Supplemental figure captions:

Fig. S1. Fiber-seq-derived ACRs show expected patterns at ATAC ACRs and transcription start sites and expected correlation with expression. (A) Schematic illustrating the calculation of the FIRE accessibility score, a measure of Fiber-seq-derived chromatin accessibility that allows direct comparisons to ATAC-seq-derived chromatin accessibility. Shown are screenshots of two FIRE ACRs with individual fibers showing FIRE elements (red) of different length, in addition to methyltransferase-sensitive patched in purple and non-methylated regions as grey lines. Black boxes mark the respective FIRE ACRs (black bars on top). For any given window, the FIRE accessibility score is calculated as the number of bases annotated as FIRE elements (red) divided by the total number of bases across all fibers mapping within this window (red, purple for methyltransferase-sensitive patches not annotated as FIRE elements, grey for not methylated). FIRE accessibility scores are shown for the two example ACRs. **(B)** Correlation between FIRE accessibility scores for Fiber-seq replicates 1 and 2. Each dot corresponds to ACRs where both replicates have >10x coverage. **(C)** Correlation of Tn5 insertions for union ACRs identified in ATAC-seq replicates 1 and 2. **(D)** m6A methylation peaked at the center of ATAC-seq derived ACRs in paired samples. **(E)** m6A methylation rate peaked immediately upstream of CAGE-defined transcription start sites (TSSs), with phased nucleosomes apparent downstream of TSSs. Average strength of m6A methylation rate upstream of TSSs was monotonically related to expression level of respective downstream genes (expression deciles). The well-phased-nucleosome signal was strongest for highly expressed genes and faded for lowly expressed ones, as expected. (**F**) Methyltransferase-sensitive patches (MSPs) larger than 100 bp constituted the majority of the m6A signal at TSSs, while MSPs shorter than 100 bp showed patterns consistent with well-positioned nucleosomes. MSP scores were calculated in aggregate for each non-overlapping 20 bp window in the region 750 bp upstream and 1 kb downstream of each TSS (see Methods). (**G**) Aggregate plot of Tn5 insertions/base in the 1 kb window upstream of TSSs stratified by downstream gene expression for paired ATAC-seq data, comparable to (**E**). (FIRE elements supporting FIRE ACRs are in red, purple indicates methylation sensitive patches (see **Fig.1**). (**H**) Aggregate plot of FIRE ACRs stratified into ten deciles based on their FIRE accessibility score. For each FIRE score accessibility decile, the number of Tn5 insertions at each bp within 2 kb of the FIRE ACR center is shown. Accessibility measured by ATAC-seq and Fiber-seq is monotonically correlated. Highly accessible FIRE ACRs tend to show neighboring FIRE ACRs (symmetric signal at highest decile). This signal is in part due to FIRE ACRs in low-mappability LTR retrotransposons (see **Fig. 2**). (**I**) The single-molecule method Fiber-seq outperforms single-cell ATAC-seq as a quantitative measure of chromatin accessibility. 39,132 ACRs were identified as shared FIRE ACRs in dark-grown maize leaves and ATAC ACRs in a pseudobulked leaf sample (GSM4696890) from Marand et al 2021 (*19*). The percentage of cells containing at least one Tn5 insertion within a shared ACR (% cells accessible) is compared to the percentage of actuated fibers (*i.e.*, with a called FIRE element, % actuated Fibers within a given ATAC ACR) underlying the same shared ACR. Each dot represents one shared ACR. Hexbin color reflects the number of dots.

