

ANNEXA report

2025-02-25

ANNEXA version: 3.2.3

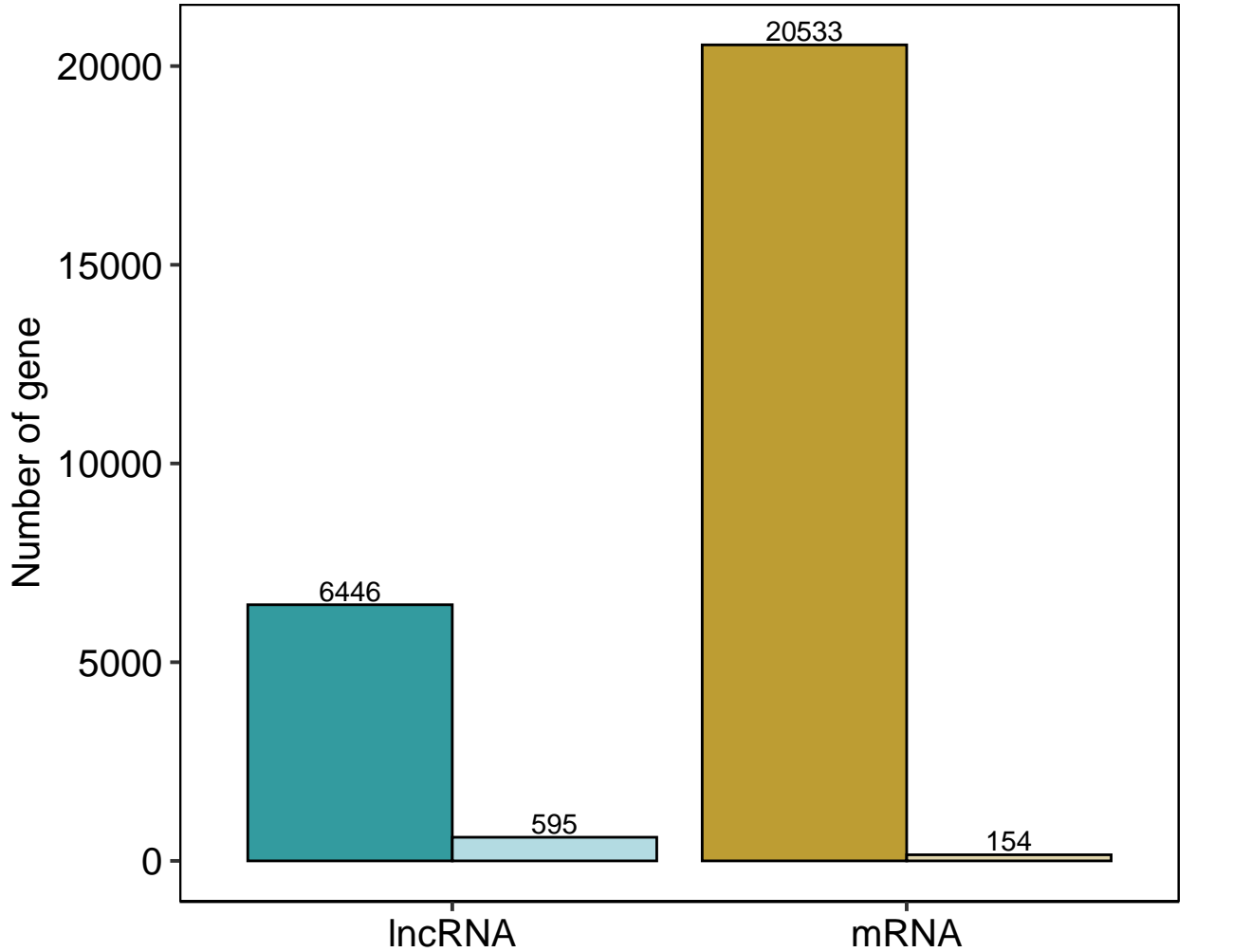
ANNEXA command line:

```
nextflow run /projects/dog/nhoffmann/ANNEXA
--profile singularity,slurm
--tx_discovery bambu
--input /projects/dog/nhoffmann/ANNEXA_BB_vs_STR/dog/canFam4/Ensembl/bambu/samples_new_bam.txt
--gtf /projects/dog/data/canFam4/annotation/ensembl/Canis_lupus_familiarisgsd.UU_Cfam_GSD_1.0.113.chr.gtf
--fa /projects/dog/data/canFam4/sequence/ensembl/Canis_lupus_familiarisgsd.UU_Cfam_GSD_1.0.dna_sm.topf
--filter true
--tfkmers_tokenizer /projects/dog/mlorthiois/TransforKmers/config/tokenizer_k6_512
--tfkmers_model /projects/dog/mlorthiois/TransforKmers/models/dog_5prime_bert_6_12-512/best
--bambu_strand false
--bambu_rec_ndr true
--bambu_singleexon false
--with-report
--resume
```

