rational, standardized methodology for measuring and characterizing EC and BC (*15*). In the meantime, models simulating the climate effect of BC must use emission inventories derived from a myriad of measurements by an unspecified combination of two poorly characterized, nonequivalent techniques. This makes for a hazy understanding of BC's climatic effect. Until these problems are addressed, the issue of BC's role in global warming and what to do about it will be difficult to resolve.

**References and Notes**

1. J. T. Houghton *et al.*, Eds., *Climate Change 2001: The Scientific Basis* (Cambridge Univ. Press, Cambridge, UK, 2001).
2. S. Arrhenius, *Worlds in the Making*, H. Borns, transl. (Harper & Brothers, New York and London, 1908).
3. V. Ramanathan *et al.*, *J. Geophys. Res.* **90**, 5547 (1985).
4. R. J. Charlson *et al.*, *Science* **255**, 423 (1992).
5. A. S. Ackerman *et al.*, *Science* **288**, 1042 (2000).
6. M. A. Jacobson, *Nature* **409**, 695 (2001).
7. J. Hansen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9875 (2000).
8. S. Menon, J. Hansen, L. Nazarenko, Y. Luo, *Science* **297**, 2250 (2002).
9. Q. Xu, *Atmos. Environ.* **35**, 5029 (2001).
10. J. S. Lighty, J. M. Veranth, A. F. Sarofim, *J. Air. Waste Manag. Assoc.* **50**, 1565 (2000).
11. R. A. Kerr, *Science* **294**, 765 (2001).
12. J. C. Chow *et al.*, *Aerosol Sci. Technol.* **34**, 23 (2001).
13. P. A. Solomon *et al.*, *J. Geophys. Res.*, in press.
14. H. Horvath, *Atmos. Environ.* **27A**, 293 (1993).
15. T. C. Bond, R. J. Charlson, J. Heintzenberg, *Geophys. Res. Lett.* **25**, 337 (1998).
16. We acknowledge support from the National Science Foundation through grants ATM-0129495 (W.L.C.) and ATM-0096245 and ATM-0080325 (M.B.), as well as from NASA through a Goddard Visiting Fellowship in Earth Sciences (W.L.C.) administered through the University of Maryland–Baltimore County and hosted by the Goddard Institute for Space Studies, where W.L.C. resided during preparation of this Perspective.

PERSPECTIVES: MOLECULAR BIOLOGY

# An RNA-Guided Pathway for the Epigenome

**Thomas Jenuwein**

Like twins separated by early evolution, DNA and RNA are the principal molecules for storing and transducing genetic information. Messenger RNAs (mRNAs) are transcribed from genes that constitute as little as 10% of the genome. Most of the remaining noncoding DNA is occupied by a diverse family of repeats including arrays of satellite sequences and transposons. Since the pioneering work of Muller (*1*) and McClintock (*2*), these apparently "nonspecific" sequences have been thought of as "epigenetic elements" that control the stability of gene expression programs and organize heterochromatic domains at centromeres and telomeres (*3*). Although this concept has been instrumental in guiding epigenetics research (*4*), it has also posed a major paradox: How can stretches of DNA that appear so heterogeneous in sequence organize similar higher order chromatin structures (*5*)?

A recent flurry of papers published in *Science* and *Cell* have started to unravel this riddle by demonstrating the presence of short RNA transcripts that are complementary to repeats in centromeres and elsewhere in the genome (*6*–*9*). These short transcripts resemble small interfering RNAs (siRNAs) that bind to complementary RNA sequences (*6*), targeting them for destruction. In the fission yeast *Schizosaccharomyces pombe*, mutants of the RNA interference (RNAi) machinery impair the definition of heterochromatin (*7*, *8*). Analogous mutants in the ciliate protozoan *Tetrahymena thermophila* affect the most dramatic reshaping of the genome, the directed elimination of DNA sequences (*9*). These exciting findings suggest that short heterochromatic RNAs (shRNAs) (*10*) are among the primary signals required for converting nonspecific sequence information into distinct chromatin states, thereby regulating the plasticity of the "epigenome." (The epigenome can be defined as a collection of biochemical modifications to chromatin that indexes genetic information.)

