DNA sequence homology induces cytosine-to-thymine mutation by a heterochromatin-related pathway in Neurospora

Eugene Gladyshev & Nancy Kleckner

Most eukaryotic genomes contain substantial amounts of repetitive DNA organized in the form of constitutive heterochromatin and associated with repressive epigenetic modifications, such as H3K9me3 and C5 cytosine methylation (5mC). In the fungus Neurospora crassa, H3K9me3 and 5mC are catalyzed, respectively, by a conserved SUV39 histone methyltransferase, DIM-5, and a DNMT1-like cytosine methyltransferase, DIM-2. Here we show that DIM-2 can also mediate repeat-induced point mutation (RIP) of repetitive DNA in N. crassa. We further show that DIM-2-dependent RIP requires DIM-5, HP1, and other known heterochromatin factors, implying a role for a repeat-induced heterochromatin-related process. Our previous findings suggest that the mechanism of repeat recognition for RIP involves direct interactions between homologous double-stranded DNA (dsDNA) segments. We thus now propose that, in somatic cells, homologous dsDNA–dsDNA interactions between a small number of repeat copies can nucleate a transient heterochromatic state, which, on longer repeat arrays, may lead to the formation of constitutive heterochromatin.

A substantial proportion of nearly every eukaryotic genome is occupied by repetitive DNA that is organized in the form of constitutive heterochromatin and associated with repressive epigenetic modifications, such as histone H3 lysine 9 di- or trimethylation (H3K9me2 and H3K9me3) and 5mC1–6. The process of de novo heterochromatin assembly has been characterized particularly well in the fungus N. crassa, where H3K9me3 and 5mC are established, respectively, by a conserved SUV39 histone methyltransferase, DIM-5, and a DNMT1-like cytosine methyltransferase, DIM-2 (ref. 7). In N. crassa, while the bulk of constitutive heterochromatin occurs at dedicated AT-rich positions in the genome. Thus, RIP represents a process that is uniquely sensitive to DNA homology. Notably, in some filamentous fungi, an analogous process leads to cytosine methylation rather than mutation17. Previously, we showed that RIP does not involve the canonical homology-recognition pathway mediated by MEI-3 (the only RecA homolog in N. crassa) and that RIP can match weakly similar DNA sequences as long as those sequences share a series of interspersed homologous base-pair triplets spaced at intervals of 11 or 12 bp18–20. These and other results are consistent with a mechanism of recombination-independent homology recognition that involves interactions between co-aligned dsDNA molecules19. Our current findings raise the possibility that such homologous interactions may occur not only in premeiotic but also in somatic cells, where they may nucleate heterochromatin formation on repetitive DNA.

**RESULTS**

RIP can occur in the absence of RID

In our earlier work18, we developed a sensitive RIP tester construct comprising one endogenous copy and one closely positioned (ectopic) copy of an arbitrarily chosen 802-bp region of the Neurospora genome (Fig. 1a and Supplementary Fig. 1). This construct induces strong mutation in a wild-type genetic background. Thus, specifically, a sample of 24 progeny spores was found to contain 3,163 mutations in the endogenous (‘left’) repeat copy, 3,153 mutations in the ectopic (‘right’) repeat copy and 524 mutations in the endogenous 600-bp segment of the linker region (Fig. 1b, rid+/+; dim-2+/+; data replotted from ref. 18). We now confirm that all of the above mutations, including those of the linker region, are induced by the presence of DNA homology: if the ectopic repeat copy is specifically omitted, no RIP activity can be detected (Fig. 1c).
In *N. crassa*, RIP has long been known to require a putative C5 cytosine methyltransferase, RID (RIP defective)\(^{25,26}\). However, the genomes of several fungal species contain clear signatures of RIP-like mutation but do not encode any apparent homologs of RID, hinting at the possibility that RIP could be mediated by other factors\(^{22,23}\). In accordance with this idea, we now find that our 802-bp tester construct can induce substantial mutation in the absence of RID (Fig. 1b, rid\(^{Δ/Δ}\); dim-2\(^{Δ/Δ}\)). Intriguingly, while the expected (mean) number of mutations in the repeated regions was decreased by nearly two orders of magnitude in the *rid*\(^{Δ/Δ}\) background, the linker region was still mutated at essentially the wild-type level (Supplementary Fig. 2a). The findings above (Fig. 1c) imply that all of these RID-independent mutations are induced by DNA homology.

**RID-independent RIP is mediated by DIM-2**

In addition to RID, *N. crassa* encodes another cytosine methyltransferase, DIM-2 (defective in methylation-2)\(^{22,23}\), which catalyzes all cytosine methylation of the 802-bp construct (Fig. 1b, rid\(^{Δ/Δ}\); dim-2\(^{Δ/Δ}\)). Furthermore, when DIM-2 was absent but RID was present, the pattern of the effects was the reciprocal of that observed when DIM-2 was present but RID was absent: the number of mutations declined strongly (by a factor of 7.6) in the linker region and only moderately (by a factor of 1.3) in the repeated regions (Fig. 1b, rid\(^{Δ/+}\); dim-2\(^{Δ/Δ}\)). Interestingly, both *rid* and *dim-2* appeared to be haploinsufficient: when the deletion allele was present in combination with a corresponding wild-type allele, the deletion allele for each gene decreased the number of mutations in the corresponding affected region(s) by a factor of three or more (Supplementary Fig. 2a).