Gene Characterization

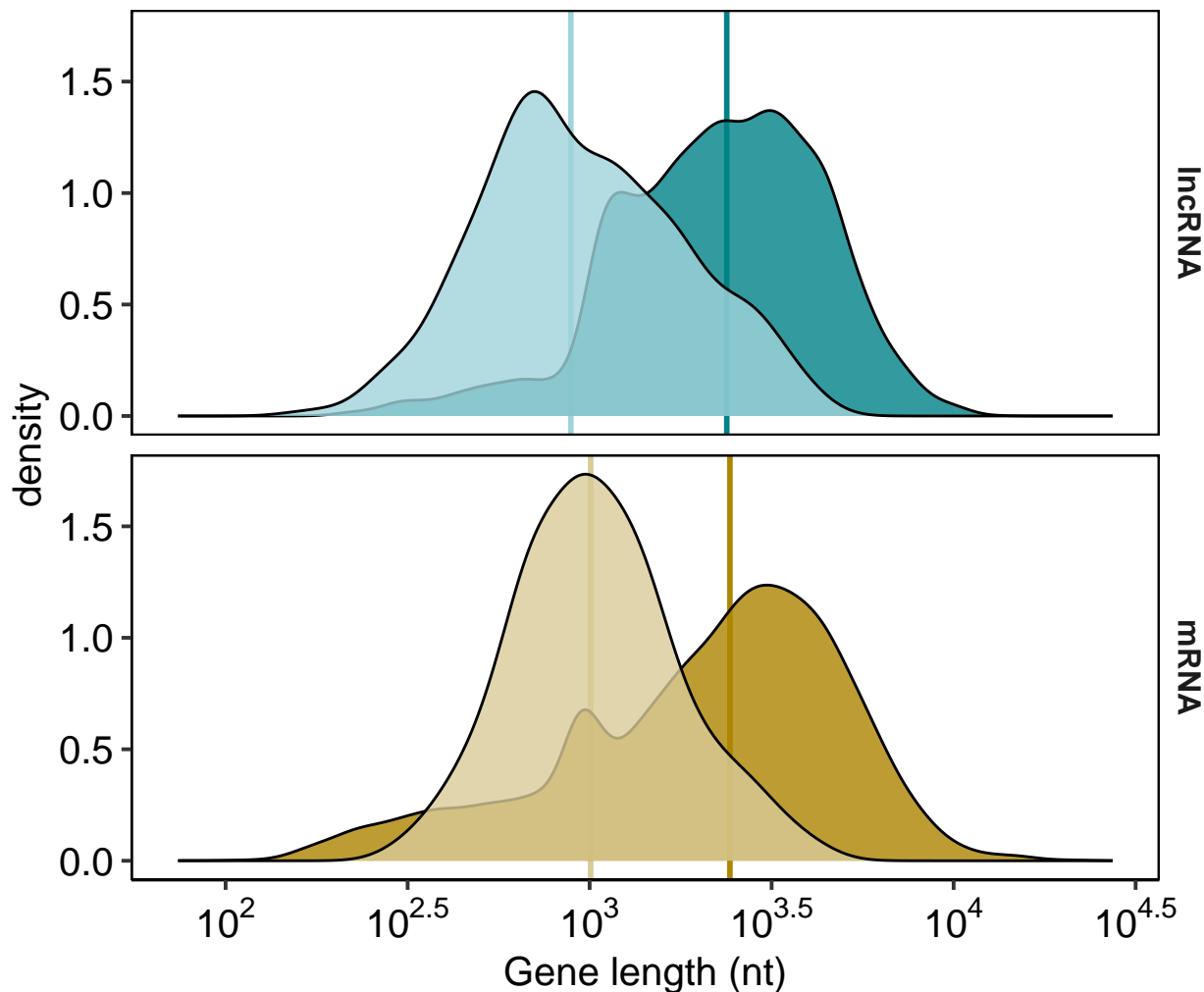
Number of genes

Source IncRNA known IncRNA novel mRNA known mRNA novel