Fig. S2. Novel FIRE ACRs comprised of short FIRE elements are bona fide regulatory elements. (A) Schematic describing short-read simulation and mappability calculation. We generated 2.1 billion fragments evenly distributed across the B73 reference genome chromosomes 1-10 (see Methods). For each simulated fragment, 50 bp paired-end reads were generated (indicated with thick black arrows). Each read matched exactly the reference sequence from which it was generated. These simulated reads were then mapped back to the genome using BWA. The 'fraction mapped' for a given region or window was calculated as the number of correctly mapped reads with mapq score > 0 divided by the total number of simulated reads with the outer end (Tn5 insertion) falling in the region. Mapq scores are indicated by blue and red boxes, incorrectly mapped simulated read shows X in red box (top row). Mappability of regions was determined as percentage of correctly mapped reads with mapq>0. (**B**) Histograms of

mappability as in (**A)** for all 21,318,473 non-overlapping 100 bp windows in the maize genome (top panel, grey), 51,817 ATAC ACRs (middle panel, gold), and 106,867 FIRE ACRs (bottom panel, purple). Low mappability explains only in part why Fiber-seq detects many more ACRs than ATAC-seq. (**C**) FIRE ACRs comprised of short FIRE elements are not detected by ATAC-seq. Correlation between FIRE accessibility scores and Tn5 insertions/ base (chromatin accessibility as measured by ATAC-seq) for FIRE ACRs comprised of FIRE elements of indicated length (see inset for legend). **Left**, LOWESS curves fitted to FIRE ACRs in respective length categories. **Right**, plots showing individual values for FIRE ACRs belonging to the five length categories. (**D**) FIRE accessibility score by Tn5 insertions/base (*i.e.*, ATAC accessibility score) for ACRs stratified into 12 categories. Each dot represents an ACR with the labeled row and column properties. As the row categories are overlapping, ACRs were sorted hierarchically as follows: all ACRs with low FIRE accessibility score were included in the 'low FIRE acc. score' rows; ACRs with FE length < 200 bp and high FIRE accessibility score were included in the 'FE length <200' rows; ACRs with mappability < 80% and both high FIRE accessibility score and FE length >=200 bp were included in the 'Unmappable' rows. (**E**) FIRE ACRs that do not overlap with ATAC ACRs show similar patterns of the m6A signal (top) and the 5mCpG signal (bottom) as FIRE ACRs that overlap with ATAC ACRs. Shifted control regions do not display these properties. FIRE element length underlying FIRE ACRs is indicated as in (**C**). (**F**) FIRE ACRs that do not overlap with ATAC ACRs show a similar distribution across genomic compartments as FIRE ACRs that overlap with ATAC ACRs.

Fig. S3. Features of FIRE ACRs within LTR retrotransposons. (**A**) Histograms for the percentage actuation (*i.e.*, the percentage of fibers with a FIRE element that comprise a FIRE ACR) for the first of two paired ACRs (putative enhancers), the second of two paired ACRs (putative promoters), and single ACRs. (**B**) Phylogeny of LTR ACRs. Branch length units are in estimated substitutions per site (*59*). Colors indicate ACR types. (**C**) Fraction of all human FIRE ACRs and human FIRE ACRs with high 5mCpG methylation (mean CpG methylation of over 50%) that overlap an annotated repeat by more than 50 bp. FIRE ACRs calls from human cell line GM12878 (*17*).

Fig S4. Examples of ACRs in intact polymorphic LTR retrotransposons (A, B) and LTR retrotransposons with non-TE internal genes (C, D). (**A**) Left, intact LTR retrotransposon with blue LTRs is absent in NAM lines: Il14H, Ki3, M37W, P39. Tracks in screenshot as in **Fig. 2**. Right, expression level of indicated gene in lines with and without the TE, B73 is labeled in yellow. (**B**) Left, intact retrotransposon with blue LTRs is absent in NAM lines: B97, CML228, CML52, Ki11, Ky21, Mo18W, P39. Tracks in screenshot as in **Fig. 2**. Right, expression level of indicated gene in lines with and without the TE, B73 is labeled in yellow. (**C**) Example of an intact LTR retrotransposon containing one annotated gene between the LTRs and lacking an ACR at the transcription start site. (**D**) Example of an intact LTR retrotransposon containing two annotated genes. For each gene, transcription begins at a FIRE ACR within the LTR.