Heterochromatin in a wide variety of organisms is characterized by high density methylation of H3-K9 (that is, methylation of histone H3 protein at amino acid lysine 9). This modification results in binding of the chromo domain protein HP1 (heterochromatin protein 1) to chromatin (*11*). In turn, H3-K9 methylation directs DNA methylation in the fungus *Neurospora crassa* (*12*) and in the weed *Arabidopsis thaliana* (*13*). In plants, DNA methylation of integrated transgenes is thought to depend on complementary RNA signals (*14*, *15*). These links prompted Volpe and colleagues (*7*) to investigate whether *S. pombe* mutants with defects in the RNAi machinery had altered H3-K9 methylation. They disrupted the genes encoding the *S. pombe* homologs of proteins involved in RNAi: dicer (*dcr*), RNA-dependent RNA polymerase (*RdRp*), and argonaute (*ago*). Deletion of these genes resulted in the transcriptional derepression of transgenes integrated at pericentric repeats in the outer *cen* region of centromeres. It also resulted in loss of H3-K9 methylation and the association of the Swi6 protein (the *S. pombe* HP1 homolog) with this chromatin domain. Importantly, the *dcr*, *RdRp*, and *ago* mutants displayed aberrant accumulation of complementary RNA transcripts (1.4 and 2.4 kb) from both strands of the outer *cen* region (*7*). These results suggest that double-stranded RNA (dsRNA) originating from the pericentric repeats triggers the nuclear production of shRNAs, which in turn initiates H3-K9 methylation and heterochromatin formation. Indeed, ~22-nucleotide long RNAs with 5′-phosphate and 3′-hydroxyl groups, complementary to both strands of the outer *cen* repeats present at all three *S. pombe* centromeres, were detected in wild-type cells (*6*).

If DNA repeats represent nuclear foci for RNAi-like processing of dsRNA and the subsequent initiation of heterochromatin formation, then this model predicts that almost any repetitive DNA transferred to an ectopic (nonphysiologic) site should induce a heterochromatic state. In an elegant series of experiments reported on page 2232 of this issue, Hall and colleagues (*8*) inserted a *cen*-related DNA fragment (a 3.6-kb *cenH* repeat) normally present at the *S. pombe* silent mating–type (*mat*) domain in a euchromatic position (*ura4* locus). This repeat was sufficient to initiate variegated silencing of a linked reporter gene and also induced H3-K9 methylation and recruitment of Swi6.