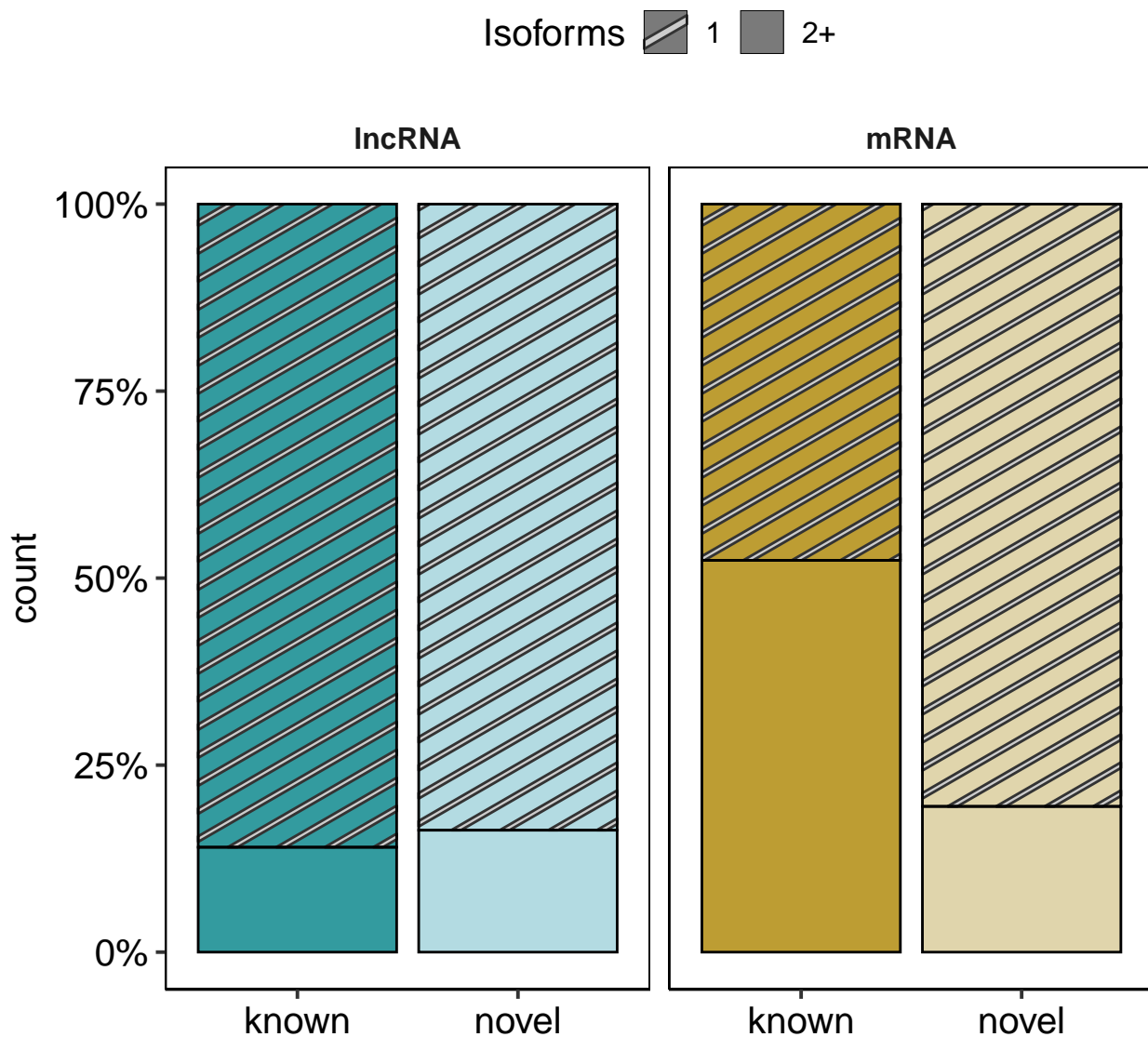


Gene length distribution

Source lncRNA known lncRNA novel mRNA known mRNA novel

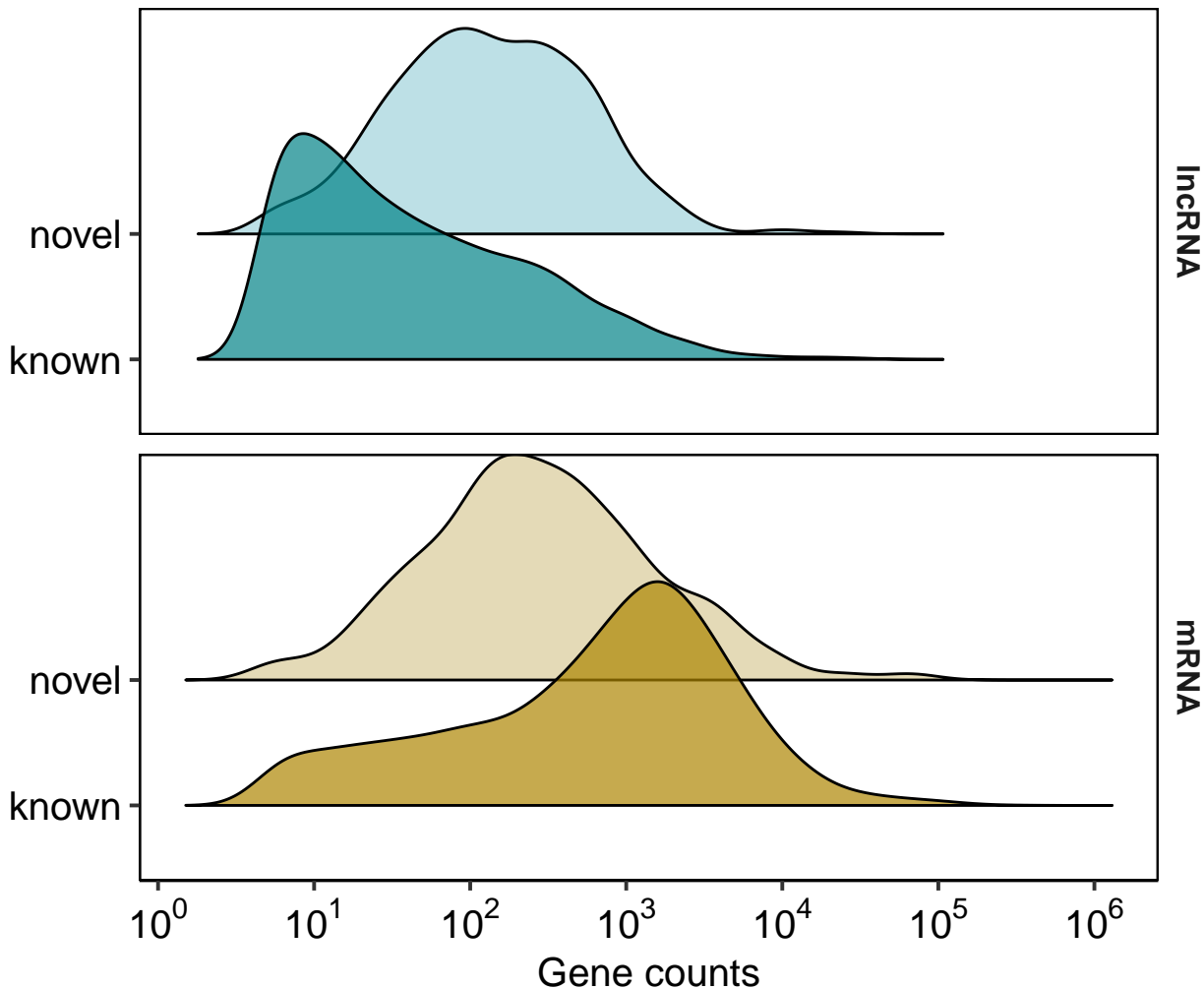


