RNA BIOCHEMISTRY

Transcriptome-wide distribution and function of RNA hydroxymethylcytosine

Benjamin Delatte,[1]*† Fei Wang,[2]* Long Vo Ngoc,[3]*‡ Evelyne Collignon,[1] Elise Bonvin,[1] Rachel Deplus,[1] Emilie Calonne,[1] Bouchra Hassabi,[1] Pascale Putmans,[1] Stephan Awe,[4] Collin Wetzel,[5]§ Judith Kreher,[4]‖ Romuald Soin,[3] Catherine Creppe,[1] Patrick A. Limbach,[5] Cyril Gueydan,[3] Véronique Kruys,[3] Alexander Brehm,[4] Svetlana Minakhina,[2]¶ Matthieu Defrance,[1] Ruth Steward,[2] François Fuks[1]#

Hydroxymethylcytosine, well described in DNA, occurs also in RNA. Here, we show that hydroxymethylcytosine preferentially marks polyadenylated RNAs and is deposited by Tet in *Drosophila*. We map the transcriptome-wide hydroxymethylation landscape, revealing hydroxymethylcytosine in the transcripts of many genes, notably in coding sequences, and identify consensus sites for hydroxymethylation. We found that RNA hydroxymethylation can favor mRNA translation. Tet and hydroxymethylated RNA are found to be most abundant in the *Drosophila* brain, and Tet-deficient fruitflies suffer impaired brain development, accompanied by decreased RNA hydroxymethylation. This study highlights the distribution, localization, and function of cytosine hydroxymethylation and identifies central roles for this modification in *Drosophila*.

In DNA, vertebrate Tet methyldioxygenases (Tet1, Tet2, and Tet3) catalyze hydroxylation of 5-methylcytosine to 5-hydroxymethylcytosine (*1*–*3*). Tet also catalyzes the formation of hydroxymethylcytosine in RNA (referred to here as hmrC) (*4*, *5*). To date, however, the distribution, localization, and functional relevance of hmrC remain unknown.

In the present study, we have sought to provide a better understanding of hmrC. For ease of interpreting results, we analyzed hmrC in *Drosophila melanogaster* because (i) cytosine methylation in *Drosophila* DNA is either absent or very low, being restricted to specific cellular contexts (*6*, *7*), and (ii) we have found no evidence of DNA hydroxymethylation in this organism (fig. S1). To detect 5hmC RNA modification, we used an antibody raised against 5-hydroxymethylcytosine (*8*, *9*). To confirm that it will bind to hmrC, we performed dot blot experiments using in vitro transcribed templates containing either unmethylated, methylated, or hydroxymethylated cytosines (table S1). The antibody to 5hmC specifically detected 5hmC-containing RNA. In addition, detection of 5hmC was abolished after ribonuclease (RNase) A treatment (fig. S2A).

We detected hmrC in dot blot experiments on total RNA extracted from *Drosophila* S2 cells (Fig. 1A and fig. S2, B and C). Isolation of polyadenylated RNA from S2 cells followed by immunoblotting showed strong enrichment in hmrC signal as compared with that of total cellular RNA (Fig. 1B and fig. S2, D and E). No signal was detected in fractions enriched in small RNAs or ribosomal RNAs (Fig. 1C and fig. S3, A and B). *Drosophila* possesses only one conserved Tet ortholog, *CG43444* (dTet) (*10*, *11*). Depletion of dTet in S2 cells, by using RNA interference for dTet (dTet KD), revealed a 54% decrease in *dTet* transcripts, as compared with control cells (Fig. 1D). Dot blotting with antibody to 5hmC showed a similar decrease in hmrC—44%—upon dTet knockdown (Fig. 1D and fig. S3, C and D).