Fig. S5. Examples of solo LTRs containing FIRE ACRs

(A-C) Solo LTRs containing FIRE ACR are colored blue. (**A**) [chr01:60,920,594-60,935,475] **(B)** [chr01:179,120,635-179,131,399] **(C)** [chr01:207,732,409-207,748,141]. See **table S9** for a comprehensive list.

Fig. S6. Diffuse chromatin accessibility and hypo-5mCpG methylation is observed at loci discovered as hAT TE insertion sites by McClintock (*5*)**. (A)** *waxy1* (Zm00001eb378140; chr09:25,127,146 - 25,129,800), one of the first genes identified by McClintock as having a hAT TE insertion, shows higher gene-body chromatin accessibility than 84.4% of other genes. McClintock identified alleles *Ds wx-m9*, *Ds wx-m6*, *Ac wx-m9*, with the *Ds* or *Ac* prefix indicating whether it was a nonautonomous or autonomous hAT TE, respectively. **(b)** *bronze1* **(**Zm00001eb374230; chr09:13,118,806-13,123,664), one of the first genes identified by McClintock as having a hAT TE insertion. McClintock identified the *Ac bz-m2* allele. The *Ac* prefix indicates insertion of an autonomous hAT TE. **(c)** *shrunken* (Zm00001eb374090; chr09:12,836,508-12,845,499), one of the first genes identified by McClintock as having a hAT TE insertion. McClintock identified two germinally-stable alleles, *Ds-4864A* and *Ds-5245*, that were “genetically indistinguishable and located just distal to the Shrunken (Sh) locus on the short arm of chromosome 9” and three germinally-unstable alleles, *sh-m6233*, *sh-m5933*, *sh-m6258*, that contain rearrangements at the Sh locus related to a hAT insertion, one of which contains a Ds-mediated 30 kb insertion (Courage-Tebbe et al., 1983). The *Ds* prefix indicates insertion of a nonautonomous hAT TE.

Supplemental Tables

Table S1. FIRE ACRs

TableS2. ATAC ACRs

Table S3. chrPt_mappedtonuc.tsv

142,724 pairs of 50 bp reads were simulated from the 142,724 bp length plastid genome, generating 100x coverage. These reads were aligned to a fasta file consisting of the ten nuclear chromosomes, the mitochondrial genome, and the plastid genome. The resulting bam (alignment) file was converted to a bed file (bedtools bamtobed) and overlapping lines were merged (bedops -m).

Table S4. chrMt_mappedtonuc.tsv

579,124 pairs of 50 bp reads were simulated from the 579,124 bp length mitochondrial genome, generating 100x coverage. These reads were aligned to a fasta file consisting of the ten nuclear chromosomes, the mitochondrial genome, and the plastid genome. The resulting bam (alignment) file was converted to a bed file (bedtools bamtobed) and overlapping lines were merged (bedops -m).

Table S5. Cell actuation vs FIBER actuation

Percent of cells containing at least one Tn5 within this peak vs percent of fibers containing at least one FIRE element overlapping peak.

Table S6. Union ATAC ACRs

Table S7. Differential ACRs (dACRs)

ATAC ACRs lacking ATAC signal in etiolated leaf protoplasts.

Table S8. All intact LTR RTs, with the number of FIRE ACRs contained within the long terminal repeats and the strand information indicated.

Table S9. Solo LTRs containing FIRE ACRs

Table S10. Motifs enriched in first-of-two paired ACRs vs second-of-two paired ACRs

Table S11. Motifs enriched in second-of-two paired ACRs vs first-of-two paired ACRs

Table S12. Motifs enriched in single ACR vs first-of-two paired ACRs (putative enhancer)

Table S13. Motifs enriched in single vs second-of-two paired ACRs (putative promoter)

Table S14. FIRE accessibility score and mean 5mCpG-methylation percentage for all FIRE ACRs within LTRs.