Proportion of mono versus multi-isoform genes



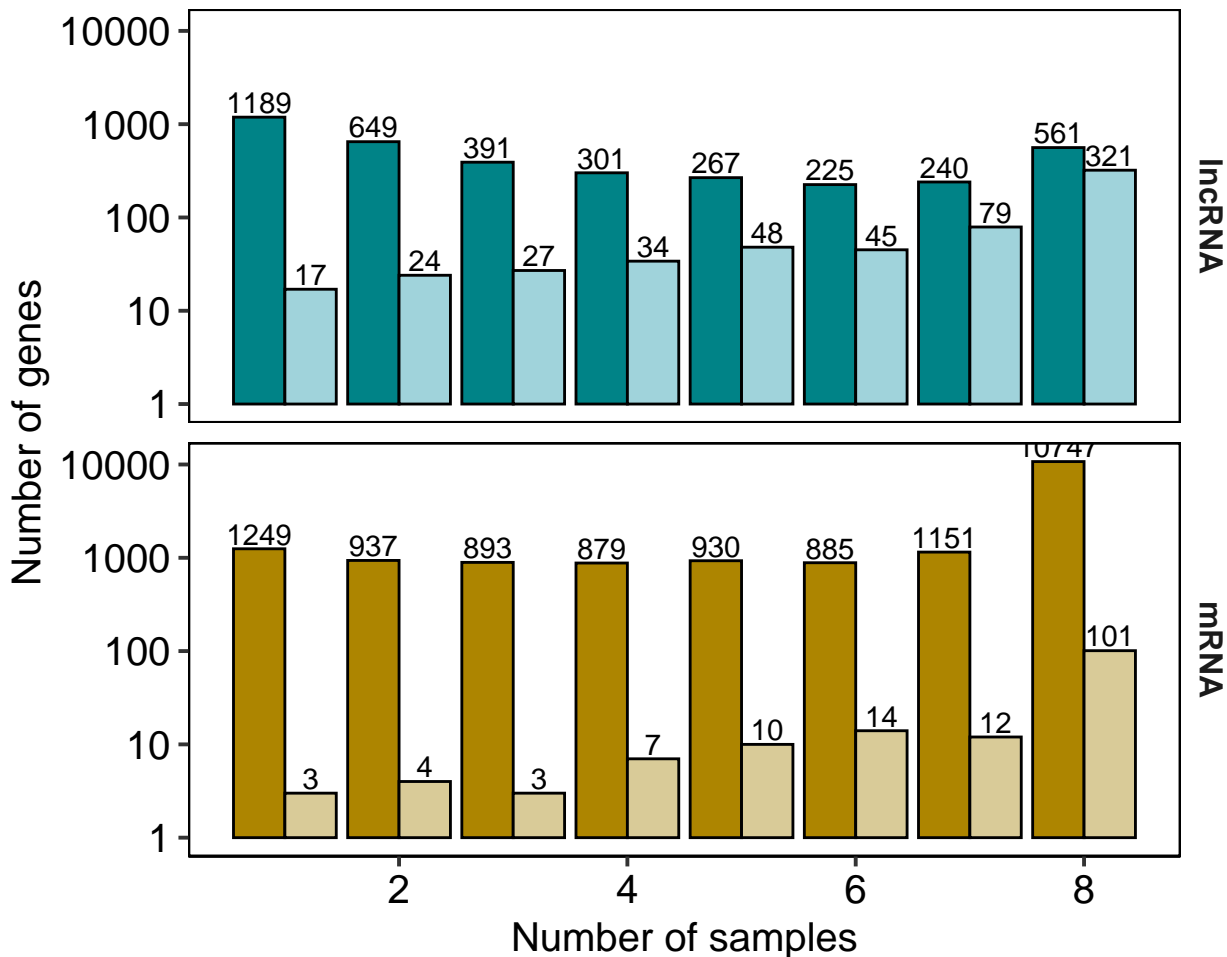
Distribution of gene counts