To map the hmrC modification landscape in a transcriptome-wide manner, we adapted a recently used method [methylated RNA immunoprecipitation followed by sequencing (MeRIP-seq)] (*12*, *13*), which we call hMeRIP-seq. This method involves immunoprecipitation of hmrC-containing RNA with the antibody to 5hmC followed by next-generation sequencing. hMeRIP-seq in S2 cells yielded 3058 significantly enriched regions ("hmrC peaks," $P < 10^{-10}$) within 1597 coding gene transcripts (fig. S4, A to C, and table S2). Examples of enrichment profiles are shown in Fig. 2A. Several key controls were performed so as to ensure the validity and stringency of our experimental approach: (i) Further bioinformatic analyses demonstrated that our hMeRIP-seq experiments did not merely coprecipitate abundant RNA fragments

[1]Laboratory of Cancer Epigenetics, Faculty of Medicine, ULB Cancer Research Center (U-CRC), Université Libre de Bruxelles (ULB), Brussels, Belgium. [2]Waksman Institute, Department of Molecular Biology and Biochemistry, Cancer Institute of New Jersey, Rutgers University, Piscataway, NJ, USA. [3]Laboratory of Molecular Biology of the Gene, Faculty of Sciences, Université Libre de Bruxelles, Gosselies, Belgium. [4]Institut für Molekularbiologie und Tumorforschung, Philipps-Universität Marburg, Marburg, Germany. [5]Department of Chemistry, University of Cincinnati, Cincinnati, OH, USA.
*These authors contributed equally to this work. †Present address: Division of Signaling and Gene Expression, La Jolla Institute for Allergy and Immunology, La Jolla, CA 92037, USA. ‡Present address: Section of Molecular Biology, University of California, San Diego, La Jolla, CA 92093, USA. §Present address: Department of Cancer Biology, University of Cincinnati, Cincinnati, OH, USA. ‖Present address: Roche Diagnostics, 68305 Mannheim, Germany. ¶Present address: Robert Wood Johnson Medical School, Department of Medicine, Rutgers University, Piscataway, NJ, USA.
#Corresponding author. E-mail: ffuks@ulb.ac.be

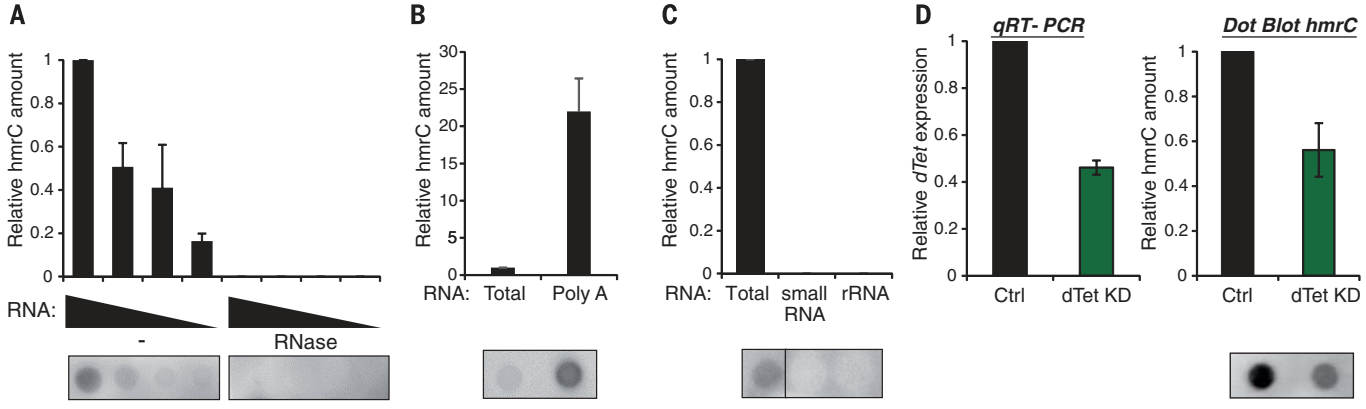

Fig. 1. RNA hydroxymethylation by dTet in *Drosophila* S2 cells. (**A**) Dot blotting on total RNA from *Drosophila* S2 cells with antibody to 5hmC, treated or not with RNase A (serially halved amounts of RNA, starting at 1 μg). Data are mean ± SD (n = 4 experiments run) with a representative blot shown. (**B**) Immunoblotting with anti-5hmC antibody was performed on polyadenylated and total RNA from S2 cells. Data are mean ± SD (n = 3 experiments run). (**C**) hmrC content of total RNA as well as fractions enriched in small RNA or rRNA was assessed by dot blotting. Data are mean ± SD (n = 3 experiments run). A vertical line indicates juxtaposition of lanes within the same blot, exposed for the same time. (**D**) dTet knockdown leads to reduced hmrC levels. (Left) Quantitative RT-PCR analysis. (Right) Dot blotting. Data are mean ± SD (n = 4 experiments run).