Table S15. CompGenes_to_Zm00001eb318460

Table S16. B73 coordinates of hAT insertions in exactly one of the other 25 NAM lines and FIRE accessibility scores.

Figure S1

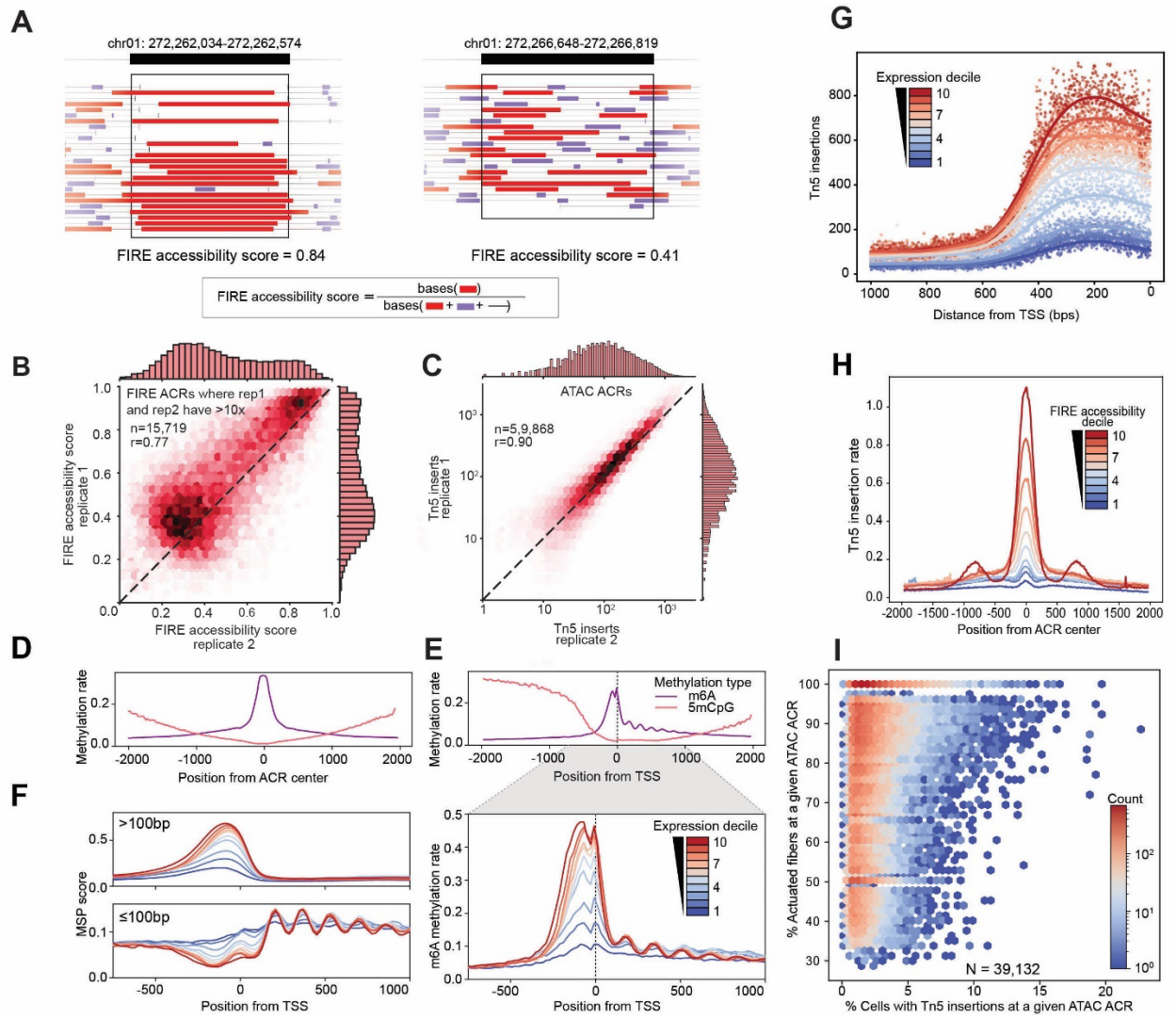

Fig. S2

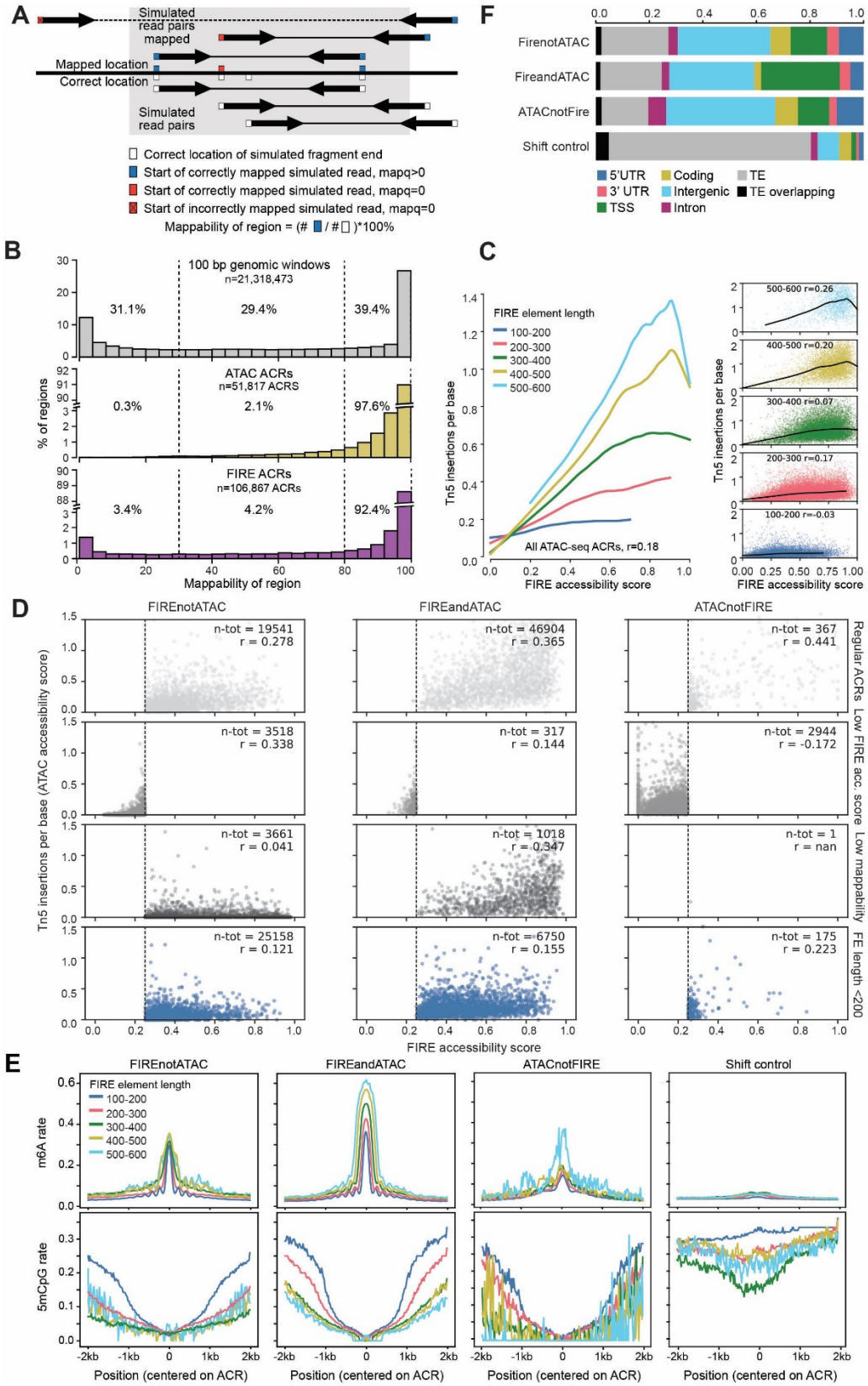

