Figure 1

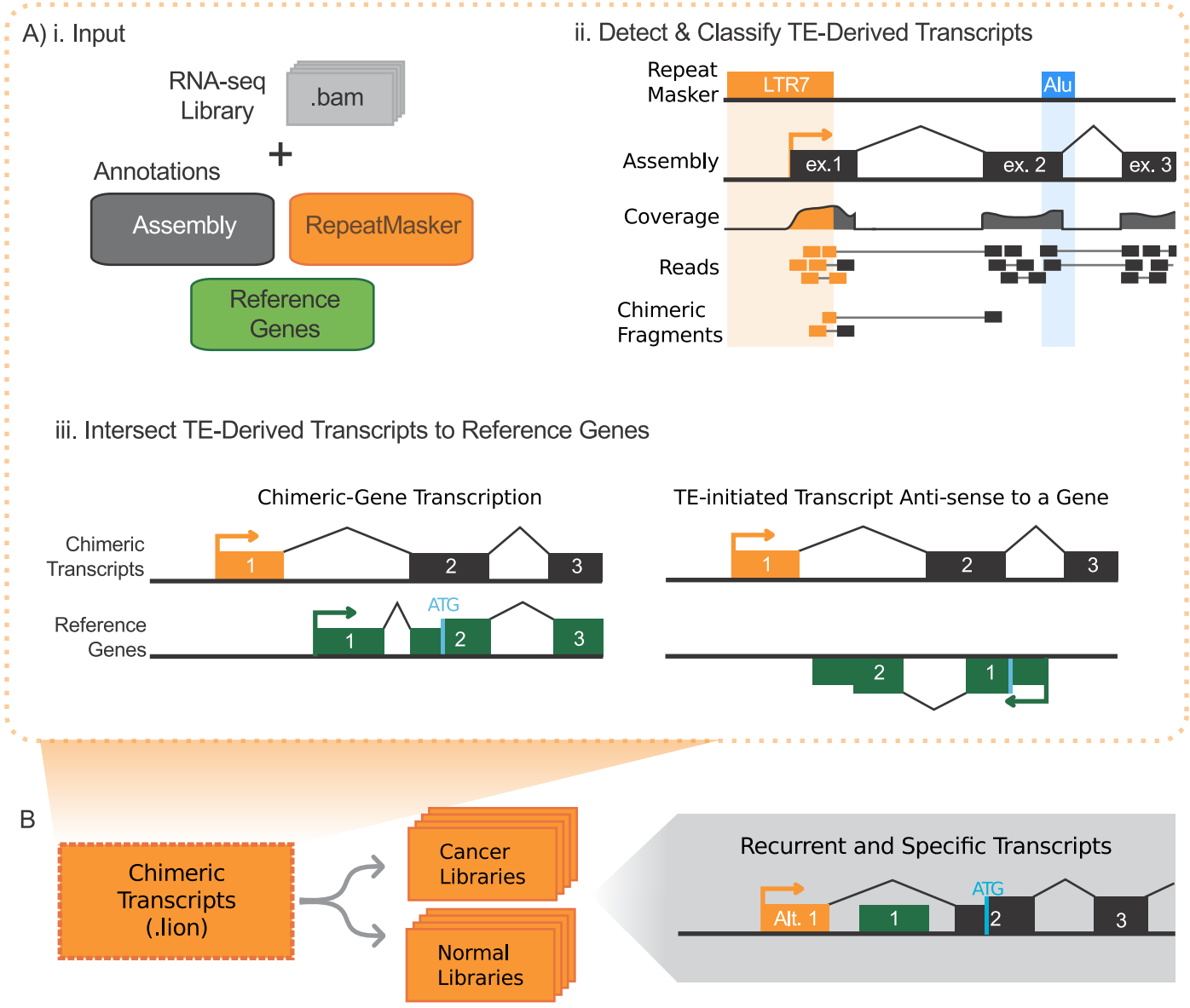

A) i. Input
RNA-seq Library
.bam
Annotations
Assembly
RepeatMasker
Reference Genes
ii. Detect & Classify TE-Derived Transcripts
Repeat Masker
LTR7
Alu
Assembly
ex.1
ex. 2
ex. 3
Coverage
Reads
Chimeric Fragments
iii. Intersect TE-Derived Transcripts to Reference Genes
Chimeric-Gene Transcription
Chimeric Transcripts
Reference Genes
ATG
TE-initiated Transcript Anti-sense to a Gene
B
Chimeric Transcripts (.lion)
Cancer Libraries
Normal Libraries
Recurrent and Specific Transcripts
Alt. 1
ATG