Source lncRNA known lncRNA novel mRNA known mRNA novel



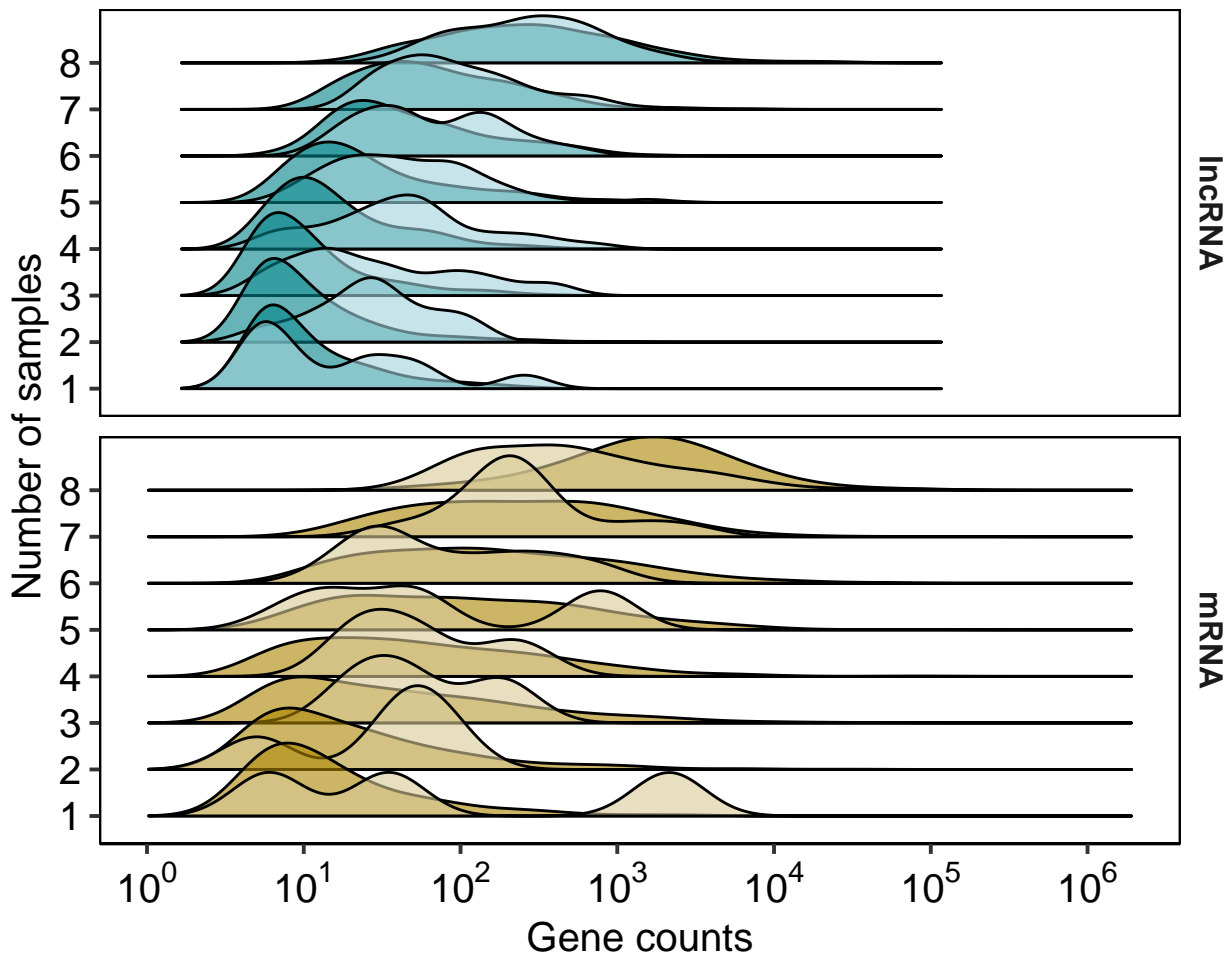
Number of genes according to his breadth of expression (gene count >1)