Fig. S3

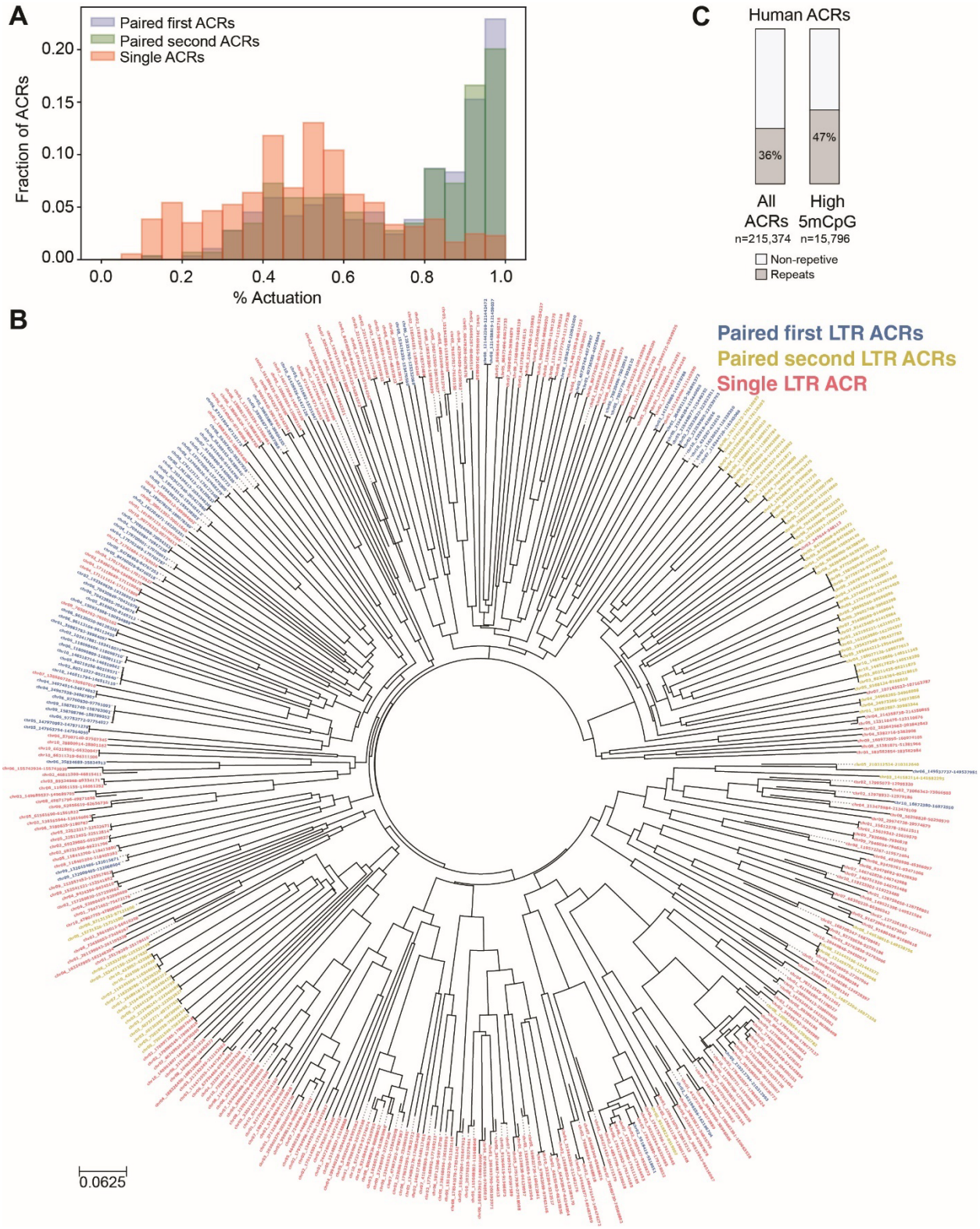

Fig. S4