Figure 2

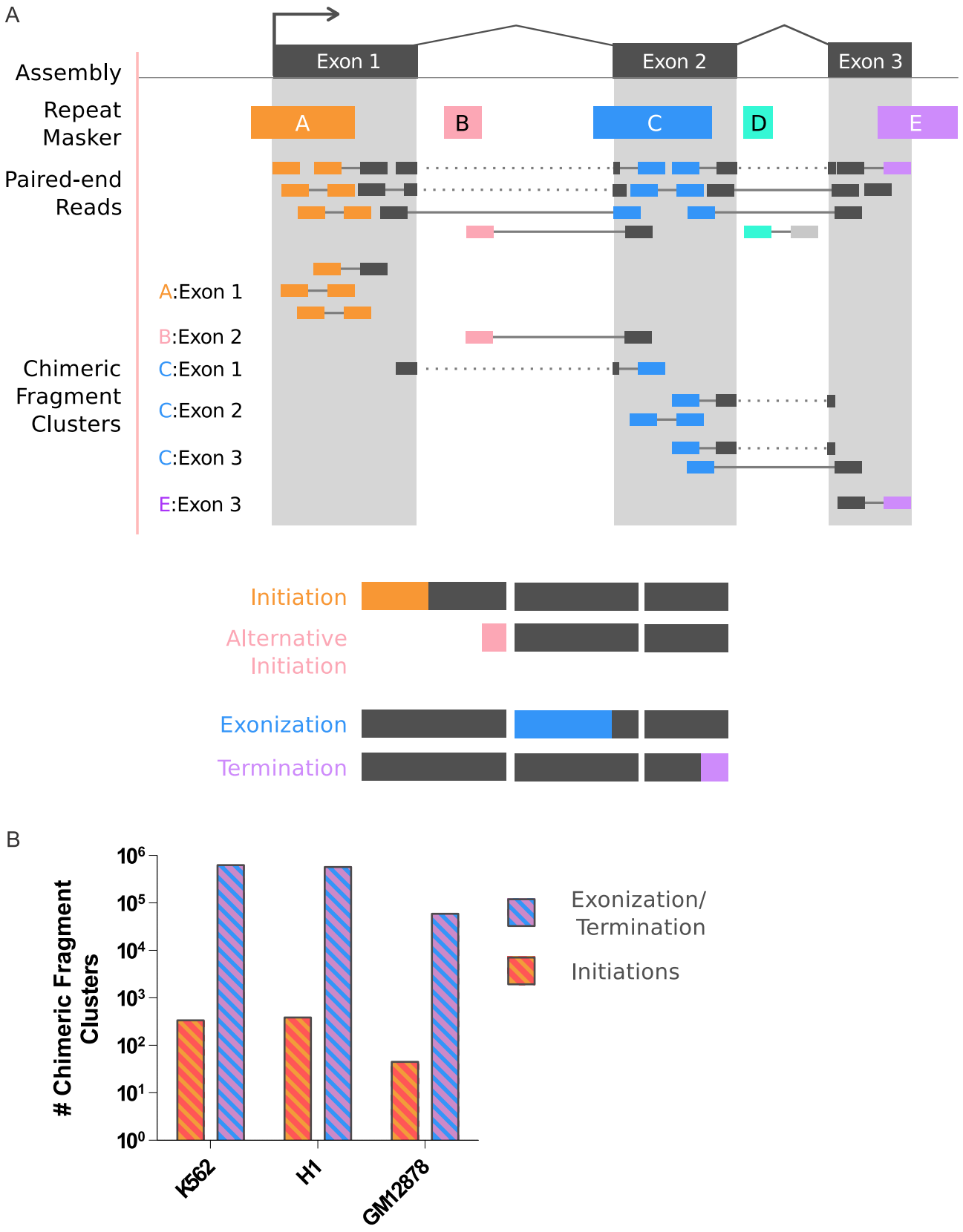
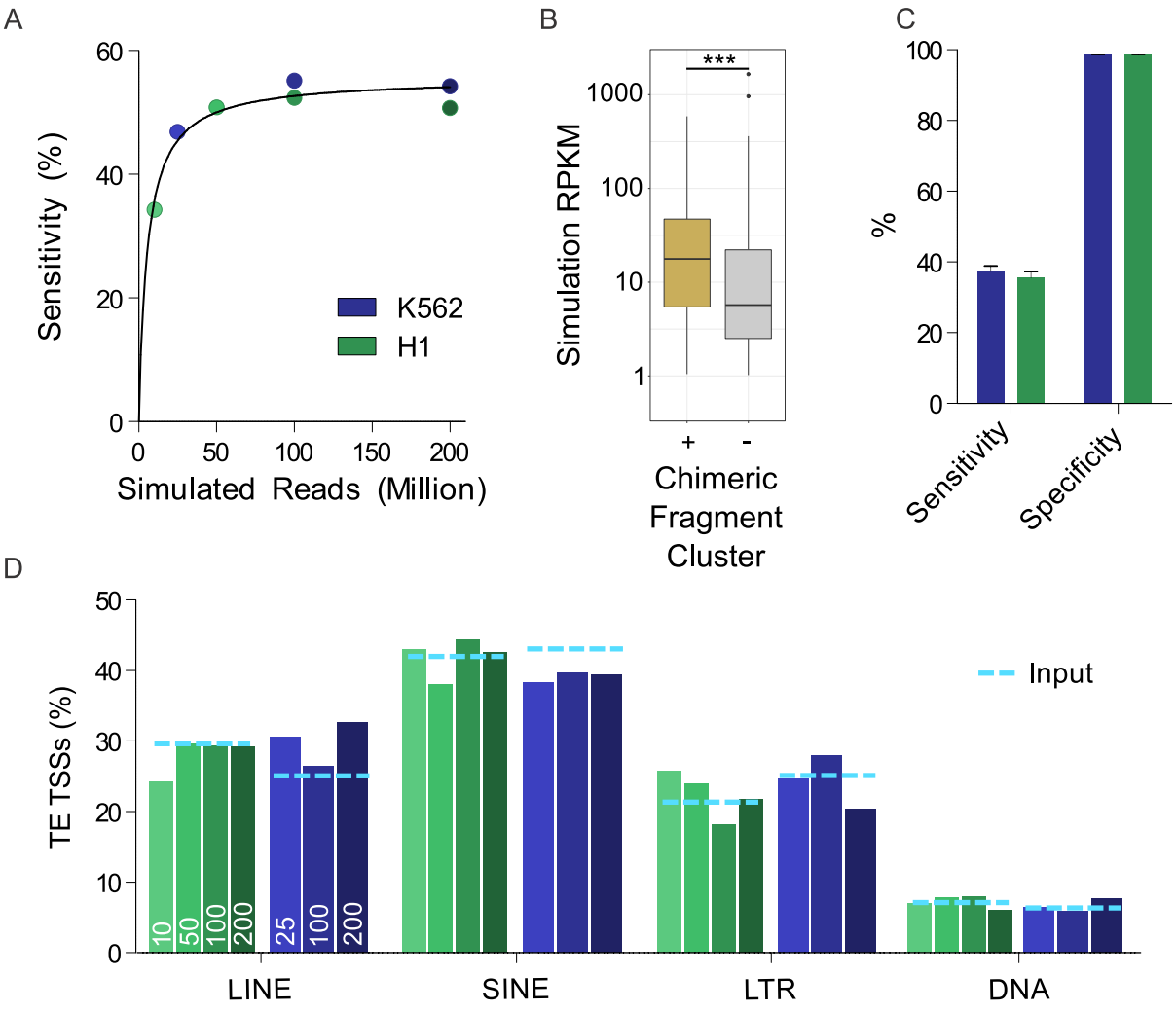

A
Assembly
Repeat Masker
Paired-end Reads
Chimeric Fragment Clusters
Exon 1
Exon 2
Exon 3
A
B
C
D
E
A:Exon 1
B:Exon 2
C:Exon 1
C:Exon 2
C:Exon 3
E:Exon 3
Initiation
Alternative Initiation
Exonization
Termination
B
Chimeric Fragment Clusters
10^6
10^5
10^4
10^3
10^2
10^1
10^0
K562
H1
GM12878
Exonization/ Termination
Initiations

Figure 3


A
Sensitivity (%)
Simulated Reads (Million)
K562
H1
B
Simulation RPKM
***
Chimeric Fragment Cluster
C
%
Sensitivity
Specificity
D
TE TSSs (%)
Input
LINE
SINE
LTR
DNA