nonspecifically (Fig. 2B and fig. S5); (ii) up to 79.4% of hmrC sites showed significant reduction levels in hMeRIP-seq upon dTet depletion as compared with that of control S2 cells, with 85.5% of the sites showing a more than fourfold reduction in hmrC levels (Fig. 2C and fig. S6); and (iii) replicate hMeRIP-seq experiments by using an additional antibody to 5hmC showed strong agreement between experiments performed with the two 5hmC antibodies (fig. S7 and table S3).

The distribution of hmrC peaks revealed by hMeRIP-seq analyses was significantly nonrandom

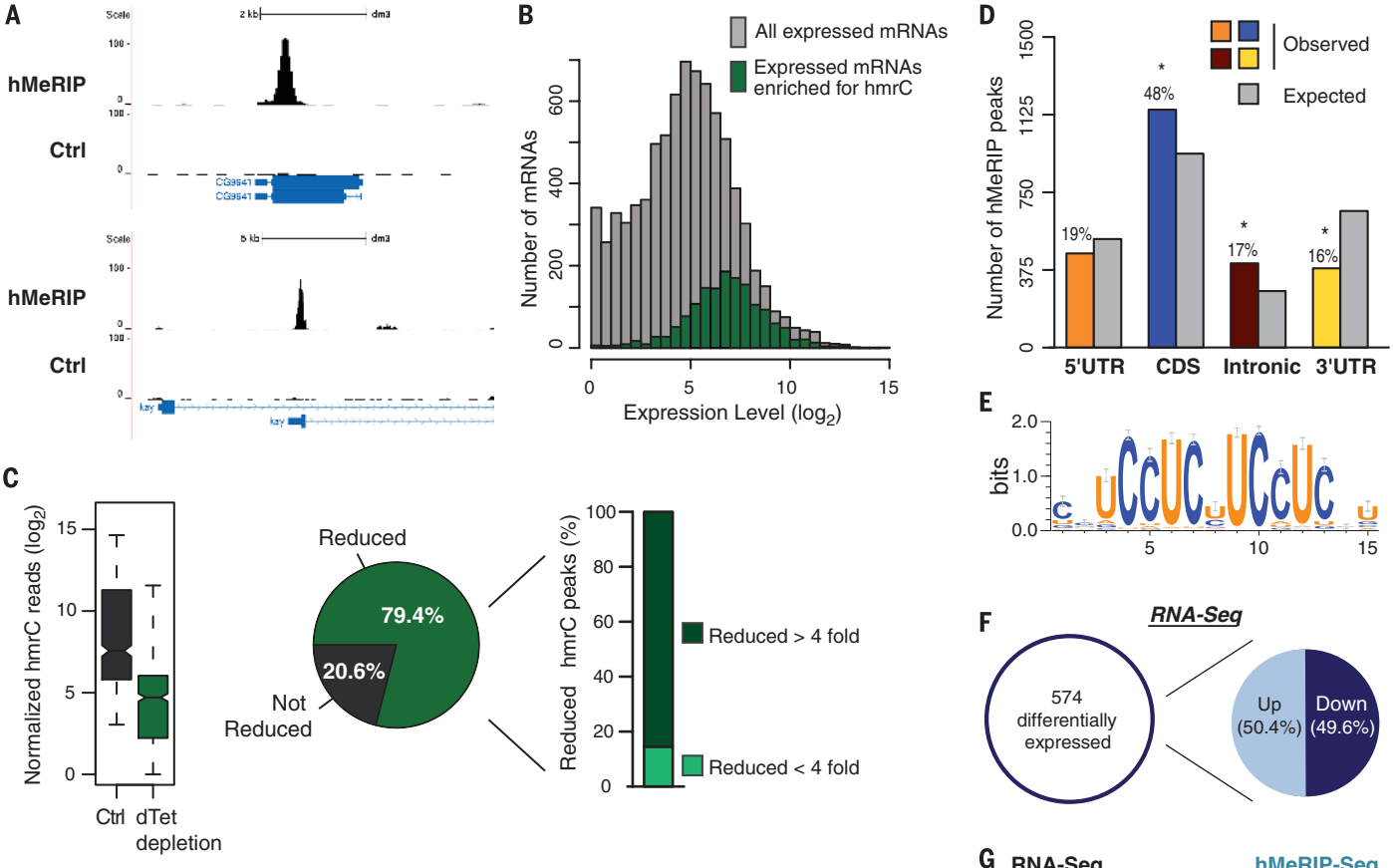

Fig. 2. Transcriptome-wide distribution of hmrC in *Drosophila* cells. (**A**) Representative UCSC Genome Browser plot from hMeRIP-seq data. (**B**) Distribution of all expressed (gray) or hmrC-enriched (green) transcripts, showing the number of mRNAs as a function of their expression levels. In hMeRIP-seq, enrichment in both abundant and less abundant fragments was observed. (**C**) hMeRIP-seq in cells depleted of dTet shows reduced hmrC levels at the majority of target regions. (Left) Box plot of the normalized number of hmrC reads in dTet-depleted