



Figures and figure supplements

Adenine methylation is very scarce in the *Drosophila* genome and not erased by the ten-eleven translocation dioxygenase

Manon Boulet et al.

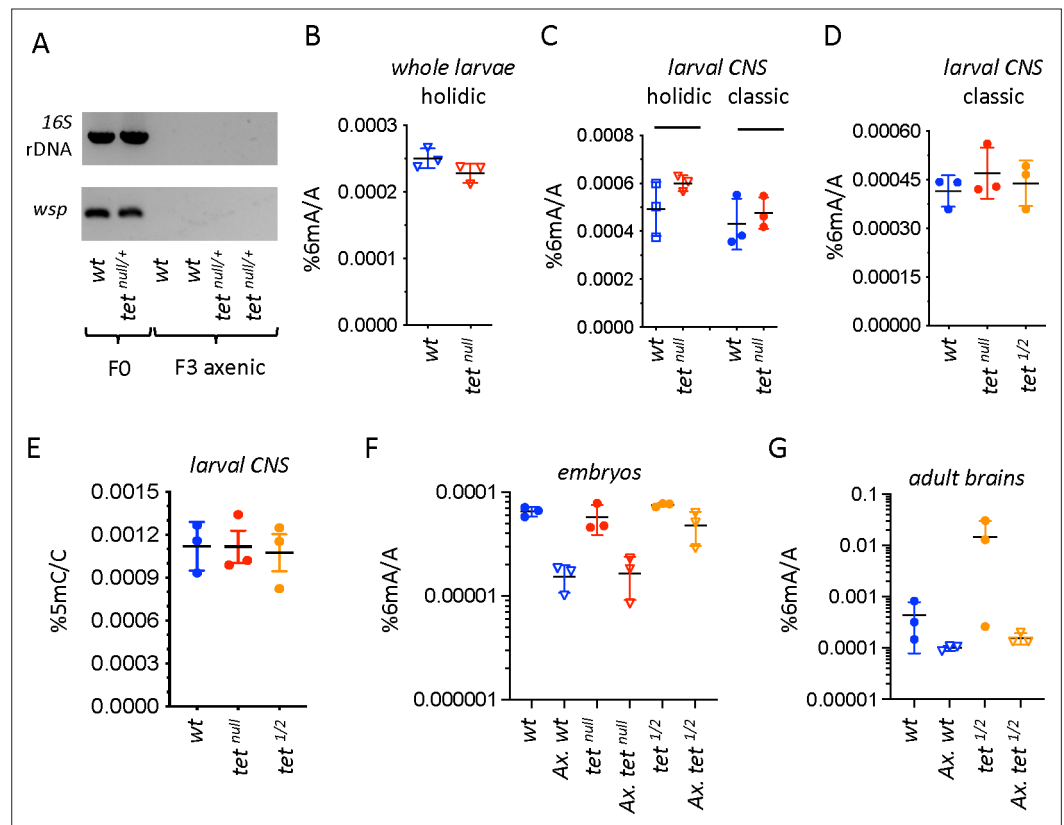


Figure 1. N6-methyladenine (6mA) levels in axenic *Drosophila* larvae are very low and not affected by ten-eleven translocation (TET) loss. **(A)** The presence of bacterial contaminations in wild-type (wt) and *tet^{null/+}* adult flies was checked by PCR using universal primers against bacterial 16 S rDNA and against the endosymbiotic bacteria gene *wsp*. PCRs were performed on DNA from parental (F0) flies and after three generations of breeding in axenic conditions (F3). **(B–D)** 6mA levels were measured by liquid chromatography with tandem mass spectrometry (LC-MS/MS) in genomic DNA (gDNA) from whole larvae **(B)** or dissected central nervous system (CNS) **(C, D)** generated from axenic flies reared on holidic medium **(B, C)** or conventional flies reared on classic medium **(C, D)**. **(E)** 5-methylcytosine (5mC) levels were measured by LC-MS/MS in gDNA from dissected CNS from flies reared on classic medium. **(F, G)** 6mA levels were measured by LC-MS/MS in embryos **(F)** and dissected adult brains **(G)** collected from crosses with conventional or axenic (Ax.) individuals raised on classic fly medium supplemented (Ax.) or not with antibiotics. wt: wild-type (*w¹¹¹⁸*); *tet^{null}*: *tet^{null/null}*; *tet^{1/2}*: *tet^{DMAD1/DMAD2}*. Filled circles: conventional flies; open triangles: axenic flies. Individual values, means and standard deviations are plotted. No statistically significant differences were observed between wt and tet mutant samples (Mann-Whitney test).