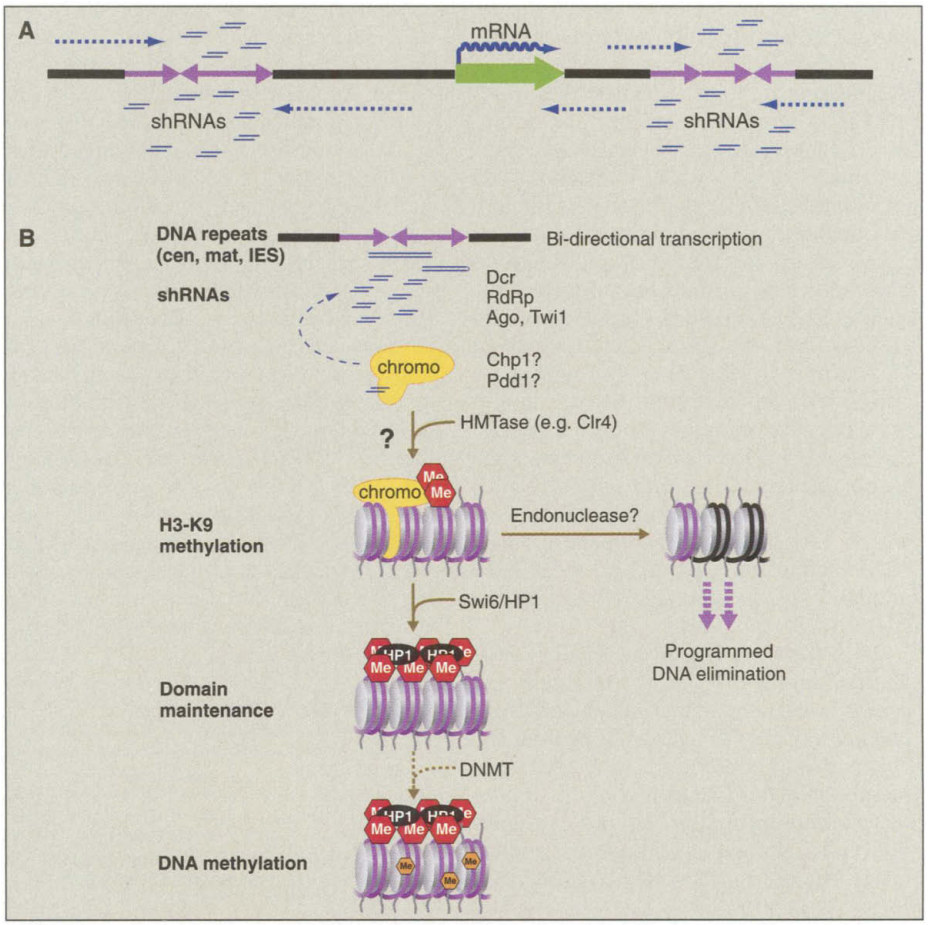
*CenH*-dependent heterochromatinization at this ectopic site was again abrogated in *dcr*, *RdRp*, and *ago* mutants. Additionally, Hall *et al.* provide evidence that the RNAi machinery is required for the initiation but not the maintenance of the heterochromatic state. This is suggested by the persistence of H3-K9 methylation at endogenous *cenH-mat* sites in mutant strains with abrogated RNAi. Furthermore, segregant analysis of offspring from diploid strains, which are heterozygous for the H3-K9 histone methyltransferase (HMTase) *clr4* but null for RNAi mutants, revealed that these cells fail to establish H3-K9 methylation at *cenH-mat*. Complementary data were obtained after treatment of mutant cells with a histone deacetylase inhibitor, which results in hyperacetylation of chromatin and transcriptional derepression of silent *cenH-mat* epialleles; rapid

The author is with the Research Institute of Molecular Pathology (IMP), Vienna Biocenter, Dr. Bohrgasse 7, A-1030 Vienna, Austria. E-mail: jenuwein@nt.imp.univie.ac.at

reversion to a repressed state only occurred in RNAi-competent strains. Interestingly, H3-K9 methylation can be initiated even in the absence of Swi6, which appears to be required for the stabilization and propagation of heterochromatin (*8*). Mutant screens for suppressors of centromeric silencing are consistent with this model because they reveal that other chromo domain genes (*chp1*) and *ago1* alleles (*16*) are important for heterochromatin formation in *S. pombe* at natural centromeres and at ectopic sites (*17*) containing centromeric repeat DNA.

How conserved is the involvement of shRNAs and components of the RNAi machinery in triggering chromatin organization of the epigenome at DNA repeats? It is not surprising that supporting evidence arrives from the study of *Tetrahymena*, which has been a particularly revealing organism for epigenetic research. *Tetrahymena* is characterized by nuclear dimorphism: It has a transcriptionally inactive micronucleus (the germ cell nucleus), which gives rise to a transcriptionally active macronucleus (the somatic cell nucleus). This developmental transition is accompanied by elimination of ~15% of genomic micro-DNA (micDNA) at more than 6000 sites called internal eliminated sequences (IESs).