cells versus control cells. (Center) Pie chart showing the percentage of reduced hmrC peaks, with (right) a more than fourfold reduction at most targets. (**D**) Distribution of hmrC peaks according to the type of structural element within the transcript. *$P < 10^{-5}$). (**E**) Sequence motif identified within hmrC peaks. (**F**) RNA-seq performed on dTet-knockdown and control S2 cells. (**G**) Partial overlap between hMeRIP-seq and RNA-seq data sets.

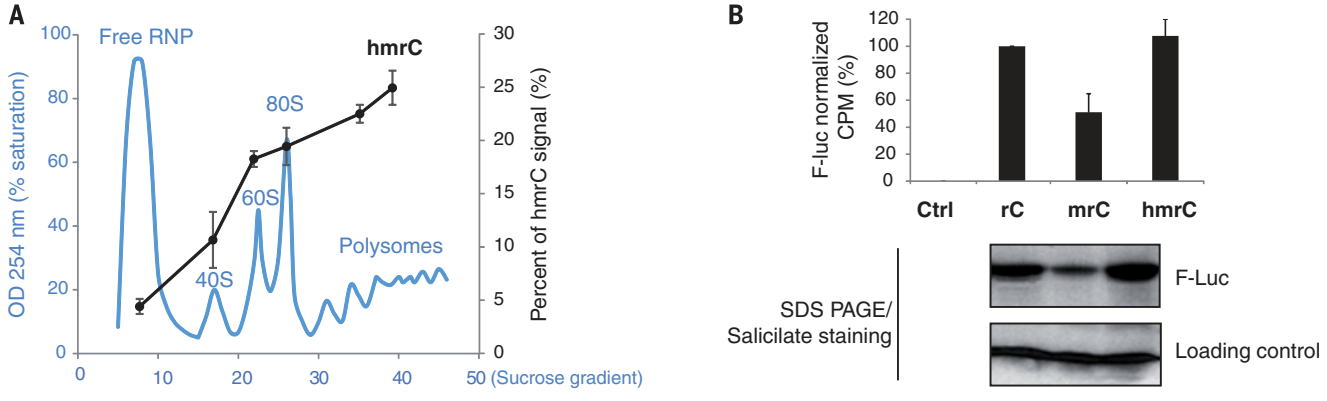

Fig. 3. RNA hydroxymethylation can favor mRNA translation. (**A**) Sucrose gradient fractionation followed by dot-blot quantification shows that active translation is associated with a high hmrC content. Data are mean ± SD (n = 3 experiments run). (**B**) In vitro translation of C-, mC-, and hmC-containing RNAs, as measured by incorporation of ^{35}S-radiolabeled methionine, shows that methylation decreases translation and hydroxymethylation restores it. Shown are the normalized scintillation counts (top) and the results of SDS–polyacrylamide gel electrophoresis (SDS-PAGE) followed by fluorography (bottom). Data are mean ± SD (n = 3 experiments run).

