Supplementary Figure 1: Positional distribution of CAGE tags around DNase I hypersensitive sites

a, The cumulative fraction (vertical axis) of capped RNA 5’ ends, as measured by Cap Analysis of Gene Expression (CAGE) in HeLa cells (control, three replicates) and exosome (RRP40) depleted HeLa cells, as a function of the distance to the midpoints (signal summits) of ENCODE HeLa DNase I hypersensitive sites (DHSs). An average of ~93% of HeLa control CAGE tag 5’ ends are within 300bp of DHS summits.

b, The cumulative fraction (vertical axis) of DHS-proximal (within 300bp of a DHS summit) HeLa control and exosome (RRP40) depleted HeLa CAGE tags as a function of the number of DHSs (horizontal axis). This shows that the large majority of transcription initiations are restricted to a small minority of DHSs.
Supplementary Figure 2: Schematic illustration of the characterization of transcribed DNase I hypersensitive sites (DHSs)

DHS-associated strand-specific expression levels in control and exosome (RRP40) depleted HeLa cells were quantified by counting of CAGE tags in genomic windows of 300bp immediately flanking the midpoints (DNase I signal summits) of DHSs. Convergent transcription was not considered. Based on strand-specific expression levels both a directionality score, measuring the strand bias in expression level, and a strand-specific exosome sensitivity score, measuring the relative amount of degraded RNAs by the exosome, were calculated. These three measures were used to summarize the transcriptional biases and properties of each DHS. The directionality ranges between 0 (100% minus strand expression) and 1 (100% plus strand expression), and 0.5 indicates a perfectly balanced bidirectional output. The sensitivity score quantifies the fraction of total (control + RRP40^{-} CAGE) expression seen only after exosome depletion. TPM = tags per million mapped tags.
Supplementary Figure 3: Exosome sensitivity determined via MTR4 (SKIV2L2) depletion with respect to DHS transcriptional directionality

Average exosome sensitivity (vertical axis) of RNAs emanating from transcribed DHSs, broken up by strand and transcriptional strand bias, measured as the fold change (log₂-transformed) between CAGE expression from MTR4 depleted cells vs. control HeLa cells as a function of the distance of capped RNA 5' ends to DHS summits (horizontal axis).
Supplementary Figure 4: Characterization of clustered transcribed DHSs

**a**, Fraction (vertical axis) and number (above each bar) of DHSs associated with GENCODEv17 annotated gene transcript TSSs. GENCODE annotation was simplified to reduce the number of different transcript biotypes (see Methods for details).

**b**, Number of DHSs (vertical axis) overlapping with ENCODE chromatin segmentation states, broken up by DHS cluster. Untranscribed DHS are included for comparison.
Supplementary Figure 5: ChIP-seq profiles separates DHS clusters
Average footprints of ENCODE ChIP-seq signals +/- 500bp around DHS midpoints (DNase signal summits), broken up by DHS clusters. Untranscribed DHSs are included for comparison. Each footprint was normalized to the average signal at all ENCODE HeLa DHSs. Hence, a signal <1 indicates less signal than average while a signal >1 indicates more signal than average. Note that weak unstable DHSs have ChIP-seq profiles of hallmark chromatin epitopes characterizing active enhancers (e.g. H3K4me1, P300, H3K27ac), and that the untranscribed DHSs have evidence of repressive marks and CTCF sites.
Supplementary Figure 6: Downstream RNA processing fates by DHS categories

a, Density plots of transcripts lengths for de novo-assembled transcripts originating from DHSs, broken up by DHS cluster. Yellow densities display spliced length of the assembled transcripts, while light blue densities display the length distribution of unspliced transcripts (genomic length). N indicates the number of transcripts in each group and the horizontal axes denote the transcript length (nt or bp). We note the difference with Ntini et al. (Polyadenylation site–induced decay of upstream transcripts enforces promoter directionality. Nat Struct Mol Biol 20, 923–928 (2013)), which is most likely due to the size constrains imposed by the RNA purification utilized in Ntini et al.

b, Fraction of de novo-assembled transcripts from each DHS category with protein-coding potential defined as obtaining a PhyloCSF score above a threshold of 100.

c-d, Box-plots showing the distribution of transcript length (c) and exon numbers (d) for each DHS cluster, separated by predicted coding and non-coding status as evaluated in (b).
Supplementary Figure 7: Polyadenylation status and cellular localization of RNAs with respect to DHS category

**a**, Boxplots of the ratios (log2) between the number of poly(A)+ and poly(A)- reads (ENCODERNA-seq) mapping to each transcript from the different DHS clusters.

**b**, Distributions of the relative fractions of CAGE tags from nuclear / cytoplasmic ('Nuclear') or polyA+ / polyA- ('polyA+') fractionations broken up by DHS class.
**Supplementary Figure 8: Evolutionary rate versus directionality**

2D densities of the sum of rejected substitutions in windows [-299,-100] (minus strand window, horizontal axis) and [101:300] (plus strand window, vertical axis) around midpoints (DNase signal summits) of DHSs with a transcriptional bias to the plus strand (directionality ≥ 0.9: blue) and minus strand (directionality ≤ 0.1: red). Quadrants I to IV are indicated. We note that unbalanced evolutionary rates (quadrants II and IV) in DHS-flanking regions are highly predictive of transcriptional strand bias.
**Supplementary Figure 9: ChIP-seq profiles separates stable and unstable mRNAs**

Average footprints of ENCODE ChIP-seq signals -500bp to +1500bp around DHS midpoints (DNase signal summits) with respect to mRNA strand, broken up by DHS clusters. Each footprint was normalized to the average signal at all ENCODE HeLa DHSs. Hence, a signal <1 indicates less signal than average while a signal >1 indicates more signal than average. Note major differences in elongation marks H3K79me2, H4K20me1 and H3K36me3 between stable and unstable mRNAs.
**Supplementary Figure 10: Expression properties of stable versus unstable mRNAs**

**a,** Frequencies of transcription termination site hexamer (pA sites) downstream (major strand) of CAGE summits of stable and unstable mRNAs. pA site frequencies downstream of PROMPTs and TSSs of weak unstable and unidirectional stable DHSs are shown for reference. Vertical axis shows the average number of predicted sites per kb within a certain window size from the TSS (horizontal axis) in which the motif search was done. 0 indicates the expected hit density from random genomic background.

**b-c,** Densities of expression specificities (1 - normalized entropy) (**b**) and max expression levels (**c**) of, in HeLa cells, classified stable and unstable mRNAs, based upon FANTOM5 primary cell CAGE samples.
Supplementary Figure 11: snoRNA hosting lincRNAs are stable
Boxplots of exosome sensitivities (vertical axis) of lincRNAs and mRNAs hosting miRNAs, snoRNAs or neither (no).