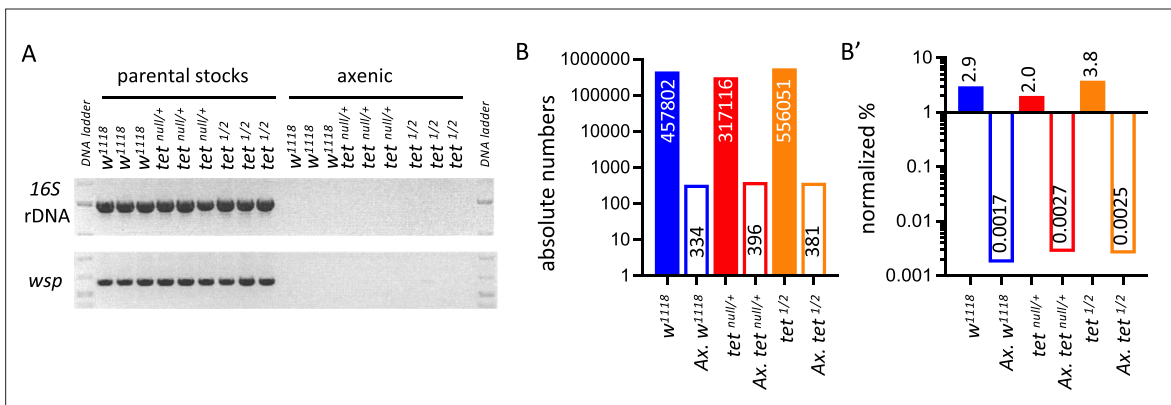


Figure 1—figure supplement 1. Detection of bacterial contamination in parental stocks and after three generations of breeding in axenic conditions. **(A)** The presence of bacterial contamination in adult flies of the indicated genotypes was checked by PCR using universal primers against bacterial 16S rDNA and against the endosymbiotic bacteria gene *wsp*. **(B, B')** Genome-wide sequencing was used to assess the presence of contamination in genomic DNA (gDNA) from adult flies of the indicated genotypes. Between 15 and 19 million reads were analyzed per sample. The absolute numbers of reads mapping to bacteria, virus, or fungi genomes are represented in B, and their proportions normalized to the total number of reads are presented in B'. *tet^{1/2}*: *tet^{DMAD1/DMAD2}*.

