Supplementary figure 1: Oligo-dT priming enhances the capture of CAGE tags on exons and 3’ UTRs.

CAGE libraries made from THP-1 cells. Data was displayed with the ZENBU genome browser (J. Severin, unpublished data). **(a)** The Actin beta gene is transcribed from right to left (violet arrow) on chromosome 7. **(b)** GAPDH gene is transcribed from left to right (green arrow) on chromosome 12. CAGE libraries were primed RT reaction with (1) random and oligodT (ratio 4:1) primers. (2) oligodT primers only and (3) random primers only. Both panels indicate that oligodT primers could enhance the capture of transcripts on 3’ exons and on internal exons, compared to random primer alone.

Supplementary data 1. The make_ctss script, which is used to cluster the CTSS (Step 65).

```
#!/bin/sh

if [ $# -eq 0 ]
then
  cat <<EOF
Usage is : $0 -q <mapping quality cutoff> <map1.bam> <map2.bam> ...
EOF
  exit 1;
fi

QCUT=


while getopts q: opt do
case ${opt} in
q) QCUT=${OPTARG};;
*) usage;;
esac
done

if [ "${QCUT}" = "" ]; then QCUT=10; fi for

var in "$@"
do
if [[ $var =~ bam$ ]]; then
foo=$var
file=${foo##*/}
base=${file%%.*}
echo working on: $base

TMPFILE="/tmp/$(basename $0).$RANDOM.txt"
     samtools view    -F 4 -u -q $QCUT -b $var |    bamToBed -i stdin > $TMPFILE
     cat    ${TMPFILE} ¥
| awk 'BEGIN{OFS="¥t"}{if($6=="+"){print $1,$2,$5}}' ¥
| sort -k1,1 -k2,2n ¥
| groupBy -i stdin -g 1,2 -c 3 -o count ¥
| awk -v x="$base" 'BEGIN{OFS="¥t"}{print $1,$2,$2+1,    x    ,$3,"+"}' >> $var.ctss.bed

cat   ${TMPFILE} ¥
```

```
| awk 'BEGIN{OFS="¥t"}{if($6=="-"){print $1,$3,$5}}' ¥
| sort -k1,1 -k2,2n ¥
| groupBy -i stdin -g 1,2 -c 3 -o count ¥
| awk -v x="$base" 'BEGIN{OFS="¥t"}{print $1,$2-1,$2, x    ,$3,"-"}' >> $var.ctss.bed

rm $TMPFILE
fi done
```

Supplement figure 1