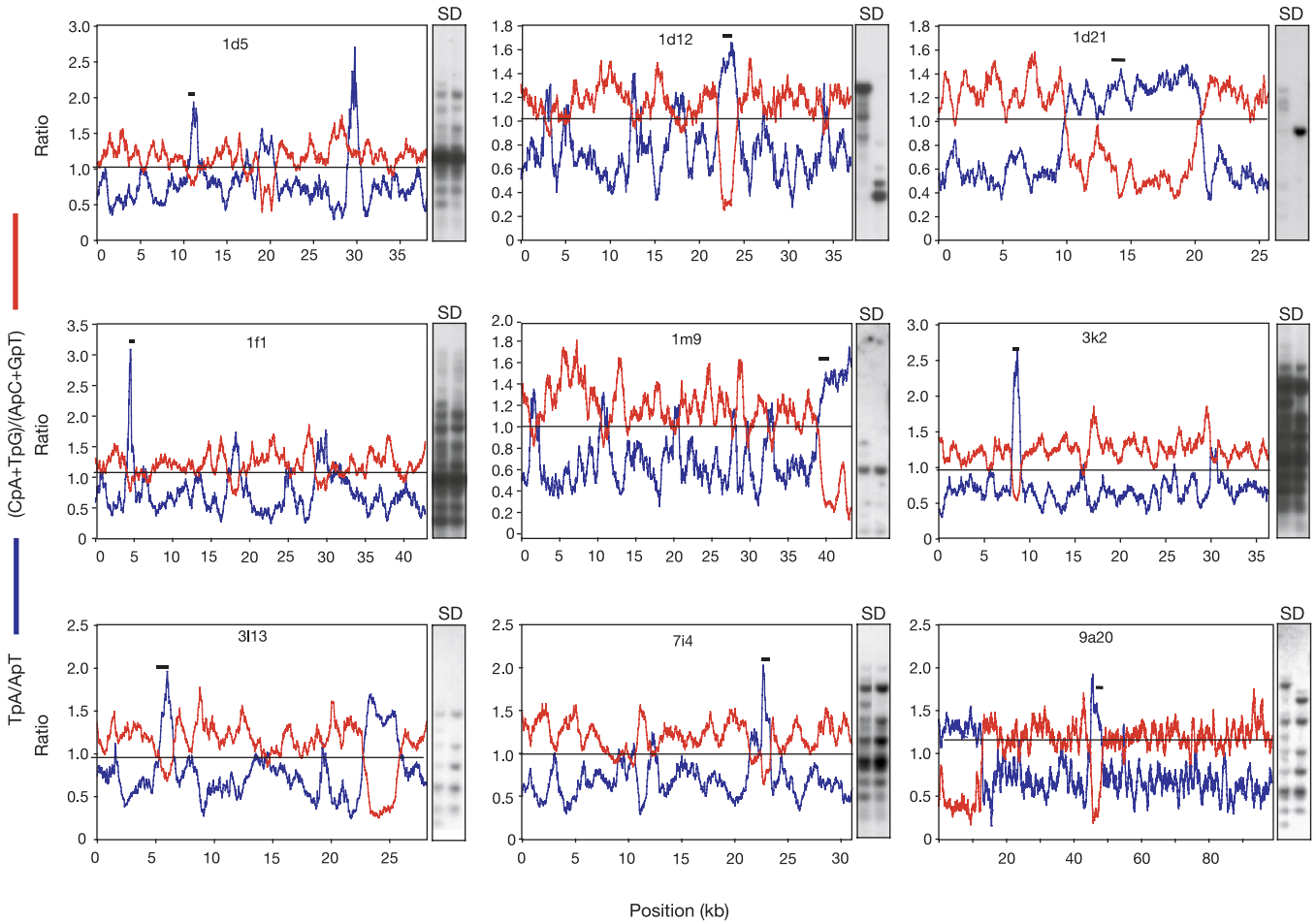

Figure 4 Methylation of sequences with extreme RIP indices. The dinucleotide composition of 3.7 Mb of non-redundant genomic DNA from linkage group V (1d5, 1d12, 1d21, 1f1, 1m9, 7i4 and 9a20) and VI (3k2 and 3l13) (http://www.mips.biochem.mpg.de/proj/neurospora/) was analysed in 200 bp windows, with 100 bp shifts in ≈40-kb segments. Sequences with TpA/ApT ratios >0.89 and/or (CpA+TpG)/(ApC+GpT) ratios <1.03 are considered relics of RIP[15]. Representative results from 9/120 contigs analysed are shown. Bars above selected TpA/ApT peaks represent polymerase chain reaction fragments (see Supplementary Table 1) generated to test for methylation of corresponding genomic DNA sequences by Southern hybridization with *Sau*3AI-(S) or *Dpn*II-(D) DNA. Resulting autoradiograms are shown to the right of the plots.