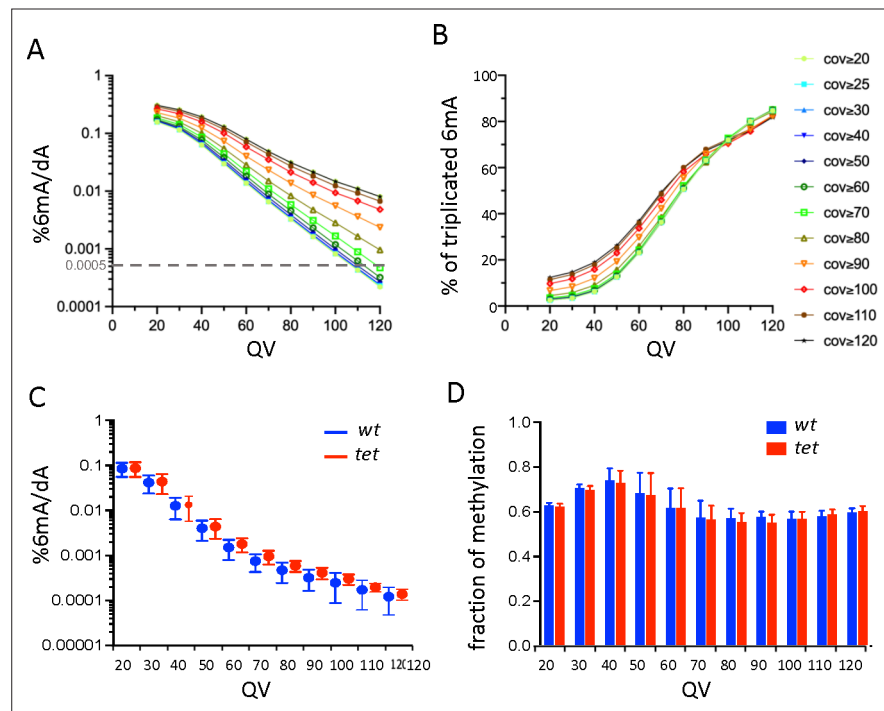


Figure 2. Single-molecule real-time sequencing (SMRT-seq) analysis of larval central nervous system (CNS) genomic DNA (gDNA) does not reveal an increase in N6-methyladenine (6mA) in the absence of ten-eleven translocation (TET). **(A)** Percentage of adenines identified as 6mA in the wild-type fusion dataset depending on the quality value (QV) and coverage values used for 6mA selection. The dashed gray line indicates the level of 6mA measured by liquid chromatography with tandem mass spectrometry (LC-MS/MS) (0.0005%). **(B)** Influence of the QV and the coverage values on the proportion of 6mA identified in the wild-type fusion dataset and in the three original samples. **(C)** Percentage of adenines covered at least 25x and identified as 6mA in each of the three wild-type (*wt*) or *tet null* (*tet*) datasets depending on the QV. Means and standard deviations are represented. **(D)** Fraction of methylation in *wt* or *tet null* datasets depending on the QV (coverage ≥ 25 x). Means and standard deviations are represented.

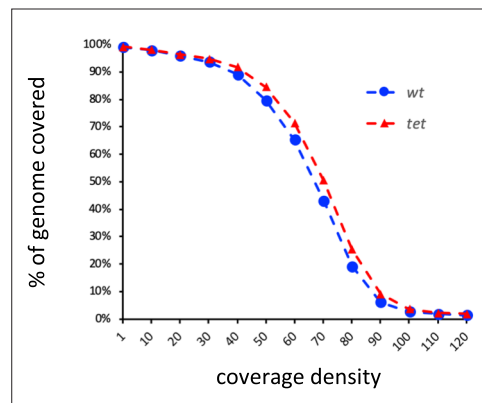


Figure 2—figure supplement 1. Proportion of the *Drosophila* genome covered in wild-type (wt) and tet^{null} (*tet*) single-molecule real-time sequencing (SMRT-seq) fusion datasets according to coverage density.

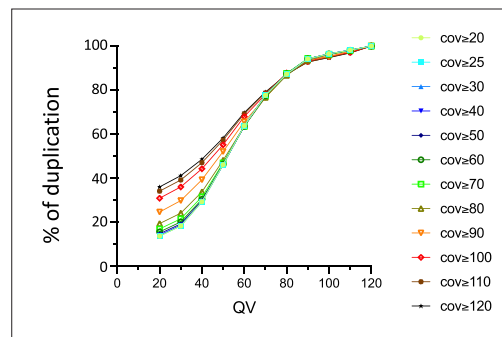


Figure 2—figure supplement 2. Influence of the quality value (QV) and the coverage values on the proportion of N6-methyladenine (6mA) identified in the wild-type (wt) fusion dataset and in the three original samples. The percentage of 6mA identified in at least 2 out of 3 replicates is shown.