The above findings show that (i) RID and DIM-2 can each individually mediate RIP; (ii) RID and DIM-2 together account for all RIP; (iii) in the context of the 802-bp tester construct, RID-mediated mutation predominantly targets the repeated sequences, whereas DIM-2-mediated mutation predominantly targets the single-copy linker region; and (iv) the effects of RID and DIM-2 are additive. Taken together, these results suggest that the RIP processes mediated by RID and DIM-2, even though both triggered by the same DNA repeats, are nonetheless functionally distinct and, to a first approximation, independent of one another.

To confirm the distinct and complementary nature of the RID- and DIM-2-mediated pathways of RIP, we defined pairwise correlations, on a per-spore basis, between the numbers of mutations occurring in different segments of the 802-bp tester construct (Supplementary Fig. 3). In situations where RID and DIM-2 activities were both strong, the total numbers of mutations in the left and right repeat copies of each individual spore clone were strongly correlated. That is, if one repeat copy exhibited a certain number of mutations, so did the other copy, on a per-spore basis. This pattern is expected if the two repeat copies are mutated by the same process. In contrast, the number of mutations in the linker region correlated less strongly with the number of mutations in either repeat copy, as expected if linker mutations and repeat mutations are mediated by two separate processes (DIM-2 and RID, respectively).

**DIM-2-mediated RIP requires DIM-5**

In vegetative cells of *N. crassa*, DIM-2 is recruited to DNA by heterochromatin protein 1 (HP1), which recognizes H3K9me3 established by the SUV39 histone methyltransferase DIM-5 (Fig. 2a)\(^{25,26}\). The above observations raised the possibility that DIM-2-mediated mutation might also require these same factors. As one test of this possibility, we investigated the functional relevance of DIM-5 for RIP. Because female sexual development of *N. crassa* *dim-5* strains is impaired in the absence of SET-7 (the H3K37 methyltransferase encoded by the *set-7* gene)\(^{25,26}\) and because RIP occurs normally in the absence of SET-7 (Supplementary Fig. 4), the role of DIM-5 in RIP was tested using *set-7* female strains (Supplementary Tables 1 and 2). We found that, when DIM-5 and RID were both absent, no mutation could be detected, even if DIM-2 was still available at the wild-type level (Fig. 2b). DIM-2-mediated mutation could readily be restored in the presence of a single functional *dim-5* allele (Fig. 2b).

In the heterozygous *dim-5*\(^{Δ/+}\) cross analyzed above (Fig. 2b), the male parent provided the wild-type *dim-5* allele together with the 802-bp construct, whereas the female parent provided the null *dim-5* allele. In such a configuration (where the *dim-5* allele and the tester construct were supplied ‘in cis’; Supplementary Fig. 1c), DIM-5 could, in principle, have acted on the repeat construct in vegetative
cells, before fertilization and the onset of RIP. To investigate whether the role of DIM-5 in RIP could be fulfilled after fertilization, we assayed the RIP-promoting activity of DIM-5 in a cross between a repeat-carrying maternal dim-5Δ strain and a repeat-lacking paternal strain that supplied the dim-5Δ allele (that is, the dim-5Δ allele and the tester construct were provided ‘in trans’). We found that the cis and trans configurations resulted in comparable levels of RIP, regardless of whether the functional dim-5Δ allele was present in the endogenous or the ectopic location (Fig. 2c). These results imply that DIM-5 can act after fertilization. Furthermore, nearly all mutation observed in the trans configuration could be abolished by a single amino acid change, p.Tyr283Phe, in DIM-5 that eliminated a catalytically important tyrosine residue27 (Fig. 2c). Moreover, the fact that the mutant DIM-5ΔY283Phe protein still exhibited very weak RIP-promoting activity in vivo corresponds to the fact that it also retains some weak catalytic activity in vitro27. Thus, DIM-5 seems to exert its role in RIP through its H3K9 methyltransferase activity.

We also examined the effect of decreasing DIM-5 levels in the presence of RID (Fig. 2d). The complete lack of DIM-5 conferred a strong and specific phenotype in the linker region (a 24.3-fold decrease in mutation) while having a much more modest impact on the repeated sequences (a 1.4-fold decrease, similar to the 1.3-fold decrease observed in the repeated sequences in the absence of DIM-2; Supplementary Fig. 2a,b). These findings provide further evidence that RID-mediated RIP and DIM-2-mediated RIP represent two largely independent processes that nevertheless respond to the same signal of DNA homology.

**DIM-2-mediated RIP also requires HP1, CUL4 and DDB1**

The finding that RIP could be mediated by the DIM-5–DIM-2 pathway raised the possibility that *Neurospora* HP1 (encoded by the *hpo* gene) might also be involved in RIP. In an attempt to test this idea, we found that lack of HP1 precluded meiotic sporogenesis, even in the set-7Δ background that normally permits sexual development in the absence of DIM-5 (refs. 25,26). This phenotype likely reflects a broader requirement for HP1 during the sexual phase in *N. crassa*. However, we were able to find a combination of two *hpo*Δ strains that produced some limited amount of late-arising spores. One of these strains (C135.3; used as a female parent) lacked the type II topoisomerase-like protein SPO11, which catalyzes the formation of double-strand DNA breaks in meiosis and is fully dispensable for RIP18–20. Our analysis showed that the apparent level of RIP was strongly decreased in the *hpo*ΔΔ background (Fig. 2e). However, (i) substantial mutation of the repeated sequences still occurred in the absence of HP1 and (ii) the relative levels of RIP in the linker region were diminished by the same amount in both *hpo*ΔΔ and *dim-2*ΔΔ