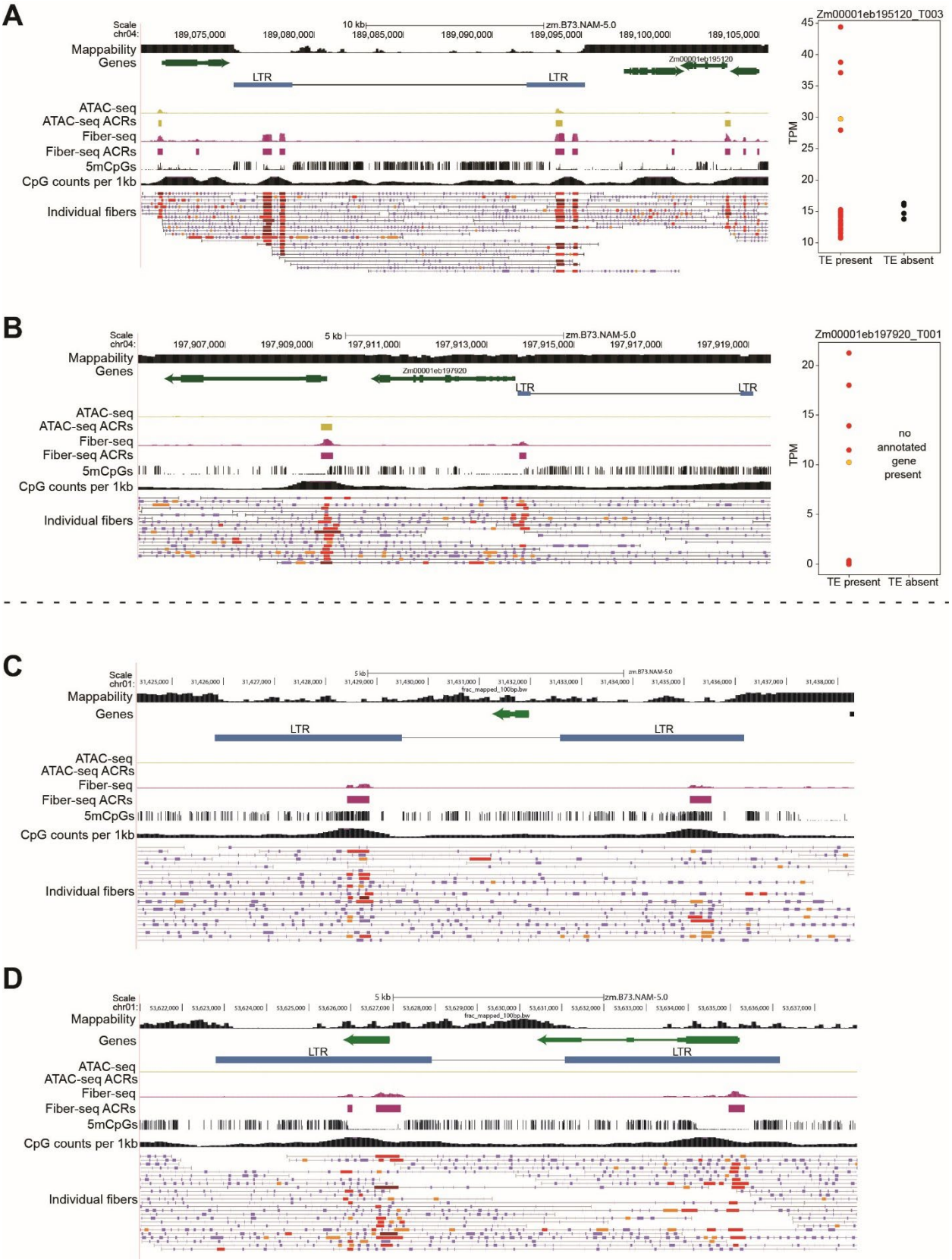

Fig. S5

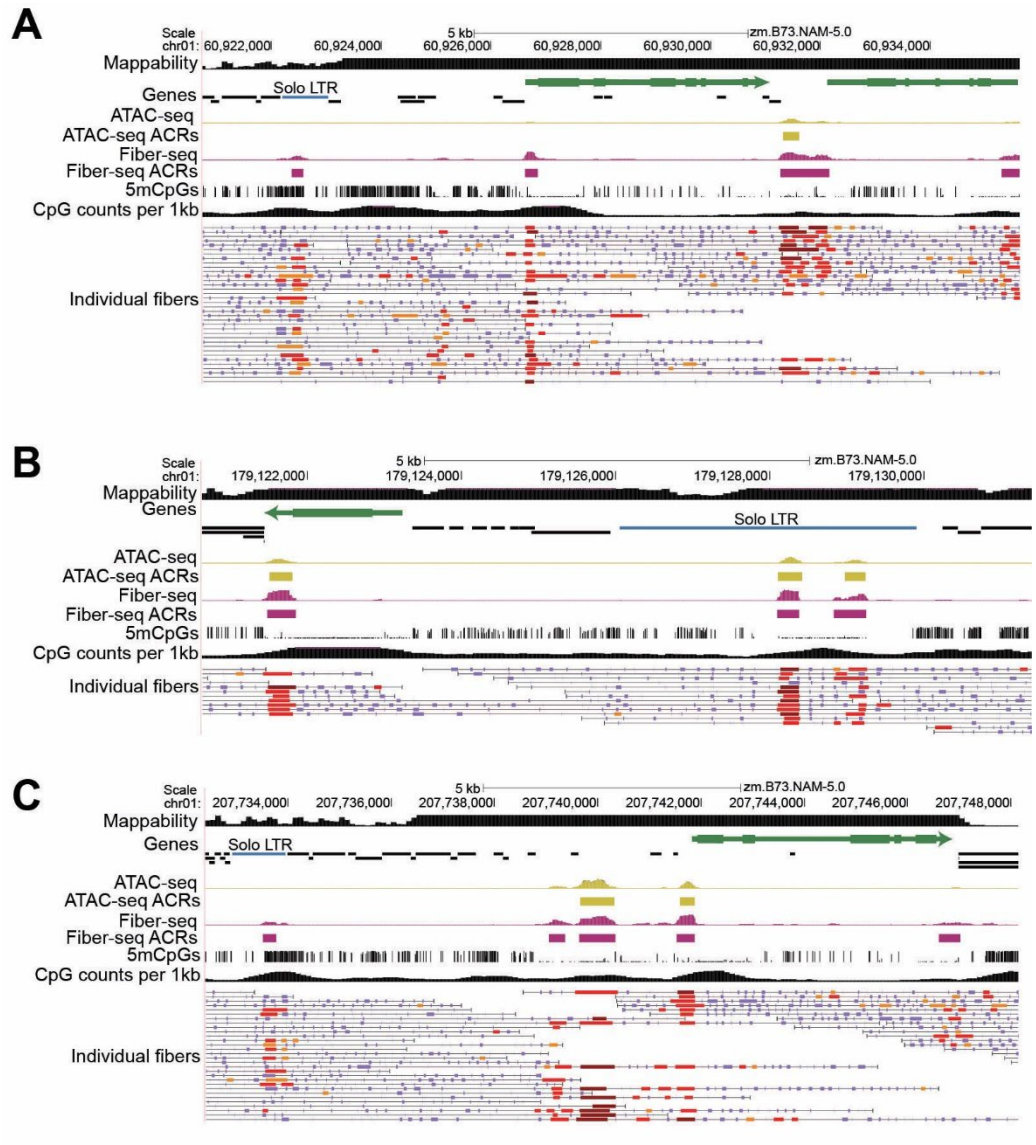

Fig. S6