Source ■ lncRNA known ■ lncRNA novel ■ mRNA known ■ mRNA novel



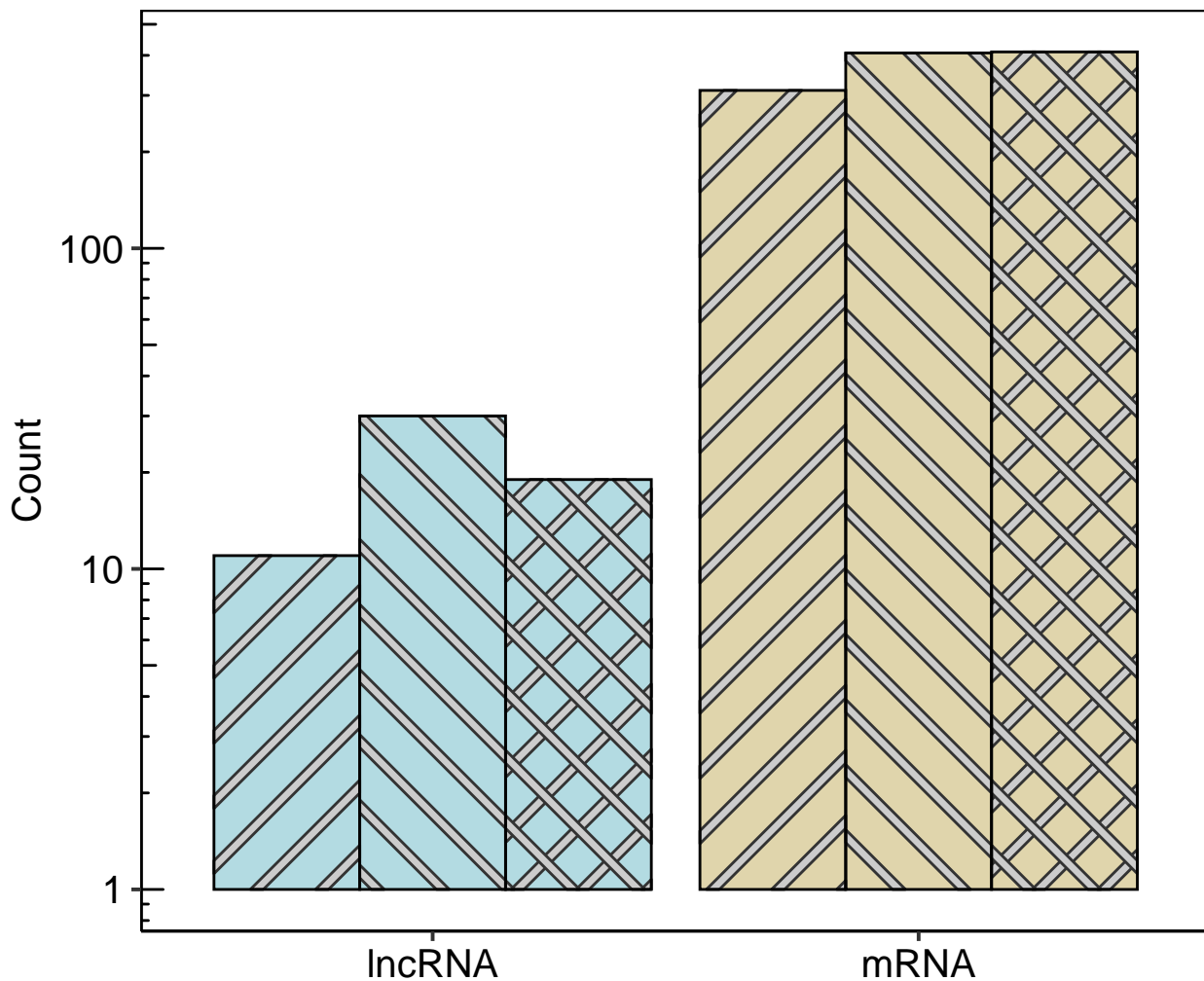
Gene counts separated by number of samples with at least 1 count