Neurospora induced by a sequence showing no evidence of RIP.

Products of RIP are frequently, but not invariably, methylated[9]. To determine whether most natural relics of RIP are methylated, we first calculated RIP indices for sequences of almost an entire *Neurospora* chromosome (Fig. 4). Most (86 of 120) of the approximately 40-kilobase (kb) segments contained at least one apparent relic of RIP. ((TpA/ApT) − (CpA + TpG)/(ApC + GpT) > 0). A total of 174 apparent relics of RIP, ranging from about 0.5 to about 13 kb (average of about 1.5 kb) were found, accounting for approximately 5% of the genomic sequences examined (roughly 0.2 out of 3.7 megabases (Mb)). To determine the fraction of putative relics of RIP that are methylated, we performed Southern hybridizations with probes generated for 20 representative sequences. Notably, 19 of 20 showed evidence of methylation (Fig. 4). Thus DNA methylation and relics of RIP are highly correlated in the *N. crassa* genome (see also ref. 22).

Several model eukaryotes, including *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Caenorhabditis elegans* and *Drosophila melanogaster* have no or little DNA methylation. At the other extreme, greater than 25% of cytosines are methylated in DNA of some higher plants[24]. Mammals and *Neurospora* fall between these extremes. Although the primary function of this modification remains controversial[25,26], it is known to prevent gene expression in animals, plants and fungi. Loss of DNA methylation has been shown to reactivate transposons in both *Neurospora*[27] and plants[28,29]. We show that in *Neurospora*, methylation is found almost exclusively in relics of transposons inactivated by RIP. Thus, all indications are that the primary, and perhaps exclusive, function of DNA methylation in *Neurospora* is to control proliferation of transposons, in conjunction with RIP. Although RIP itself may be limited to certain fungi[9], methylation is primarily associated with transposons in plants[30], raising the possibility that genome defence was the original function of this epigenetic process[25]. That normal development depends on DNA methylation in some organisms, such as mammals and *Arabidopsis*, may reflect newer roles that methylation has assumed[26]. □

Methods

Preparation of *Neurospora* DNA

A wild-type strain of *N. crassa* (74-OR23-IVA; Fungal Genetics Stock Center (FGSC) 2489) was grown in stationary liquid Vogel's minimal medium N at 31 °C until saturation (3 days). Genomic DNA was isolated[23] and purified further by phenol/chloroform extractions, CsCl-ethidium bromide equilibrium gradient centrifugation, extraction with isoamyl alcohol and ethanol precipitation. DNA (12 μg) was digested for 2 h with an excess of *Sau*3AI, extracted with phenol:chloroform:isoamyl alcohol (49:49:2), precipitated, and suspended in 40 μl 10 mM Tris (pH 8.0) in preparation for chromatography.

Enrichment of methylated DNA on MBD columns

Neurospora DNA digested with *Sau*3AI (10 μg) was fractionated using an MBD column as described[8]. DNA was ethanol-precipitated from pooled pairs of fractions (2 ml), and aliquots were analysed by gel electrophoresis and Southern hybridization as described[23].

© 2003 Nature Publishing Group

Cloning and identification of methylated sequences

Approximately 25% of fractions 17–18, 19–20, 21–24, 25–28, 29–32 and 33–36 were cloned separately into phosphatase-treated, *Bam* HI-digested pBLUESCRIBE (Stratagene), and approximately 20% of each ligation was electroporated into cells of *Escherichia coli* PKL-F′ (*recA*, *lac*, *mcrA*, *mcrB1*, *hsdR2*, *supE44*, *galK2*, *galT22*, *metB1* [*F′ proAB*, *lacI9Z ΔM15*, Tn*10*]). Representative white colonies on 5-bromo-4-chloro-3-indoyl-β-D-galactoside (X-Gal) plates were picked (approximately 400 from fractions 21–24; 400 from fractions 25–28; and 100 from each of the other pools) for analysis. Clones were grown in 96-well microtiter dishes with 0.2 ml LB medium supplemented with 10% glycerol and 50 $\mu g\,ml^{-1}$ ampicillin, and then transferred to −70 °C for storage. Clones representing methylated DNA were tentatively identified by colony hybridization by comparing signals resulting from probing replica blots with labelled DNA from column fractions 7–8 (lane 4 in Fig. 1) and 17–22 (lanes 9–11 in Fig. 1). Purified DNA of promising clones was used to probe *Sau*3AI- or *Mbo* I-digested *Neurospora* DNA by Southern hybridization[23]. Inserts were sequenced at the University of Oregon Biotechnology facility.