Reporting in *Cell*, Mochizuki *et al.* (*9*) show that programmed DNA rearrangement in *Tetrahymena* is severely impaired in a mutant containing a defect in the *twi1* gene (a *piwi*-related gene of the *argonaute* family). They also report that wild-type cells contain ~26- to 31-nucleotide RNAs (5′-phosphate, 3′-hydroxyl groups) that preferentially hybridize to genomic mic-DNA. These short RNAs are largely absent in *twi1* mutants and their accumulation is delayed in another mutant *pdd1* (programmed DNA degradation) (*9*), previously shown to have impaired elimination of DNA. Mochizuki and colleagues suggest that these short RNAs could serve as "scanning molecules" to specify target sequences through a pairing-like mechanism. In a related report, Taverna *et al.* (*18*) demonstrate that Pdd1 links H3-K9 methylation to programmed DNA degradation. Pdd1 contains two chromo domains and one additional domain with sequence similarity to a motif that may associate with RNA (*19*). Pdd1 binds to H3-K9 methylated peptides in vitro, and *pdd1* mutants abrogate bulk H3-K9 methylation in *Tetrahymena* micronuclei. Importantly, Pdd1 is present at endogenous IES sites and is sufficient to trigger DNA elimination from ectopic IES targets when recruited in the form of a lexA-fusion protein (*18*). Together, these data suggest that Pdd1 contains multiple chromo domains because it may




**shRNAs regulate heterochromatin formation.** (**A**) DNA repeats may be nuclear foci for the generation and accumulation of short heterochromatic RNAs (shRNAs). These RNAs are important for epigenetic processes associated with the initiation and formation of heterochromatin. Shown is a hypothetical genomic region (black) containing DNA repeats (purple) and a gene or transcription unit (green). Promoter-driven mRNA is indicated by the wavy line (dark blue), and spurious intergenic transcription is represented by the blue dashed lines. The directed abundance of mRNAs and their export to the cytoplasm greatly diminishes the potential to form short nuclear dsRNAs. (**B**) Simplified model highlighting the hierarchy of processes that induce epigenetic transitions triggered by shRNAs originating from DNA repeats. (DNMT, DNA methyltransferase.)

be a "bridging molecule" for connecting a target DNA sequence (through RNA pairing) to a distinct histone modification (H3-K9 methylation). Because bidirectional transcription occurs across IES repeats (*20*), the *Tetrahymena* data indicate a remarkably similar mechanism (as proposed for induction of *S. pombe* heterochromatin), in which focal generation of short RNAs may be one of the primary signals promoting H3-K9 methylation and subsequent epigenetic modification of the genome.

The beauty of the new discoveries is that they allow a unifying mechanism for the targeted restructuring of the epigenome. A central tenet of this model is that the "somatic" genome displays spurious transcriptional activity (*21*), allowing for continuous interplay between DNA and RNA molecules. Possibly, the relative abundance of dsRNA versus promoter-driven transcripts may have been adopted during early evolution to discriminate between "genome surveillance" and "genome activation" systems. If there is spurious and bidirectional transcription on a genome-wide scale, then the clustered presence of repeat sequences will greatly favor dsRNA formation and the focal nuclear accumulation of shRNAs (see the figure). In contrast, authentic promoter sequences will direct the stabilization of abundant mRNAs, which are then exported to the cytoplasm. Thus, the pathway to a silenced chromatin domain may be initiated by a combination of two generic signals: an array of DNA repeats and the local concentration of shRNAs where neither of these signals requires a heterochromatin-specific sequence.

Although the generation of shRNAs depends upon components of the RNAi machinery, their apparent nuclear prevalence distinguishes them from siRNAs or other small or micro-RNAs (*22*, *23*). Inter-

estingly, RdRp is present at the chromatin regions comprising the outer *cen* repeats in *S. pombe* (*7*). Rather than inducing degradation or translational inhibition of mRNA, the shRNAs target chromatin modifications, presumably through a pairing mechanism (*24*), and interaction with chromatin-associated proteins. Although this recognition step is currently unclear, good candidates for linking shRNAs to chromatin are chromo domain proteins, such as Chp1 (*17*), Pdd1 (*18*), and the Clr4 HMTase itself. Intriguingly, the *Drosophila* Clr4 homolog, Su(var)3-9, contains an amino-terminal extension that resembles the elongation initiation factor eIF2γ (*25*). Following targeted H3-K9 methylation, the chromatin region can either be stably silenced by Swi6/HP1 incorporation (*8*), which may entail further repression by DNA methylation (*12*, *13*) or, in its most dramatic form, result in DNA elimination (*9*, *18*) (see the figure).