($P < 10^{-5}$), with many of these peaks found in coding sequences (48%) (Fig. 2D). Further analyses identified an overrepresented motif with occurrences in a large proportion of peaks (64%) and which tends to be highly UC-rich, containing UCCUC repeats (Fig. 2E and fig. S8A). Gene Ontology analysis of the hmrC targets showed enrichment for genes involved in basic cellular processes and notably in the regulation of embryogenesis and development (fig. S8B and table S6). In addition, exons were more enriched in hmrC than introns (fig. S9D). Thus, the above data present the hmrC-enriched transcriptome, revealing the presence of this modification in specific mRNA regions and in specific sequence contexts.

Next, we knocked down *dTet* expression in S2 cells and performed RNA-sequencing (RNA-seq) experiments in order to assess how many mRNAs might be regulated by dTet. We found the expression of 574 mRNAs to change after dTet depletion (50.4% were more abundant, whereas 49.6% were less abundant) (Fig. 2F; fig. S9, A to C; and table S4). Examples of mRNAs whose levels either increased or decreased upon dTet knockdown are shown in fig. S9, A and B. A positive correlation was observed between transcript abundance and the presence of hmrC peaks (Fig. 2B and fig. S5). We then compared these dTet-regulated mRNAs with the targets identified by hMeRIP-seq and found a slight but significant percentage (26%, $P < 10^{-43}$) of the dTet-regulated mRNAs to contain at least one hmrC peak (Fig. 2G, figs. S9C and S11, and table S5). It is worth mentioning that dTet contains an N-terminal CXXC Zn-finger domain, which likely explains the reported ability of mammalian Tets to regulate gene expression independently of their catalytic activity (*14*, *15*). It thus remains possible that dTet might affect gene expression via its CXXC domain, independently of its ability to hydroxylate methylated RNA (fig. S10, A and B). This domain, however, is likely not required for the role of dTet in *Drosophila* brain development (see below) because flies where dTet CXXC is deleted are still viable and show no specific phenotype (*11*).

To examine how cytosine hydroxymethylation might affect mRNA function, we examined the distribution of hmrC as a function of the mRNA translational status in *Drosophila* S2 cells. For this, we performed standard sucrose-gradient fractionation followed by dot blotting. A correlation between hmrC abundance and active mRNA translation was observed; fractions with low translation activity (free mRNAs and monosomes) were found to be poor in hmrC, whereas mRNAs heavily loaded with ribosomes (polysome fractions) showed a high hmrC content (Fig. 3A). We also assessed how mrC distributes across polysome fractions and found it to be high in monosomes and low in polysomes (fig. S12). Next, we examined whether mRNA hydroxymethylation might affect mRNA translation. In vitro translation of