Computer analyses

To investigate the occurrence of relics of RIP in the genome, we used the Windows program of the GCG Wisconsin Package (Accelrys) to scan 3.7 Mb of linkage group V and VI sequence from the German *Neurospora* genome sequencing consortium (http://www.mips.biochem.mpg.de/proj/neurospora/) for hallmarks of RIP, and plotted the data using Excel software (Microsoft). The *Neurospora* genome (http://www-genome.wi.mit.edu/annotation/fungi/neurospora/) and GenBank sequences (http://www.ncbi.nlm.nih.gov/BLAST/) were searched for matches of 46 sequenced methylated regions (see Table 1) using BLASTN, BLASTX or TBLASTN[16]. Alignments were performed using nucleic acid and protein sequences at the Biology Workbench website (http://workbench.sdsc.edu/) with ClustalW.

Received 24 December 2002; accepted 14 March 2003; doi:10.1038/nature01564.

1. Bird, A. P. CpG-rich islands and the function of DNA methylation. *Nature* **321,** 209–213 (1986).
2. Russell, P. J., Rodland, K. D., Cutler, J. E., Rachlin, E. M. & McCloskey, J. A. in *Molecular Genetics of Filamentous Fungi* (ed. Timberlake, W.) 321–332 (Alan Liss, New York, 1985).
3. Foss, H. M., Roberts, C. J., Claeys, K. M. & Selker, E. U. Abnormal chromosome behavior in *Neurospora* mutants defective in DNA methylation. *Science* **262,** 1737–1741 (1993).
4. Selker, E. U., Fritz, D. Y. & Singer, M. J. Dense non-symmetrical DNA methylation resulting from repeat-induced point mutation (RIP) in *Neurospora*. *Science* **262,** 1724–1728 (1993).
5. Li, E., Bestor, T. H. & Jaenisch, R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69,** 915–926 (1992).
6. Kouzminova, E. A. & Selker, E. U. *Dim-2* encodes a DNA-methyltransferase responsible for all known cytosine methylation in *Neurospora*. *EMBO J.* **20,** 4309–4323 (2001).
7. Tamaru, H. & Selker, E. U. A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. *Nature* **414,** 277–283 (2001).
8. Cross, S. H., Charlton, J. A., Nan, X. & Bird, A. P. Purification of CpG islands using a methylated DNA binding column. *Nature Genet.* **6,** 236–244 (1994).
9. Selker, E. U. Repeat-induced gene silencing in fungi. *Adv. Genet.* **46,** 439–450 (2002).
10. Meehan, R. R., Lewis, J. D. & Bird, A. P. Characterization of MeCP2, a vertebrate DNA binding protein that binds methylated DNA. *Nucleic Acids Res.* **20,** 5085–5092 (1992).
11. Selker, E. U. Trichostatin A causes selective loss of DNA methylation in *Neurospora*. *Proc. Natl Acad. Sci. USA* **95,** 9430–9435 (1998).
12. Foss, H. M., Roberts, C. J. & Selker, E. U. Reduced levels and altered patterns of DNA methylation caused by mutations in *Neurospora crassa*. *Mol. Gen. Genet.* **259,** 60–71 (1998).
13. Metzenberg, R. L., Stevens, J. N., Selker, E. U. & Morzycka-Wroblewska, E. Identification and chromosomal distribution of 5S rRNA genes in *Neurospora crassa*. *Proc. Natl Acad. Sci. USA* **82,** 2067–2071 (1985).
14. Krumlauf, R. & Marzluf, G. A. Genome organization and characterization of the repetitive and inverted repeat DNA sequences in *Neurospora crassa*. *J. Biol. Chem.* **255,** 1138–1145 (1980).
15. Margolin, B. S. *et al.* A methylated Neurospora 5S rRNA pseudogene contains a transposable element inactivated by RIP. *Genetics* **149,** 1787–1797 (1998).
16. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25,** 3389–3402 (1997).
17. Bibbins, M., Cummings, N. J. & Connerton, I. F. DAB1: a degenerate retrotransposon-like element from *Neurospora crassa*. *Mol. Gen. Genet.* **258,** 431–436 (1998).
18. Cambareri, E. B., Helber, J. & Kinsey, J. A. Tad1-1, an active LINE-like element of *Neurospora crassa*. *Mol. Gen. Genet.* **242,** 658–665 (1994).
19. Cambareri, E. B., Aisner, R. & Carbon, J. Structure of the chromosome VII centromere region in *Neurospora crassa*: degenerate transposons and simple repeats. *Mol. Cell Biol.* **18,** 5465–5477 (1998).
20. Kempken, F. & Windhofer, F. The hAT family: a versatile transposon group common to plants, fungi, animals, and man. *Chromosoma* **110,** 1–9 (2001).
21. Plasterk, R. H., Izsvak, Z. & Ivics, Z. Resident aliens: the Tc1/mariner superfamily of transposable elements. *Trends Genet.* **15,** 326–332 (1999).
22. Galagan, J. E. *et al.* The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* **422,** 859–868 (2003).
23. Miao, V. P., Freitag, M. & Selker, E. U. Short TpA-rich segments of the zeta-eta region induce DNA methylation in *Neurospora crassa*. *J. Mol. Biol.* **300,** 249–273 (2000).
24. Paszkowski, J. & Whitham, S. A. Gene silencing and DNA methylation processes. *Curr. Opin. Plant Biol.* **4,** 123–129 (2001).
25. Yoder, J. A., Walsh, C. P. & Bestor, T. H. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* **13,** 335–340 (1997).
26. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16,** 6–21 (2002).
27. Zhou, Y., Cambareri, E. B. & Kinsey, J. A. DNA methylation inhibits expression and transposition of the *Neurospora* Tad retrotransposon. *Mol. Gen. Genet.* **265,** 748–754 (2001).
28. Miura, A. *et al.* Mobilization of transposons by a mutation abolishing full DNA methylation in *Arabidopsis*. *Nature* **411,** 212–214 (2001).
29. Singer, T., Yordan, C. & Martienssen, R. A. Robertson's Mutator transposons in *A. thaliana* are regulated by the chromatin-remodeling gene Decrease in DNA Methylation (DDM1). *Genes Dev.* **15,** 591–602 (2001).
30. Tompa, R. *et al.* Genome-wide profiling of DNA methylation reveals transposon targets of CHROMOMETHYLASE3. *Curr. Biol.* **12,** 65–68 (2002).