---

**Figure 2** DIM-2-mediated RIP requires the SUV39 histone methyltransferase DIM-5. (a) A cartoon representation of the canonical heterochromatin pathway in *N. crassa*15. (b) RIP mutation profiles of the 802-bp construct. Crosses X_8(60) (top) and X_9(60) (bottom). (c) RIP mutation profiles of the 802-bp construct. Crosses (left to right) X_1(24), X_4(24) and X_17(24) and X_18(48). (e) RIP mutation profiles of the 802-bp construct. Crosses (left to right) X_18(24), X_19(24) and X_20(24). (g) A fraction of linker mutations reports the relative activity of DIM-2-mediated RIP. Each fraction value corresponds to the number of mutations in the 600-bp segment of the linker region normalized by the total number of mutations found in this 600-bp segment and in the repeated sequences. Analyzed crosses (left to right): X_1, X_4, X_17, X_18, X_19 and X_20. The difference between any two fraction values was evaluated for significance using the χ² homogeneity test (*P* value indicated below the line) and Fisher’s exact test (*P* value indicated below the line). The complete lack of DIM-5 conferred a substantial mutation of the repeated sequences in the absence of DIM-2; however, we were able to find a combination of two *hpo*Δ strains that produced some limited amount of late-arising spores. One of these strains (C135.3; used as a female parent) lacked the type II topoisomerase-like protein SPO11, which catalyzes the formation of double-strand DNA breaks in meiosis and is fully dispensable for RIP18–20. Our analysis showed that the apparent level of RIP was strongly decreased in the *hpo*ΔΔ background (Fig. 2e). However, (i) substantial mutation of the repeated sequences still occurred in the absence of HP1 and (ii) the relative levels of RIP in the linker region were diminished by the same amount in both *hpo*ΔΔ and *dim-2*ΔΔ.
DIM-2-mediated RIP responds to weak interspersed homology

The discovery of a heterochromatin-related pathway of RIP now raises the issue of the precise nature of the homology requirements for this process. Our previous results provided no indication that the canonical recombination-mediated mechanism of homology recognition (mediated by MEI-3) had a role in either the RID- or DIM-2-mediated pathway. We have also previously shown that, in the wild-type situation, when both pathways are active, RIP can respond to the presence of weak, interspersed homology as long as it comprises short islands of homology (23 bp) spaced at regular intervals of 11 (or 12) bp along a pair of participating DNA segments. This and other results permitted a model in which homology recognition for RIP involves direct interactions between co-aligned DNA duplexes.

One particular series of experiments that led to this conclusion used repeat constructs in which 200-bp regions of interspersed homology were adjacent to a 220-bp region of perfect homology (Fig. 3a,b). These direct ‘200 + 220’ repeats were linked by a single-copy region of 537 bp. The 220-bp region of perfect homology was incorporated to provide a stable, permanent point of interaction, thereby facilitating the detection of weak effects induced by interspersed homologies.

In this context, different homology patterns could be created by manipulating only the right 200-bp segment while leaving the rest of the construct unchanged. We have now used this same repeat system to specifically address the homology requirements for DIM-2-mediated RIP. We first reanalyzed our published data (obtained in the rid(0); dim-2(0) background) with respect to mutation of the 450-bp segment of the linker region expected to differentially report the effects of the DIM-5–DIM-2-mediated process (Fig. 3a). The analyzed instances of interspersed homology involved homologous units of 3 or 4 bp spaced at regular intervals of 11 bp. We found that the presence of these weak, interspersed homologies corresponded to a significant increase in RIP mutation throughout the 450-bp linker segment (Fig. 3c and Supplementary Fig. 2d). We then asked whether these same homology patterns could promote RIP in the absence of RID, when only the DIM-5–DIM-2 pathway was active. Here we found again that RIP in the linker segment was increased significantly in the presence of the assayed interspersed homologies (Fig. 3c and Supplementary Fig. 2d).

Most of the analyses examined DIM-2-mediated RIP as triggered by closely positioned repeats. It was of interest to determine whether the heterochromatin-related pathway of RIP could also mutate DNA repeats at widely separated genomic positions. To address this question, we designed a sensitive genetic system that can detect RIP between a single pair of homologous sequences located 2.7 Mb apart.
on the same chromosome (Fig. 5a; also see ref. 20). The csr-1 gene encodes a cyclophilin protein that has a high affinity for cyclosporin A, and the presence of a single active copy of csr-1 is sufficient to render a Neurospora strain sensitive to cyclosporin28. Thus, if two csr-1 copies are present, one of which is active and the other of which is not, RIP mutation of the active copy (induced by the ectopic inactive copy) will generate cross progeny resistant to cyclosporin (Fig. 5b).