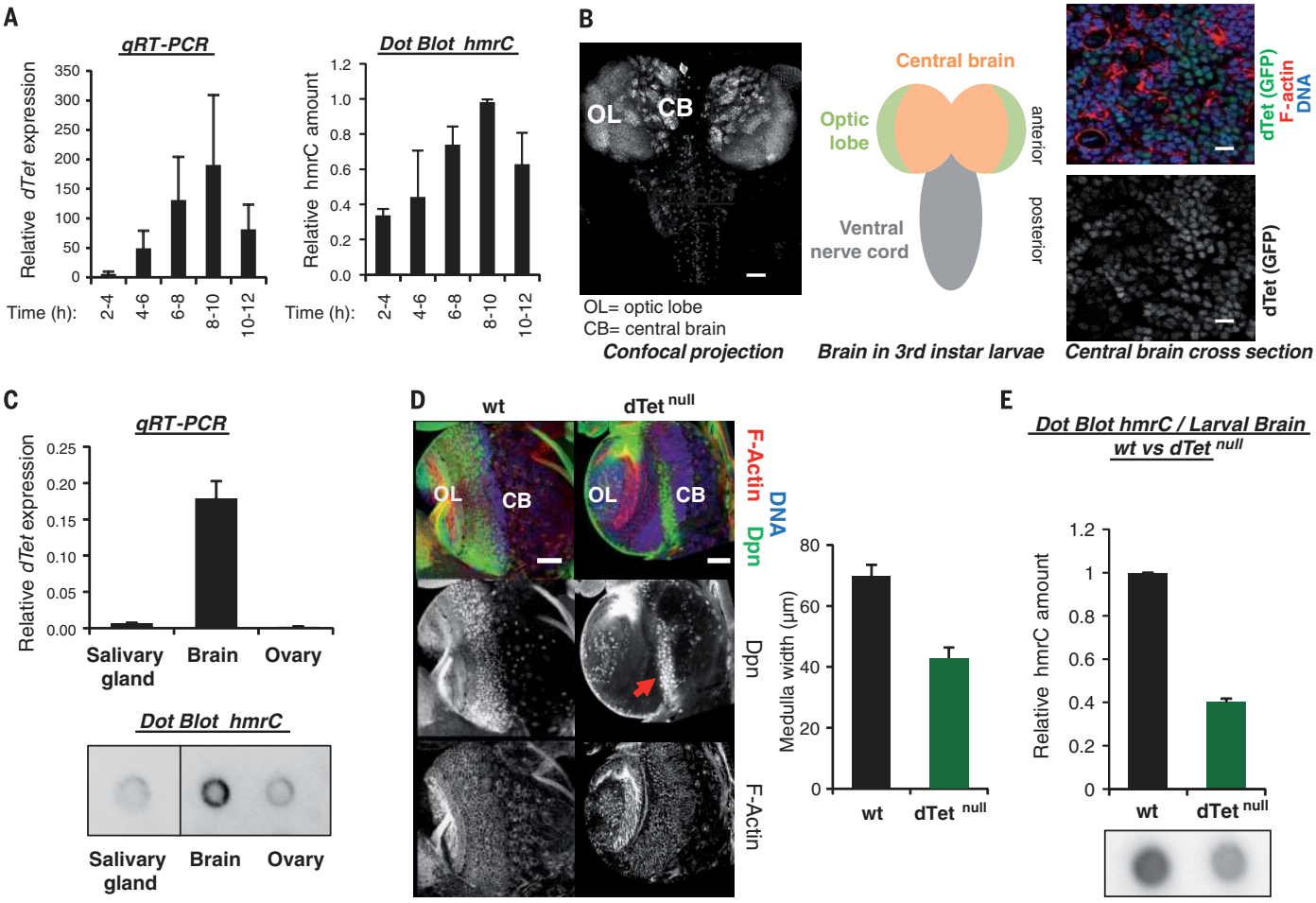

Fig. 4. dTet-deficient fruitflies show impaired brain development, accompanied by decreased RNA hydroxymethylation. (**A**) *dTet* expression and hmrC levels during *Drosophila* embryogenesis. Data are mean ± SD (n = 4 experiments run). (**B**) (Left) Pattern of endogenous GFP-tagged dTet in the larval brain. Scale bar, 50 μm. (Center) Scheme of the larval brain. (Right) Confocal brain section showing the expression of endogenous dTet (green) and F-actin (red). DNA, blue. Scale bar, 10 μm. (**C**) (Top) *dTet* expression in the salivary gland, brain, and ovary. Data are mean ± SD (n = 3 experiments run). (Bottom) Immunoblotting with 5hmC antibody in RNA from salivary gland, brain, and ovary. Vertical line indicates juxtaposition of lanes within the same blot, exposed for the same time. (**D**) (Left) $dTet^{null}$ brains are smaller than wild-type brains. Dpn (neuroblasts), green; F-actin, red; DNA, blue. Scale bar, 50 μm. (Right) Average width of the medulla region containing the neuroblasts (left, red arrow). Error bars represent 95% confidence intervals ($p < 2.4\ 10^{-9}$) for 20 brain lobes. (**E**) Brains of dTet-deficient larvae show a decrease in hmrC. Data are mean ± SD (n = 3 experiments run).

unmodified, methylated, and hydroxymethylated Firefly Luciferase–encoding RNA templates in rabbit reticulocyte lysate showed a decrease in translation of methylated RNAs. In contrast, hmrC-modified templates gave rise to near-control protein levels (Fig. 3B), suggesting that hydroxymethylation can restore the translation efficiency of previously methylated substrates.