Source IncRNA known IncRNA novel mRNA known mRNA novel



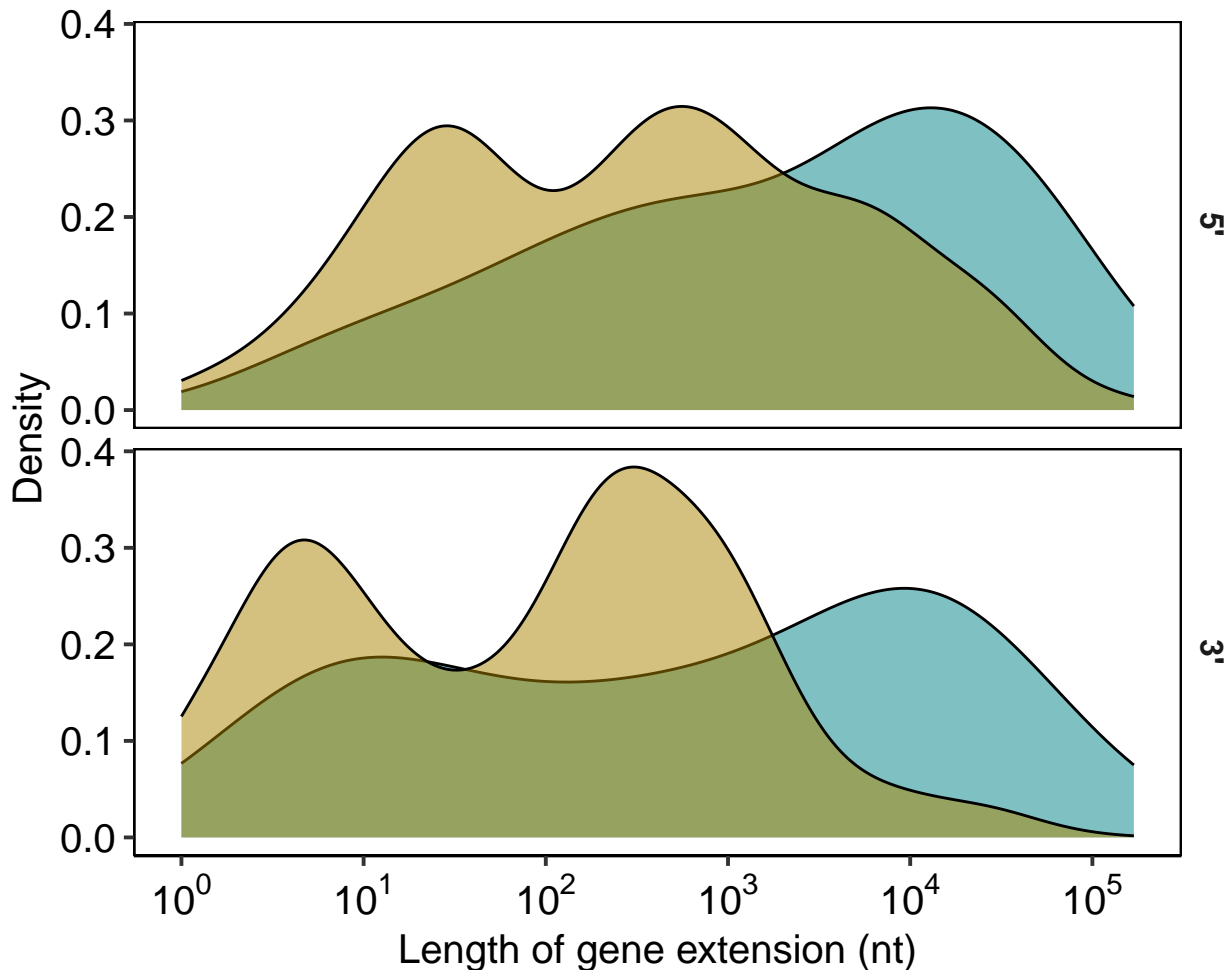
Number of 5' and 3' gene extensions

Gene extension 3' 5' 5'-3'