We found that, in the absence of the ectopic copy, no cyclosporin-resistant progeny could be detected among nearly 5 × 10^6 spores produced by a cross between two wild-type strains, implying that the frequency of spontaneous mutation of the csr-1 gene is extremely low. In contrast, in the presence of the ectopic copy, a large proportion of cross progeny became resistant to cyclosporin (Fig. 5c). This proportion decreased by nearly four orders of magnitude in the absence of cross progeny became resistant to cyclosporin (Fig. 5b). The above results demonstrate that relatively short DNA repeats can trigger RIP by the heterochromatin-related pathway. Our previous studies18–20, extended above, suggest that recognition of repeats for this pathway is mediated by direct homologous dsDNA–dsDNA interactions. Taken together, these findings raise the possibility that such dsDNA–dsDNA interactions may be involved in the formation of heterochromatin in vegetative cells of N. crassa but result in cytosine methylation instead of mutation.

**Supplementary Fig. 5.** An additional interesting difference between the two mutational pathways is the even stronger preference for CpA sites associated with DIM-2-mediated RIP (Fig. 5e). Taken together, these results demonstrate the capacity of the heterochromatin-related pathway to mediate mutation of DNA repeats that are separated by large genomic distances.

**Short repeat arrays induce transient somatic 5mC**

The above results demonstrate that relatively short DNA repeats can trigger RIP by the heterochromatin-related pathway. Our previous studies18–20, extended above, suggest that recognition of repeats for this pathway is mediated by direct homologous dsDNA–dsDNA interactions. Taken together, these findings raise the possibility that such dsDNA–dsDNA interactions may be involved in the formation of heterochromatin in vegetative cells of N. crassa but result in cytosine methylation instead of mutation.

The most straightforward test of this idea involves asking whether, in vegetative cells, the DIM-5–DIM-2 pathway can mediate 5mC at short repeat arrays that also trigger DIM-2-mediated RIP (Fig. 5e). The most straightforward test of this idea involves asking whether, in vegetative cells, the DIM-5–DIM-2 pathway can mediate 5mC at short repeat arrays that also trigger DIM-2-mediated RIP (Fig. 5e). The most straightforward test of this idea involves asking whether, in vegetative cells, the DIM-5–DIM-2 pathway can mediate 5mC at short repeat arrays that also trigger DIM-2-mediated RIP (Fig. 5e).
of the 4x array, in a single-copy genomic region that shows strong DIM-5–DIM-2-dependent RIP (Fig. 4a,b, site A).

We analyzed 5mC at BstU site A under several conditions, using two additional sites (B and C) that were not expected to undergo heterochromatinization as normalization controls. We found that cleavage of site A was partially inhibited when the 4x array was present, as indicated by enrichment of the corresponding PCR product (Fig. 4c; raw data are provided in Supplementary Fig. 6). Notably, this enrichment effect did not occur in the absence of DIM-5 or if a copy of the csr-1 gene was integrated instead of the 4x array. The level of the diagnostic PCR signal induced by the 4x array remained low and corresponded to a 5mC frequency of ~2–3%. We conclude that our assay is detecting rare (and, thus, likely transient) events of the repeat-induced heterochromatinization as normalization controls. We found that cleavage of site A was partially inhibited when the 4x array was present, as indicated by enrichment of the corresponding PCR product (Fig. 4c; raw data are provided in Supplementary Fig. 6). Notably, this enrichment effect did not occur in the absence of DIM-5 or if a copy of the csr-1 gene was integrated instead of the 4x array. The level of the diagnostic PCR signal induced by the 4x array remained low and corresponded to a 5mC frequency of ~2–3%. We conclude that our assay is detecting rare (and, thus, likely transient) events of the repeat-induced heterochromatinization as normali...
DIM-2 can also methylate non-repetitive parts of the genome. If premeiotic phase. It can be noted, however, that in vegetative cells, DIM-2 methylates cytosines in all dinucleotide contexts, whereas cytosine methylation, without obvious deamination. Interestingly, DIM-2. In vegetative cells of Neurospora crassa, RIP can be mediated by the canonical cytosine methyltransferase. This is consistent with the idea that RIP is initiated by DNA methylation events, as supported by the fact that 5mC (red bars). Such heterochromatin nucleation sites are intrinsically unstable. However, on longer repeat arrays, multiple concomitant nucleation events may be stabilized to yield constitutive heterochromatin. During the premeiotic phenomenon of RIP, the same transient localized interactions result in permanent C-to-T mutations, instead of 5mC. These mutations accumulate during several cell cycles over a period of several days to give a high level of RIP in analyzed spores.

in another fungus, Ascobolus immersus. An alternative proposal was also put forward by which RID alone could mediate mutation, via modulation of its catalytic activity. Our current findings show that RIP can be mediated by the canonical cytosine methyltransferase, DIM-2. In vegetative cells of N. crassa, DIM-2 catalyzes all known cytosine methylation, without obvious deamination. Interestingly, DIM-2 methylates cytosines in all dinucleotide contexts, whereas DIM-2-mediated mutation exhibits a very strong preference for CpA dinucleotides. These findings are fully compatible with a two-step mechanism. By this hypothesis, repeat recognition would trigger 5mC (by either DIM-2 or RID) that would be converted to DNA deaminase that is active specifically during the premeiotic phase. It can be noted, however, that in vegetative cells DIM-2 can also methylate non-repetitive parts of the genome. If that same process takes place in premeiotic nuclei, then a two-step mechanism would further require the deamination step to be dependent on the presence of DNA homology as well. A one-step mechanism involving premeiotic modulation of the catalytic activity of DIM-2 would not have such a requirement.