With the exception of the budding yeast *Saccharomyces cerevisiae*, which does not appear to contain components of the RNAi machinery or of H3-K9 methylation, the new pathway is predicted to trigger heterochromatin formation in complex organisms including mammals. For example, bidirectional transcription across mouse major satellite repeats has been described (*26*), suggesting that there are shRNAs in pericentric heterochromatin. In agreement with this interpretation, the definition of pericentric H3-K9 methylation is abrogated after ribonuclease treatment of permeabilized mouse cells (*27*). To extend these parallels further, it is conceivable that the initiating mechanisms for X-chromosome inactivation and imprinting, or even for transcriptional silencing mediated by Polycomb-group (Pc-G) proteins, may also use short RNAs. X-chromosome inactivation requires the *Xist* RNA as part of a large RNA scaffold; in addition, several Pc-G proteins are associated with the inactive X chromosome (*28*). Pc-G gene mutations impair RNAi in the worm *Caenorhabditis elegans* (*29*), and small RNAs direct dosage compensation in *Drosophila* (*30*). Thus, the new discoveries in *S. pombe* (*6*–*8*) and *Tetrahymena* (*9*, *18*) may reveal a key unifying signal for inducing chromatin alterations in most of the epigenetic transitions that occur during cellular differentiation.

**References and Notes**


1. H. J. Muller, *Proc. Int. Congr. Genet.* **1**, 213 (1932).
2. B. McClintock, *Cold Spring Harbor Symp. Quant. Biol.* **16**, 13 (1951).
3. M. L. Pardue, J. G. Gall, *Science* **168**, 1356 (1970).
4. G. H. Karpen, R. C. Allshire, *Trends Genet.* **13**, 489 (1997).
5. A. K. Csink, S. Henikoff, *Trends Genet.* **14**, 200 (1998).
6. B. J. Reinhart, D. P. Bartel, *Science* **297**, 1831 (2002); published online 22 August 2002 (10.1126/science.1077183).
7. T. A. Volpe *et al.*, *Science* **297**, 1833 (2002); published online 22 August 2002 (10.1126/science.1074973).
8. I. M. Hall *et al.*, *Science* **297**, 2232 (2002).
9. K. Mochizuki, N. A. Fine, T. Fujisawa, M. A. Gorovsky, *Cell*, published online 14 August 2002 (10.1016/S0092867402009091).
10. F. E. Vaistij, L. Jones, D. C. Baulcombe, *Plant Cell* **14**, 857 (2002).
11. M. Lachner, T. Jenuwein, *Curr. Opin. Cell Biol.* **14**, 286 (2002).
12. H. Tamaru, E. U. Selker, *Nature* **414**, 277 (2001).
13. J. P. Jackson, A. M. Lindroth, X. Cao, S. E. Jacobsen, *Nature* **416**, 556 (2002).
14. A. J. Hamilton, D. C. Baulcombe, *Science* **286**, 950 (1999).
15. M. F. Mette, W. Aufsatz, J. van der Winden, M. A. Matzke, A. J. Matzke, *EMBO J.* **19**, 5194 (2000).
16. R. Allshire, personal communication.
17. J. F. Partridge *et al.*, *Curr. Biol.* in press.
18. S. D. Taverna, R. S. Coyne, C. D. Allis, *Cell*, published online 6 September 2002 (10.1016/S0092867402009418).
19. A. Akthar, D. Zink, P. B. Becker, *Nature* **407**, 405 (2000).
20. D. L. Chalker, M.-C. Yao, *Genes Dev.* **15**, 1287 (2001).
21. A. Bird, *Trends Genet.* **11**, 94 (1995).
22. M. Matzke, A. J. Matzke, J. M. Kooter, *Science* **293**, 1080 (2001).
23. G. Hutvagner, P. D. Zamore, *Curr. Opin. Genet. Dev.* **12**, 225 (2002).
24. J. A. Birchler, M. Pal-Bhadra, U. Bhadra, *Curr. Opin. Genet. Dev.* **10**, 211 (2000).
25. B. Tschiersch *et al.*, *EMBO J.* **13**, 3822 (1994).
26. F. Rudert, S. Bronner, J. M. Garnier, P. Dolle, *Mamm. Genome* **6**, 76 (1995).
27. C. Maison *et al.*, *Nature Genet.* **30**, 329 (2002).
28. W. Mak *et al.*, *Curr. Biol.* **12**, 1016 (2002).
29. N. R. Dudley, J. C. Labbe, B. Goldstein, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 4191 (2002).
30. R. L. Kelley, M. I. Kuroda, *Cell* **103**, 9 (2000).