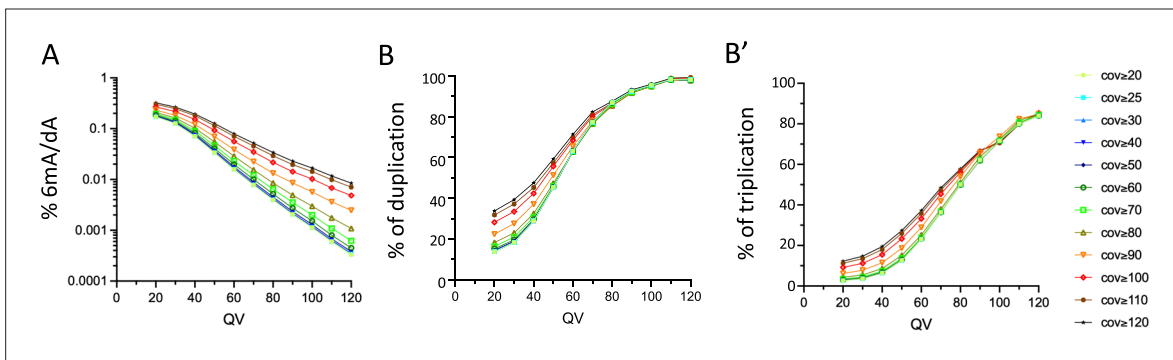


Figure 2—figure supplement 3. Influence of the quality and coverage values on SMRT-seq based detection of 6mA in the CNS of *tet^{null}* larvae. **(A)** Percentage of all adenines identified as N6-methyladenine (6mA) by SMRT-seq in the *tet^{null}* fusion dataset depending on the quality value (QV) and coverage values used for 6mA selection (log10 scale). **(B, B')** Influence of the QV and the coverage values on the proportion of 6mA identified in the *tet^{null}* fusion dataset and in the three original samples. **(B)** percentage of 6mA identified in at least 2 out of 3 samples; **(B')** percentage of 6mA identified in all three samples.

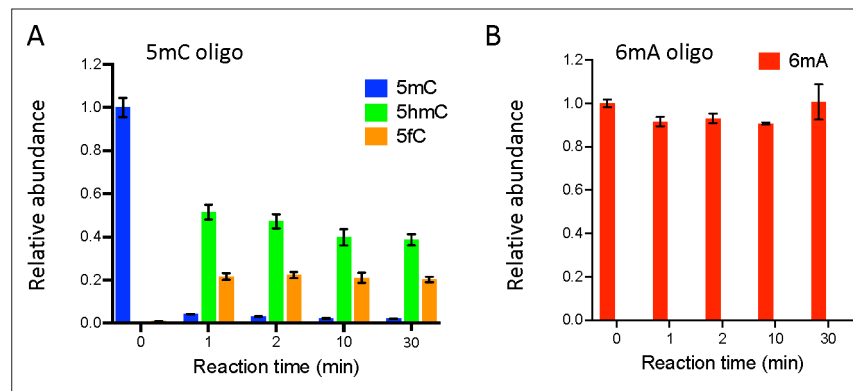


Figure 3. Ten-eleven translocation (TET) does not oxidize N6-methyladenine (6mA). **(A, B)** In vitro assays showing TET activity profile on 5-methylcytosine (5mC) **(A)** and 6mA **(B)** containing double-stranded oligonucleotide substrates. The levels of 5mC and its oxidized products (5-hydroxymethylcytosine: 5hmC and 5fC) are represented relative to the 5mC level at $t=0$. The levels of 6mA are represented relative to 6mA level at $t=0$. Error bars denote standard deviations from three independent experiments.