Finally, we note that the characteristic spread of DIM-5–DIM-2-mediated genetic and epigenetic modifications from repeats into neighboring non-repetitive genomic regions implicates this homology-directed pathway in the transcriptional silencing and accelerated evolution of pathogenic genes that are often found in clusters with repetitive elements in the genomes of many filamentous fungi.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank D. Zickler for critically reading the manuscript. This work was supported by grant GM044794 from the US National Institutes of Health to N.K. and research fellowships from the Helen Hay Whitney Foundation, the Howard Hughes Medical Institute and the Charles A. King Trust to E.G.

AUTHOR CONTRIBUTIONS

E.G. designed, performed and analyzed the data from all experiments. E.G. and N.K. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html. Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Almouzni, G. & Probst, A.V. Heterochromatin maintenance and establishment: lessons from the mouse pericentromere. Nucleus 2, 332–338 (2011).
2. Becker, J.S., Nicetal, D. & Zaret, K.S. H3K9me3-dependent heterochromatin: barrier to cell fate changes. Trends Genet. 32, 29–41 (2016).
3. Martienssen, R. & Moazed, D. RNAi and heterochromatin assembly. Cold Spring Harb. Perspect. Biol. 7, a019323 (2015).
4. Padeken, J., Zeller, P. & Gasser, S.M. Repeat DNA in genome organization and stability. Curr. Opin. Genet. Dev. 21, 12–19 (2015).
5. Plio, M., Meštrović, N. & Mravinac, B. Centromere identity from the DNA point of view. Chromosoma 123, 313–325 (2014).
6. Saksov, J., Simboeck, E. & Dejardin, J. Constitutive heterochromatin formation and transcription in mammals. Epigenetics 8, 3 (2015).
7. Aramayo, R. & Selker, E.U. Neurospora crassa, a model system for epigenetics research. Cold Spring Harb. Perspect. Biol. 5, a017921 (2013).
8. Bull, J.H. & Wootton, J.C. Heavily methylated amplified DNA in transformants of Neurospora crassa. Nature 310, 701–704 (1984).
9. Selker, E.U. & Garrett, P.W. DNA sequence duplications trigger gene inactivation in Neurospora crassa. Proc. Natl. Acad. Sci. USA 85, 6870–6874 (1988).
10. Freitag, M. et al. DNA methylation is independent of RNA interference in Neurospora. Science 304, 1939 (2004).
11. Chicas, A., Forrest, E.C., Scepich, S., Cogoni, C. & Macino, G. Small interfering RNAs that trigger posttranscriptional gene silencing are not required for the histone H3 lysine 9 methylation necessary for transgenic tandem repeat stabilization in Neurospora crassa. Mol. Cell. Biol. 25, 3793–3801 (2005).
12. Garrick, D., Fiering, S., Martin, D.I. & Whitelaw, E. Repeat-induced gene silencing in mammals. Nat. Genet. 18, 56–59 (1998).
13. Preis, J.I., Downes, M., Oates, N.A., Rasko, J.E. & Whitelaw, E. Sensitive flow cytometric analysis reveals a novel type of parent-of-origin effect in the mouse genome. Curr. Biol. 13, 955–959 (2003).
14. van der Maarel, S.M., Tawil, R. & Tapsic, S.J. Facioscapulohumeral muscular dystrophy and DUX4: breaking the silence. Trends Mol. Med. 17, 252–258 (2011).
15. Galagan, J.E. & Selker, E.U. RIP: the evolutionary cost of genome defense. Trends Genet. 20, 417–423 (2004).
16. Hane, J. et al. in Genetic Transformation Systems in Fungi (eds. van den Breg, M.A. & Maruthachalam, K.) Vol. 2, 55–68 (Springer International Publishing, Switzerland, 2015).
17. Rossignol, J.L. & Faugeron, G. Gene inactivation triggered by recognition between DNA repeats. Experientia 50, 307–317 (1994).
18. Gladyshev, E. & Kleckner, N. Direct recognition of homology between double helices of DNA in Neurospora crassa. Nat. Commun. 5, 3509 (2014).
19. Gladyshev, E. & Kleckner, N. Recombination-independent recognition of DNA homology for repeat-induced point mutation (RIP) is mediated by the underlying nucleotide sequence. PLoS Genet. 12, e1006015 (2016).
20. Gladyshev, E. & Kleckner, N. Recombination-independent recognition of DNA homology for repeat-induced point mutation. Curr. Genet. http://dx.doi.org/10.1007/s00294-016-0649-4 (2016).
21. Freitag, M., Williams, R.L., Kothe, G.O. & Selker, E.U. A cytosine methyltransferase homologue is essential for repeat-induced point mutation in Neurospora crassa. Proc. Natl. Acad. Sci. USA 99, 8802–8807 (2002).
22. Clutterbuck, A.J. Genomic evidence of repeat-induced point mutation (RIP) in filamentous ascomycetes. Fungal Genet. Biol. 48, 306–326 (2011).
23. Amselem, J., Lebrun, M.H. & Quesneville, H. Whole genome comparative analysis of transposable elements provides new insight into mechanisms of their inactivation in fungal genomes. BMC Genomics 16, 141 (2015).
24. Kouzminova, E. & Selker, E.U. dim-2 encodes a DNA methyltransferase responsible for all known cytosine methylation in Neurospora. EMBO J. 20, 4309–4323 (2001).
25. Jamieson, K. et al. Loss of HP1 causes depletion of H3K27me3 from facultative heterochromatin and gain of H3K27me2 at constitutive heterochromatin. Genome Res. 26, 97–107 (2016).
26. Basenko, E.Y. et al. Genome-wide redistribution of H3K27me3 is linked to genotoxic stress and defective growth. Proc. Natl. Acad. Sci. USA 112, E6339–E6343 (2015).
27. Zhang, X. et al. Structure of the Neurospora SET domain protein DIM-5, a histone H3 lysine methyltransferase. Cell 111, 117–127 (2002).
28. Tropschug, M., Barthelmess, I.B. & Neupert, W. Sensitivity to cyclosporin A is mediated by cyclophilin in Neurospora crassa and Saccharomyces cerevisiae. Nature 342, 953–955 (1989).