We next sought to assess hmrC in vivo in fruitflies. To determine the timing of *dTet* expression and of the appearance of hmrC during early embryogenesis of *D. melanogaster*, we performed quantitative reverse transcription polymerase chain reaction (RT-PCR) to measure *dTet* transcript levels, and immunoblotting to estimate levels of hmrC. We found that *dTet* levels correlate positively with hmrC levels during fruitfly embryogenesis (Fig. 4A and fig. S13A). We also analyzed a publicly available database of RNA-seq results from different stages of *D. melanogaster* development and found that in third-instar larvae, *dTet* expression is highest in the central nervous system (fig. S13B). These findings suggest that dTet-mediated hydroxymethylation of RNA could play a role in the fruitfly brain. To confirm and extend these observations, we generated transgenic *Drosophila* flies expressing a *GFP-dTet* fusion construct under the control of the endogenous dTet promoter (called *dTet-Mi*). The green fluorescent protein (GFP)–tagged dTet protein was detected throughout the larval brain, the highest levels being detected in the optic lobe and central brain (Fig. 4B).

It was of interest to assess whether the fly brain contains high levels of hmrC. To this end, in addition to the brain, we used ovary [because this was the organ chosen to show the role of dTet in DNA m6A demethylation (*11*)]. We also used another organ, the salivary gland, from which we could extract enough RNA to measure hmrC levels. Quantitative RT-PCR and dot blotting showed higher dTet expression and hmrC content in the brain than in the ovary or in the salivary gland of the fruitfly (Fig. 4C). Hence, by revealing that the hmrC signal is highest in the brain, these data support the argument for the importance of hmrC in this organ.

We wished to evaluate RNA hydroxymethylation levels in a complete loss-of-function mutant of dTet (fig. S13, C and D). In agreement with recent observations (*11*), dTet-deficient animals survived through the larval stages but died at the pupal stage (no adult animals survived; $n > 5000$ dTet-null animals analyzed) (fig. S13E). Morphological defects were observed at larval stages: The brains of mutant larvae were smaller than those of normal larvae and showed abnormal organization of neuroblasts in their central part (Fig. 4D). To quantify the observed changes, we measured the width of the medulla region (based on 20 examined brains). We found it to be significantly reduced in $dTet^{null}$ animals ($P < 2.4 \times 10^{-9}$), likely reflecting a lower number of neuroblasts (Fig. 4D). To measure the effects of dTet loss on the RNA hydroxymethylation level, we performed immunoblotting analyses with antibody to 5hmC on brains from $dTet^{null}$ and wild-type larvae. These experiments showed a decreased hmrC content in the brains of dTet-deficient larvae (Fig. 4E and fig. S14).

Our understanding of the posttranscriptional modifications that decorate RNA, a regulatory layer positioned between DNA and proteins, is in its infancy. We have conducted a study addressing the distribution, localization, and function of cytosine hydroxymethylation in RNA, using *Drosophila melanogaster* as a model. Our work has yielded the following key findings: (i) It provides a picture of the hydroxymethylated transcriptome, (ii) reveals an unrecognized function for hmrC, and (iii) suggests a central role for this RNA modification and dTet in the *Drosophila* brain. All in all, we expect this study to change the way we think about the roles played by cytosine hydroxymethylation and the Tet proteins. Our findings open new research prospects in an emerging realm of biological regulation: epitranscriptomics.

REFERENCES AND NOTES

1. S. Kriaucionis, N. Heintz, *Science* **324**, 929–930 (2009).
2. M. Tahiliani *et al.*, *Science* **324**, 930–935 (2009).
3. R. M. Kohli, Y. Zhang, *Nature* **502**, 472–479 (2013).
4. I. Rácz, I. Király, D. Lásztily, *Planta* **142**, 263–267 (1978).
5. L. Fu *et al.*, *J. Am. Chem. Soc.* **136**, 11582–11585 (2014).
6. G. Raddatz *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 8627–8631 (2013).
7. S. Takayama *et al.*, *Genome Res.* **24**, 821–830 (2014).
8. E. M. Kallin *et al.*, *Mol. Cell.* **48**, 266–276 (2012).
9. H. Wu *et al.*, *Genes Dev.* **25**, 679–684 (2011).
10. T. L. Dunwell, L. J. McGuffin, J. M. Dunwell, G. P. Pfeifer, *Cell Cycle* **12**, 3357–3365 (2013).
11. G. Zhang *et al.*, *Cell* **161**, 893–906 (2015).
12. K. D. Meyer *et al.*, *Cell* **149**, 1635–1646 (2012).
13. D. Dominissini *et al.*, *Nature* **485**, 201–206 (2012).
14. R. Deplus *et al.*, *EMBO J.* **32**, 645–655 (2013).
15. K. Williams *et al.*, *Nature* **473**, 343–348 (2011).

