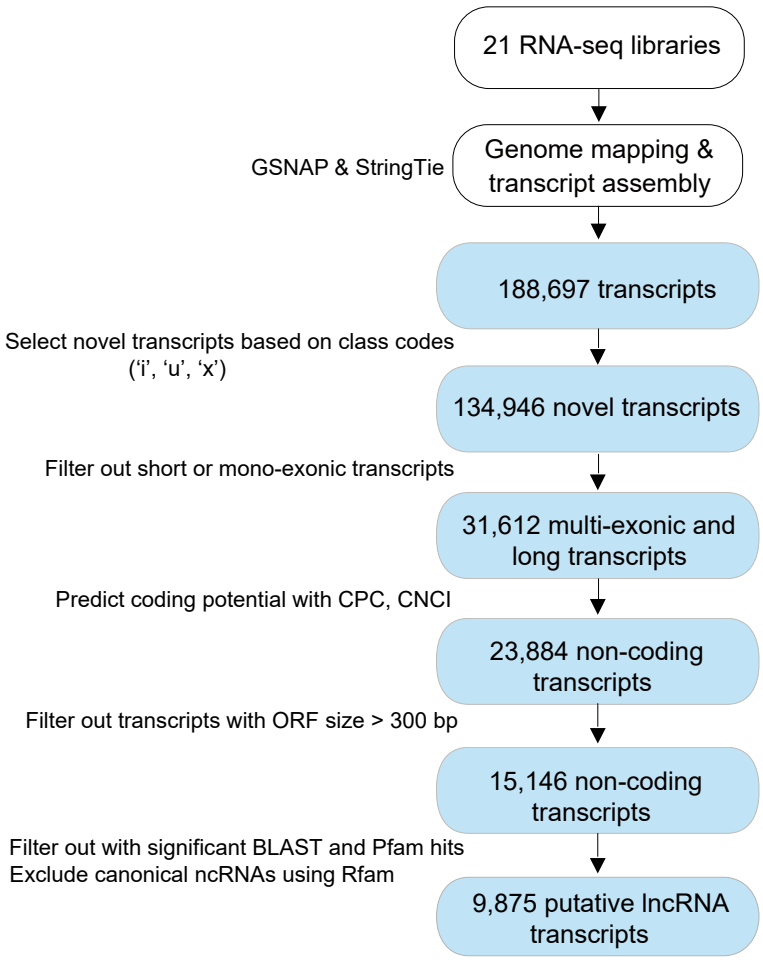
21 RNA-seq libraries
GSNAP & StringTie
Genome mapping & transcript assembly
188,697 transcripts
Select novel transcripts based on class codes ('i', 'u', 'x')
134,946 novel transcripts
Filter out short or mono-exonic transcripts
31,612 multi-exonic and long transcripts
Predict coding potential with CPC, CNCI
23,884 non-coding transcripts
Filter out transcripts with ORF size > 300 bp
15,146 non-coding transcripts
Filter out with significant BLAST and Pfam hits
Exclude canonical ncRNAs using Rfam
9,875 putative lncRNA transcripts