29. Roberts, R.J., Vincze, T., Postai, J. & Macelis, D. REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. Nucleic Acids Res. 43, D298–D299 (2015).

30. Pontecorvo, G. Structure of heterochromatin. Nature 153, 365–367 (1944).

31. Dorer, D.R. & Henikoff, S. Expansions of transgene repeats cause heterochromatin formation and gene silencing in Drosophila. Cell 77, 993–1002 (1994).

32. Malagnac, F. et al. A gene essential for de novo methylation and development in Ascobolus reveals a novel type of eukaryotic DNA methyltransferase structure. Cell 91, 281–290 (1997).

33. Selker, E.U., Fritz, D.Y. & Singer, M.J. Dense nonsymmetrical DNA methylation resulting from repeat-induced point mutation in Neurospora. Science 262, 1724–1728 (1993).

34. Belden, W.J., Lewis, Z.A., Selker, E.U., Laros, J.J. & Dunlap, J.C. CHD1 remolds chromatin and influences transient DNA methylation at the clock gene frequency. PLoS Genet. 7, e1002166 (2011).

35. Ohm, R.A. et al. Diverse lifestyles and strategies of plant pathogenesis encoded in the genomes of eighteen Dothideomycetes fungi. PLoS Pathog. 8, e1003037 (2012).
ONLINE METHODS

Plasmids. Plasmids pEAG66, pEAG115A, pEAG186B, pEAG186K and pEAG186L are based on the pCSR1 plasmid36. Plasmids pEAG82B (inactive copy of the csr-1 gene), pEAG236A (the 4× construct), pEAG244B (active copy of the dim-5 gene) and pEAG244G (inactive copy of the dim-5 gene) are based on the pMF280 (ref. 37) and pMF334 (ref. 38) plasmids. Annotated maps of pEAG66, pEAG115A, pEAG186B, pEAG186K and pEAG186L were published previously18. Annotated maps of pEAG82B, pEAG236A, pEAG244B and pEAG244G are provided in Supplementary Data 1 (in GenBank format; individual plain-text files are compressed with tar or gz).

Manipulation of Neurospora strains. All strains used in this study are listed in Supplementary Table 1. All FGSC# strains were obtained from the Fungal Genetics Stock Center39. Transformation of the endogenous gene was integrated into the endogenous parental nuclei that become assorted into dikaryotic ascogenous cells40. RIP is likely activated in haploid parental nuclei that become assorted into dikaryotic ascogenous cells40. Manipulation of Neurospora strains was integrated into the endogenous gene) and pEAG244G (inactive copy of the dim-5 gene) were based on the pCSR1 plasmid36. Plasmids pEAG66, pEAG115A, pEAG186B, pEAG186K and pEAG186L are based on the pMF280 (ref. 37) and pMF334 (ref. 38) plasmids. Annotated maps of pEAG66, pEAG115A, pEAG186B, pEAG186K and pEAG186L were published previously18. Annotated maps of pEAG82B, pEAG236A, pEAG244B and pEAG244G are provided in Supplementary Data 1 (in GenBank format; individual plain-text files are compressed with tar or gz).

Analysis of the cyclosporin-resistant phenotype. This assay was designed to detect very low levels of RIP activity involving two widely separated repeats of the csr-1 gene20,28. A copy of the csr-1 gene (including the promoter region) was integrated into the his-3 locus, 2.7 Mb away from the endogenous csr-1. The ectopic copy carries a single guanine-to-adenine mutation in the 5′-splice site, which confers resistance to cyclosporin. Thus, inactivation of the endogenous csr-1 allele was sufficient to confer a cyclosporin-resistant phenotype. All ejected spores (both early and late arising) were collected into 1 ml of distilled water. Spores were heat-shocked for 30 min at 60 °C to induce germination. To estimate the total number of viable spores, 10 µl of the original 1× suspension was diluted 100-fold (back to 1 ml) and 50 µl of the 1:100 dilution was plated onto non-selective sorbose agar solidified in standard Petri dishes (100 mm × 15 mm). Three replicate dishes were plated in total, consuming 3 × 50 µl of the 1:100 dilution. The number of viable spores was estimated as the average number of colonies per dish multiplied by the overall dilution factor. To estimate the number of cyclosporin-resistant progeny in crosses with strong RIP, an additional 3 × 50 µl from the same 1:100 dilution was plated in triplicate on sorbose agar containing cyclosporin A (5 µg/ml; Sigma, 30024–25MC). The total number of cyclosporin-resistant progeny was estimated by the same formula. In cases of weak RIP (expected in the absence of RIP in the absence of the ectopic csr-1 copy), the remaining 0.99 µl of the original 1:1 suspension was plated directly on selective medium and the number of cyclosporin-resistant progeny was determined by the total count of growing colonies.