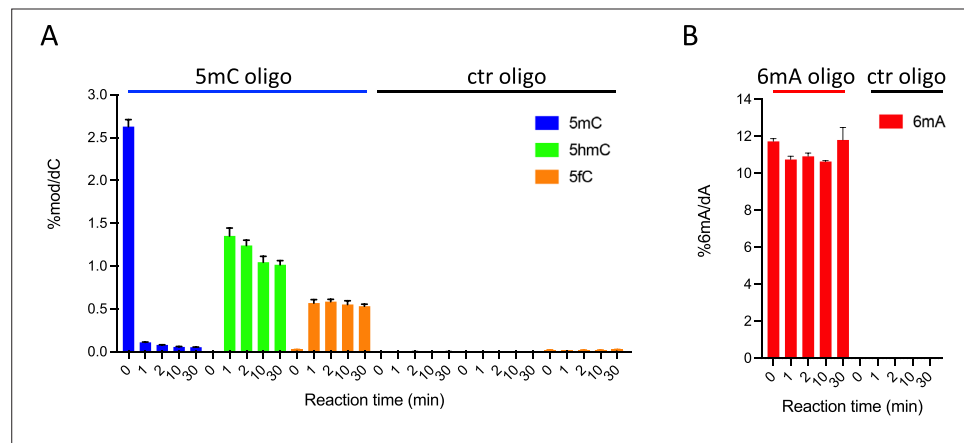


Figure 3—figure supplement 1. In vitro assays showing ten-eleven translocation (TET) activity profile on double-stranded oligonucleotide substrates containing or not 5-methylcytosine (5mC) (A) or N6-methyladenine (6mA) (B). The levels of 5mC, 5-hydroxymethylcytosine (5hmC), and 5fC normalized to dC (A) or 6mA normalized to dA (B) are represented. Error bars denote standard deviations from three independent experiments. Only background levels of modified nucleosides were detected when purified recombinant TET catalytic domain was incubated with unmodified (ctr) oligonucleotides.

ccTET	318	IVNRATPIHRETS* [*] IEGMDILVTGGN	344
JBP1	205	NERFR [*] TASHTDNGDFDNGYGLAVLKG	230
dTET	1939	CLDFCAHSHRDLHNM [*] QDGCTVHVALLK	1965
hTET1	1664	CLDFCAHPH [*] RDIHMNNGSTVVCTLR	1690
hTET2	1374	CLDFCAH [*] AHRDLHNM [*] QGSTLVCTLR	1400
hTET3	1069	CMDFCAH [*] AHKDQHNLYNGCTVVCTLR	1095

Figure 3—figure supplement 2. Multiple sequence alignment of ten-eleven translocation (TET)/JBP family members. The sequence surrounding the HxD motif in *Coprinopsis cinerea* TET (ccTET), *Trypanosoma brucei* JBP1, *Drosophila melanogaster* TET (dTET), and *Homo sapiens* TET1, TET2, TET3 is shown. Conserved amino acids between TET homologs are boxed in yellow. The two key amino acids required for N⁶-methyladenine (6mA) oxidation by ccTET are labeled with a star.