Supplementary Information accompanies the paper on *Nature*'s website (http://www.nature.com/nature).

Acknowledgements We thank F. K. Selker and C. B. Matsen for help with computer analyses; T. Wolfe for technical assistance; G. Mannhaupt for providing some accession numbers; and J. Galagan for comments on the manuscript. E.U.S. acknowledges the hospitality of the Bird laboratory when he initiated this project while on sabbatical, and is thankful to D. Macleod, R. Meehan and F. Antequera for their advice. This work was supported by a US Public Health Service grant to E.U.S. from the National Institutes of Health, and a Senior International Fellowship from the Fogarty International Center of the National Institutes of Health.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to E.U.S (e-mail: selker@molbio.uoregon.edu). Sequences are deposited in GenBank. Accession numbers for sequences isolated with the MBD column are listed in Table 1 and accession numbers for the MIPS sequences (Fig. 4) are: AL670542 (1d5), AL451013 (1d12), AL807371 (first 6 kb of 1d21), BX294013 (last 23.55 kb of 1d21), AL513410 (1f1), AL513411 (1m9), BX295538 (3k2), BX295540 (3l13), AL807369 (7i4), BX295539 (9a20).

Cell fusion is the principal source of bone-marrow-derived hepatocytes

Xin Wang*, Holger Willenbring*, Yassmine Akkari*, Yumi Torimaru*, Mark Foster*, Muhsen Al-Dhalimy*, Eric Lagasse†, Milton Finegold‡, Susan Olson* & Markus Grompe*

** Department of Molecular and Medical Genetics, Oregon Health & Science University, Portland, Oregon 97239, USA*
† Stem Cells Inc., Palo Alto, California 94304, USA
‡ Department of Pathology, Texas Children's Hospital, Houston, Texas 77030, USA

Evidence suggests that haematopoietic stem cells might have unexpected developmental plasticity, highlighting therapeutic potential. For example, bone-marrow-derived hepatocytes can repopulate the liver of mice with fumarylacetoacetate hydrolase deficiency and correct their liver disease[1]. To determine the underlying mechanism in this murine model, we performed serial transplantation of bone-marrow-derived hepatocytes. Here we show by Southern blot analysis that the repopulating hepatocytes in the liver were heterozygous for alleles unique to the donor marrow, in contrast to the original homozygous donor cells. Furthermore, cytogenetic analysis of hepatocytes transplanted from female donor mice into male recipients demonstrated 80,XXXY (diploid to diploid fusion) and 120,XXXXYY (diploid to tetraploid fusion) karyotypes, indicative of fusion between donor and host cells. We conclude that hepatocytes derived form bone marrow arise from cell fusion and not by differentiation of haematopoietic stem cells.

Recent reports have highlighted the broad developmental potential of bone-marrow-derived stem cells, a phenomenon termed 'stem cell plasticity'. Bone marrow contains haematopoietic stem cells (HSCs)[2] as well as mesenchymal stem cells[3] and multipotent adult progenitor cells[4]. HSCs produce not only all of the blood lineages, but also skeletal muscle[5], neurons[6], cardiac muscle[7], and pulmonary[8] and liver epithelium[9–11]. It has been shown that transplantation of HSCs can act as a substitute for hepatocyte transplantation in a murine model of tyrosinaemia, and HSC transplantation can correct this metabolic liver disease[1]. Although this indicates that

© 2003 Nature Publishing Group