Sampling RIP products for sequence analysis. RIP is likely activated in haploid parental nuclei that become assorted into dikaryotic ascogenous cells40. Approximately 6–7 rounds of RIP may occur as these premeiotic nuclei continue to divide by mitosis. After the nuclei undergo karyogamy and meiosis, the haploid meiotic products are packaged into progeny spores and ejected from the mature peritheium (the fruiting body). RIP mutation tends to be intrinsically weak in ‘early-arising’ spores (ejected within the first 1–3 d after the onset of sporogenosis) but then increases in ‘late-arising’ spores (ejected over the next 7–10 d). 100–200 spore-producing perithecia can normally develop in a single test tube under our experimental conditions18. Each peritheium represents an autonomous anatomical structure; it hosts several independent lineages of ascogenous cells and ultimately produces several hundred progeny spores. All ejected spores correspond to one statistical population. Late-arising spores are sampled without any knowledge of their RIP status. Because the number of sampled spores (24–60 from each cross) is much smaller than the estimated number of ascogenous lineages, each spore effectively represents an independent measure of RIP activity. Despite the apparent spore-to-spore variability in RIP levels, our studies have indicated that the expected number of RIP mutations (the arithmetic mean) represents an accurate and useful measure of RIP activity18,19.

Genomic DNA extraction, PCR amplification and sequencing. Genomic DNA was extracted from individual spore clones as described previously18. PCR products were purified with the Omega Bio-Tek kit and sequenced directly by primer walking. The following primers were used for PCR and sequencing (primer sequences are provided in Supplementary Table 3). The 802-bp construct was amplified with primers P66_Seq1 and RIP2_R1 and sequenced with primers P66_Seq1, P66_Seq3, LNK_SeqR1 and RIP2_R2. All (220 + 200)-bp constructs were amplified with primers P66_Seq3 and RIP2_R1 and sequenced with primers P66_Seq17 and P66_Seq18. The 4× con- struct was amplified with primers NcHis3_R6 and NcHis3_F7 and sequenced with primers NcHis3_F4, NcHis3_F7, NcHis3_R1, NcHis3_R4, NcHis3_R6, P236A_Seq1, P236A_Seq2, P236A_Seq4, P236A_Seq5 and P236A_Seq8. If the entire 4× construct could not be amplified as a single fragment, piecewise applications were carried out using pairwise combinations of the above sequencing primers. The endogenous csr-1 gene was amplified with primers P66_Seq10 and RIP2_R1 and sequenced with primers P66_Seq10, P66_Seq12, CSR1_Seq1, CSR1_SeqR and RIP2_R2. Eight progeny spores with particularly strong levels of DIM-2-mediated mutation were selected for further sequence analysis of the adjoining genomic regions. For these eight clones, the extended 5.4-kb region containing the endogenous csr-1 gene was amplified with primers P66_Seq3 and RIP2_R3 and sequenced with additional primers P66_Seq3, P66_Seq9, P66_Seq4, CSR1_SeqR2, CSR1_SeqR3 and RIP2_R3. If a primer site appeared to be affected by RIP, additional ad hoc primers were used instead. Sequencing reactions were read with an ABI3730xl DNA analyzer at the DNA Resource Core of the Dana-Farber/Harvard Cancer Center (funded in part by NCI Cancer Center support grant 2P30CA06516–48). Individual chromatograms were assembled into contigs with Phred/Phrap. All assembled contigs were validated manually using Consed.

Sequence and statistical analysis. For each cross, assembled contigs were aligned with the parental (reference) sequence using ClustalW. All alignments analyzed in this study are provided in Supplementary Data 2 in ClustalW format; individual plain-text files are compressed with tar/gz. Mutations were detected and analyzed as described previously18. The absolute level of RIP in a given region of interest was expressed (i) as the expected number of mutations (the arithmetic mean) and, if necessary, (ii) as the percentage of mutated strands, where all cytosine-to-thymine mutations were considered to be on the top strand and all guanine-to-adenine mutations were considered to be on the bottom strand31, s.e.m. was used as a measure of variation. The following regions were analyzed for the 802-bp construct: the endogenous left repeat copy (positions 348–1149), the 600-bp segment of the linker (positions 1227–1826) and the ectopic right repeat copy (positions 1879–2680). The following region was analyzed for the (220 + 200)-bp construct: the 450-bp segment of the linker region (positions 627–1071). Empirical distributions of mutation counts (the numbers of cytosine-to-thymine and guanine-to-adenine mutations found together within a given region on a per-sphere basis) were compared by Kolmogorov–Smirnov test18. The pairwise differences in the percentage of mutated strands and the proportion of linker mutations were compared for significance as the original raw counts by (i) Pearson’s χ2 homogeneity test with Yates’ continuity correction and (ii) Fisher’s exact test (as implemented in R).