31. I thank M. Matzke (Salzburg), A. Bird (Edinburgh), and K. Nasmyth (Vienna) for helpful comments.

PERSPECTIVES: BIOMEDICINE

# Gluten and the Gut—Lessons for Immune Regulation

**Detlef Schuppan and Eckhart G. Hahn**

Celiac sprue is an inflammatory disease that leads to destruction of the microscopic fingerlike projections of the small intestine called villi. The disease is triggered by ingestion of the gluten proteins contained in wheat, barley, and rye, and symptoms range from minor complaints to severe nutrient malabsorption (*1*). The report by Shan *et al.* on page 2275 of this issue (*2*) significantly enhances our understanding of celiac sprue pathogenesis and hints at potential intervention strategies to treat this common disease.

There are three remarkable features of celiac sprue: (i) It usually remits upon strict dietary avoidance of gluten; (ii) it requires a unique genetic background for antigen presentation—expression of the human leukocyte antigen (HLA) class II molecules DQ2 or DQ8; and (iii) patients have characteristic circulating mucosal (immunoglobulin A) autoantibodies to the ubiquitous enzyme tissue transglutaminase (tTGase) (*3*, *4*). What is intriguing is the connection between these three features. Gluten peptides presented in the context of HLA-DQ2 or HLA-DQ8 molecules elicit proliferation of intestinal T cells from celiac sprue patients and induce these cells to release inflammatory cytokines (*5*). The autoantigen, tTGase, catalyzes transamidation between a glutamine residue of peptide 1 (glutamine donor) and a lysine residue of peptide 2 (glutamine acceptor), creating stable covalent complexes among a limited set of mostly extracellular matrix proteins (*6*). This enzyme is highly expressed in the subepithelial cells of the gut, where it is stored in an intracellular inactive form. It is released in response to mechanical or inflammatory stress and is activated by high extracellular calcium levels. The strong affinity of tTGase for gluten reflects the fact that 30 to 50% of the amino acids in gluten are glutamine. This enzyme induces formation of aggregates of gluten and other antigens, which seems to be important for efficient antigen uptake by antigen-presenting cells of the immune system (*7*). At low pH or in the absence of glutamine acceptors, tTGase deamidates certain glutamine residues of gluten to glutamic acid. This posttranslational modification enhances binding of gluten epitopes to HLA-DQ2 or HLA-DQ8 and potentiates their ability to stimulate T cells (*8*). Several immunodominant gluten peptides (all substrates of tTGase) have been identified, but it is unclear to what extent these peptides reach the small intestinal mucosa after exposure to gastric and duodenal proteases.

In the new work, Shan and colleagues (*2*) isolated a unique 33–amino acid peptide from the 266–amino acid α2-gliadin (the homologous gliadins represent the major storage proteins of wheat and harbor most of gluten's antigenic epitopes). They tested the resulting peptide fractions against different gliadin-reactive T cell lines in the context of HLA-DQ2. Currently, more than

The authors are in the First Department of Medicine, University of Erlangen-Nuernberg, Ulmenweg 18, 91054 Erlangen, Germany. E-mail: detlef.schuppan@med1.imed.uni-erlangen.de