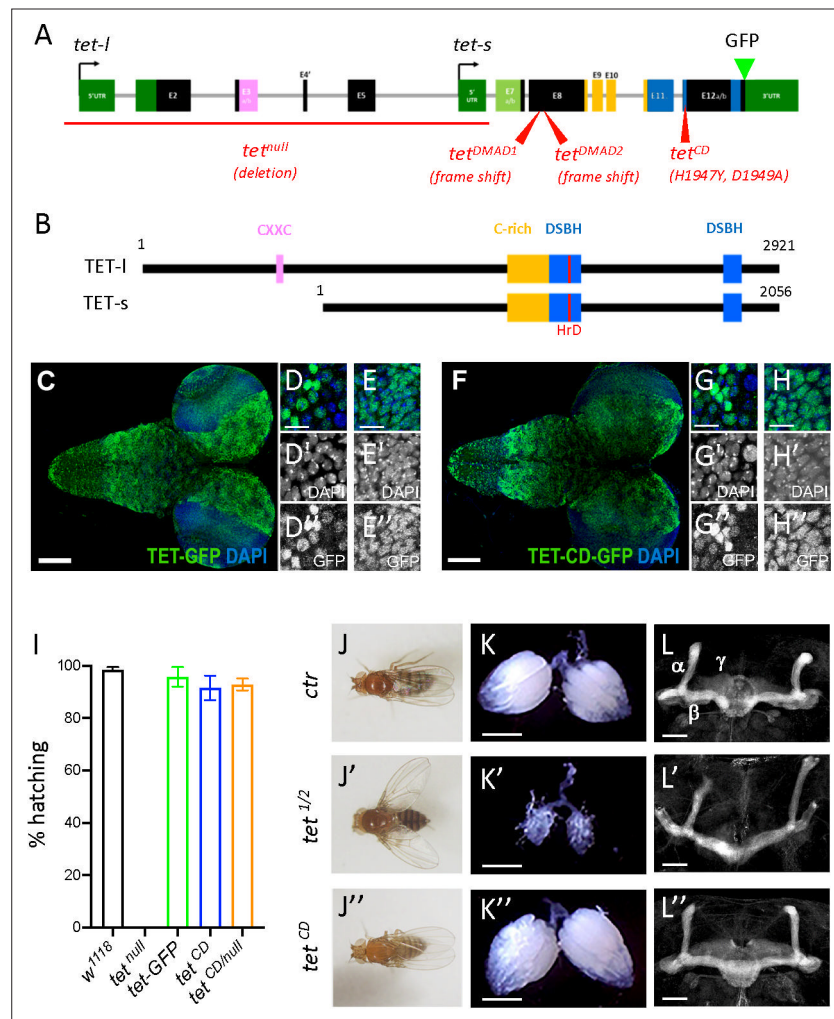


Figure 4. Ten-eleven translocation (TET) catalytic activity is largely dispensable in *Drosophila*. **(A, B)** Schematic representation of *tet* locus **(A)** and main protein isoforms **(B)**. **(A)** *tet* is transcribed from two alternative promoters giving rise to *tet-long* (*tet-l*) and *tet-short* (*tet-s*) isoforms. Filled boxes represent exons; non-coding exons (UTR) are depicted in green, and coding exons in black or according to their domain-associated color. Introns are represented as gray lines (not to scale). The *tet null*, *DMAD1*, *DMAD2*, and *catalytic dead* (*CD*) alleles are depicted in red. The location of the GFP insertion generated by CRISRP/Cas9-mediated knock-in is also indicated. **(B)** The conserved domains of TET are colored; pink: CXXC DNA binding domain, orange: Cystein-rich domain, blue: double-stranded β helix (DSBH) domain, red: HxD (iron binding motif). Amino acid positions are indicated according to the longest TET-I and TET-s isoforms. **(C–H)** Expression pattern of the wild-type and catalytic-dead versions of TET in the larval central nervous system (CNS). *tet-GFP* **(C–E)** and *tet^{CD}-GFP* **(F–H)** knock-in lines were used to detect TET proteins by confocal imaging after immunostaining against GFP (green). Nuclei were labeled with DAPI (blue). **(C, F)**: stitched images showing dorsal views of the entire CNS. Scale bar: 100 μ m. **(D, E, G, H)**: high-magnification views of TET expression in the ventral nerve cord **(D, G)** or the central brain **(E, H)**. DAPI-only and GFP-only channels are presented in the middle (') and lower (") panels, respectively. Scale bar: 10 μ m. **(i)** Percentage of adult flies of the indicated genotypes hatching from their pupal case. Means and standard deviations from four independent experiments. **(J–L)** Wing positioning **(J)**, ovaries **(K)**, and mushroom bodies **(L)** of wild-type adult flies (*ctr*: *tet-GFP*) as compared to flies lacking TET expression (*tet^{1/2}*: *tet^{DMAD1/DMAD2}* adult escapers) or TET enzymatic activity (*tet^{CD}*). **(L–L'')** Immunostaining against Fas2 on adult brains was used to label mushroom body α , β , and γ lobes. **(K–K'')** Scale bar 500 μ m. **(L–L'')** Scale bar 50 μ m.

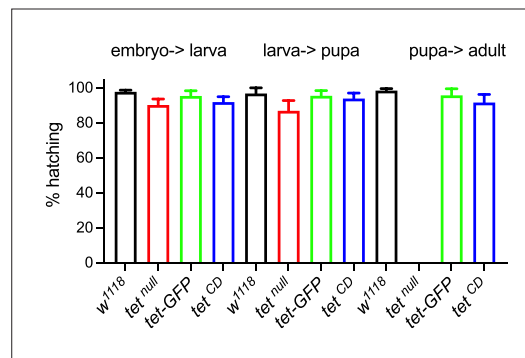


Figure 4—figure supplement 1. Survival assays shows the percentage of hatching embryos, larvae, and adults of the indicated genotypes. Means and standard deviations are from at least six independent experiments.