Analysis of cytosine methylation in vegetative cells of N. crassa. Purification of genomic DNA. Crude preparations of genomic DNA were obtained by the method used to extract DNA for RIP analysis18, except that phenol–chloroform was replaced with chloroform in all purification steps. DNA was precipitated with isopropanol, washed with ethanol, resuspended in 80 µl of water and mixed with 20 µl of the 5× loading buffer (20% glycerol, 0.25% SDS, 5 mM EDTA, 1 mg/ml xylene cyanol). The entire sample (100 µl) was loaded into a single well of a 0.75% agarose mini-gel (NuSieve GTG Agarose, 50080) containing ethidium bromide. The gel was run at 3 V/cm in 1× TAE (Tris–acetate–EDTA) buffer for 1.5 h. A region of the gel containing high-molecular-weight DNA was excised with a sterile razor blade under long UV light and digested with β-agarase I (NEB, M0392L; 3 h at 42 °C). Following a single
Restriction digest with BstUI. The BstUI recognition sequence (5′-CGCG-3′) contains four cytosines that can potentially be methylated by DIM-2 (two cytosines on the top strand and two cytosines on the bottom strand), and methylation of any one (or more) of these cytosines protects the site from cleavage by BstUI (ref. 29). Each restriction reaction was set up in a total volume of 50 µl, 0.2 µg of purified genomic DNA from the previous step and 4 µl of BstUI (10 U/µl, NEB, R0518S). Reactions were incubated for 3 h at 50 °C. BstUI was inactivated with proteinase K (by adding 2 µl of proteinase solution, 20 mg/ml, and incubating at 50 °C for an additional 1.5 h). Proteinase K was heat inactivated (95 °C for 10 min), and the volume of each sample was adjusted to 200 µl (by adding 148 µl of water). Mock-digested samples were processed in exactly the same way, except that 4 µl of 50% glycerol was added to each reaction instead of BstUI.

Semiquantitative PCR. Each PCR was run in a total volume of 10 µl. Three genomic regions were amplified separately for each DNA sample: region A (primers REG_A_F1 and REG_A_R1), region B (primers REG_B_F1 and REG_B_R1) and region C (primers REG_C_F1 and REG_C_R1). DNA sample concentrations were adjusted (by no more than a factor of 1.5) to produce comparable PCR yields of region B. No additional adjustments were implemented. Each DNA sample (corresponding to 20–36 µg of purified genomic DNA from the previous step) was used to generate one principal master mix (PMM) for 18 individual PCRs. Each PMM also received 18 µl of the 10× reaction buffer (EconoTaq DNA Polymerase, Lucigen, F93366-1), 7.2 µl of the dNTP solution (NEB, N0447S) and 5.4 µl of EcoTaq polymerase (in a total volume of 162 µl). Each PMM was split equally into three 54-µl region-specific master mixes (RSMMs). Each RSMM received 6 µl of a primer solution (containing one pair of primers corresponding to a given PCR region, at 5 pmol/µl each), bringing its total volume to 60 µl. Each RSMM was then split into six individual 10-µl reactions.

All PCRs were run in a C1000 Touch thermal cycler (Bio-Rad) using the following program: incubation at 94 °C for 2 min followed by N cycles of 94 °C for 20 s, 59 °C for 20 s and 72 °C for 20 s and a final incubation at 72 °C for 2 min. For all digested DNA samples, the value of N (the number of amplification cycles) was set to 27. For all mock-digested samples, the value of N was set to 22. The number of cycles was determined empirically to ensure that all PCRs remained in the exponential range. Each PCR product was mixed with 2 µl of 5× buffer and loaded onto a 2% regular agarose gel in a total volume of 12 µl. All PCRs corresponding to a given DNA sample were analyzed side by side on the same gel. Gels were run for 30 min in 0.5× TAE at 6 V/cm and scanned with 50-micron resolution using a Molecular Imager FX instrument (Bio-Rad). Luminosities of PCR products were measured in ImageJ (with the Analyze → Measure tool) using fixed region-of-interest parameters. A constant value corresponding to the image background (estimated separately on the basis of five independent measurements) was subtracted from the raw luminosity value for each PCR band. The relative number of uncleavable BstUI sites in the digested DNA sample (in region A) was estimated using a calibration curve based on a series of twofold dilutions of the corresponding mock-digested DNA sample.

Data availability. All data generated or analyzed during this study are included in this published article and its supplementary information files.

36. Bardiya, N. & Shiu, P.K. Cyclosporin A–resistance based gene placement system for Neurospora crassa. Fungal Genet. Biol. 44, 307–314 (2007).
37. Freitag, M., Hickey, P.C., Raju, N.B., Selker, E.U. & Read, N.D. GFP as a tool to analyze the organization, dynamics and function of nuclei and microtubules in Neurospora crassa. Fungal Genet. Biol. 41, 897–910 (2004).
38. Freitag, M. & Selker, E.U. Expression and visualization of red fluorescent protein (RFP) in Neurospora crassa. Fungal Genet. Newsl. 52, 14–17 (2005).
39. McCluskey, K., Wiest, A. & Plamann, M. The Fungal Genetics Stock Center: a repository for 50 years of fungal genetics research. J. Biosci. 35, 119–126 (2010).
40. Arnaisse, S., Zickler, D., Bourdais, A., Dequard-Chablat, M. & Debuchy, R. Mutations in mating-type genes greatly decrease repeat-induced point mutation process in the fungus Podospora anserina. Fungal Genet. Biol. 45, 207–220 (2008).