ACKNOWLEDGMENTS

We thank L. Droogmans for advice, as well as A. Rao and J. Kadonaga for continued support. We also thank Y. N. Jan for providing the antibody to Dpn. R.D., B.H., C.C., and E.B. were supported by the Belgian Fonds de la Recherche Scientifique (FNRS). B.D. was supported by the FNRS (« Aspirant ») and a Belgian American Educational Foundation (BAEF) fellowship. L.V.N. was supported by the Belgian Fund for Research Training in Industry and Agriculture and a BAEF fellowship. E.Co. was supported by a "l'Oréal" fellowship, and M.D. was supported by the WBHealth grant (CanDx) from the Walloon Region. F.F. is a ULB Professor. F.F.'s lab was funded by grants from the FNRS and Télévie, the Interuniversity Attraction Poles (P7/03) program, the Action de Recherche Concertée (ARC) (AUWB-2010-2015 ULB-No 7), the WB Health program, the Belgian Fondation contre le Cancer, and the Fonds Gaston Ithier. F.W. was supported by a Charles and Johanna Busch graduate fellowship and the work at Rutgers by a NIH grant (RO1 GM089992) to R.S. V.K. was supported by the ARC (AV.12/17) as well as Brachet and Van Buuren Funds. P.A.L. received financial support from the National Science Foundation (NSF CHE 1212625 and 1507357). The authors declare no conflicts of interest. All sequencing data have been deposited to the Gene Expression Omnibus as series GSE66090.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/351/6270/282/suppl/DC1
Materials and Methods
Figs. S1 to S14
Tables S1 to S6
References (*16*–*35*)

8 May 2015; accepted 8 December 2015
10.1126/science.aac5253

GENE REGULATION

Transcription factors LRF and BCL11A independently repress expression of fetal hemoglobin

Takeshi Masuda,[1] Xin Wang,[2] Manami Maeda,[1] Matthew C. Canver,[3] Falak Sher,[3] Alister P. W. Funnell,[4] Chris Fisher,[5] Maria Suciu,[5] Gabriella E. Martyn,[4] Laura J. Norton,[4] Catherine Zhu,[1] Ryo Kurita,[6] Yukio Nakamura,[6,7] Jian Xu,[3,8] Douglas R. Higgs,[5] Merlin Crossley,[4] Daniel E. Bauer,[3] Stuart H. Orkin,[3,9] Peter V. Kharchenko,[2]* Takahiro Maeda[1]*

Genes encoding human β-type globin undergo a developmental switch from embryonic to fetal to adult-type expression. Mutations in the adult form cause inherited hemoglobinopathies or globin disorders, including sickle cell disease and thalassemia. Some experimental results have suggested that these diseases could be treated by induction of fetal-type hemoglobin (HbF). However, the mechanisms that repress HbF in adults remain unclear. We found that the LRF/ZBTB7A transcription factor occupies fetal γ-globin genes and maintains the nucleosome density necessary for γ-globin gene silencing in adults, and that LRF confers its repressive activity through a NuRD repressor complex independent of the fetal globin repressor BCL11A. Our study may provide additional opportunities for therapeutic targeting in the treatment of hemoglobinopathies.

During human development, the site of erythropoiesis changes from the embryonic yolk sac to the fetal liver and then, in newborns, to the bone marrow, where it persists through adulthood. Coincidentally, there is a "globin switch" from embryonic to fetal globin genes in utero, and then a second switch from fetal to adult globin expression soon after birth. This process has been studied for more than 60 years (*1*). The latter transition from fetal to adult