Distribution of 5' and 3' gene extensions (at genomic level)

Source ■ lncRNA known ■ mRNA known

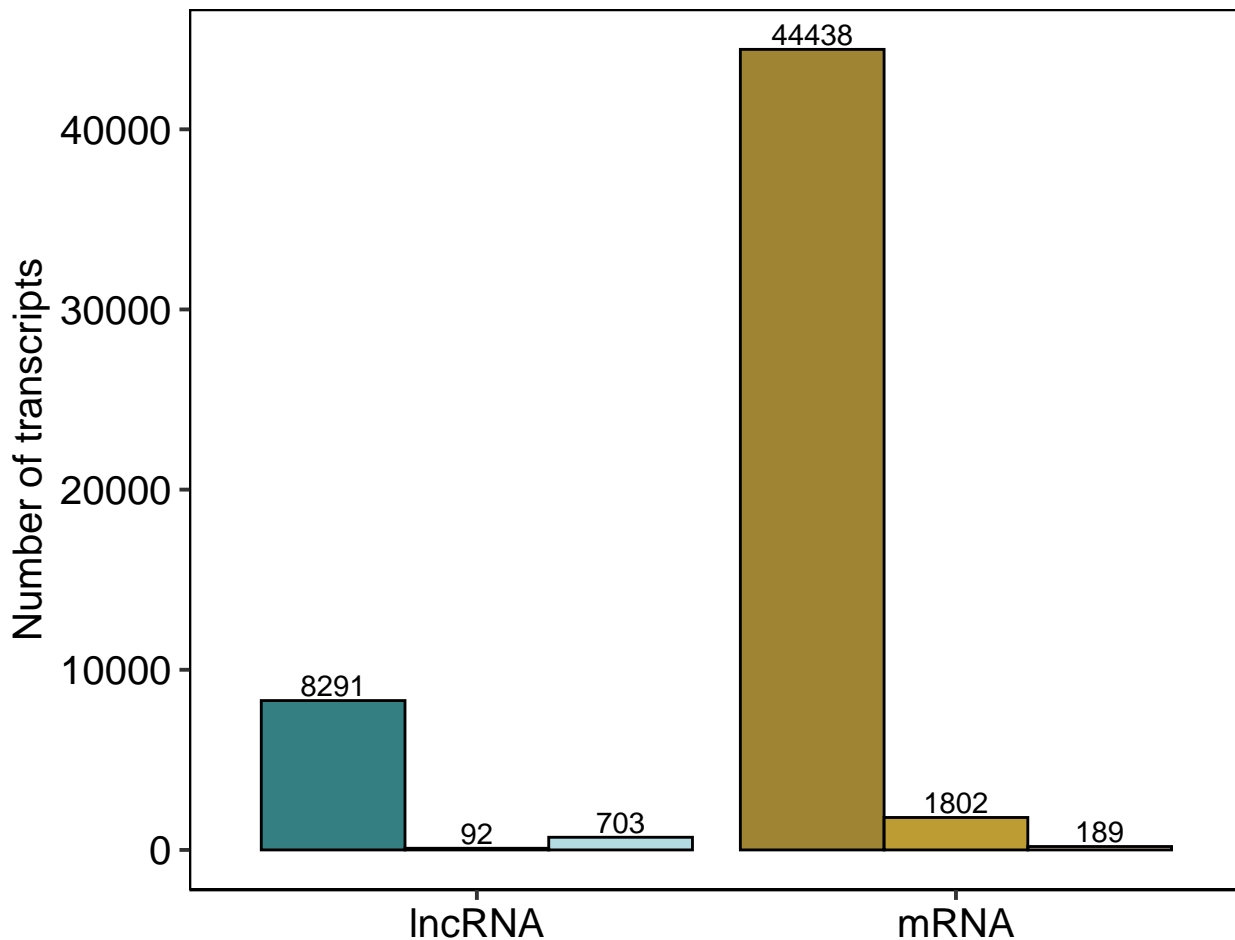


Transcript Characterization

Number of transcripts

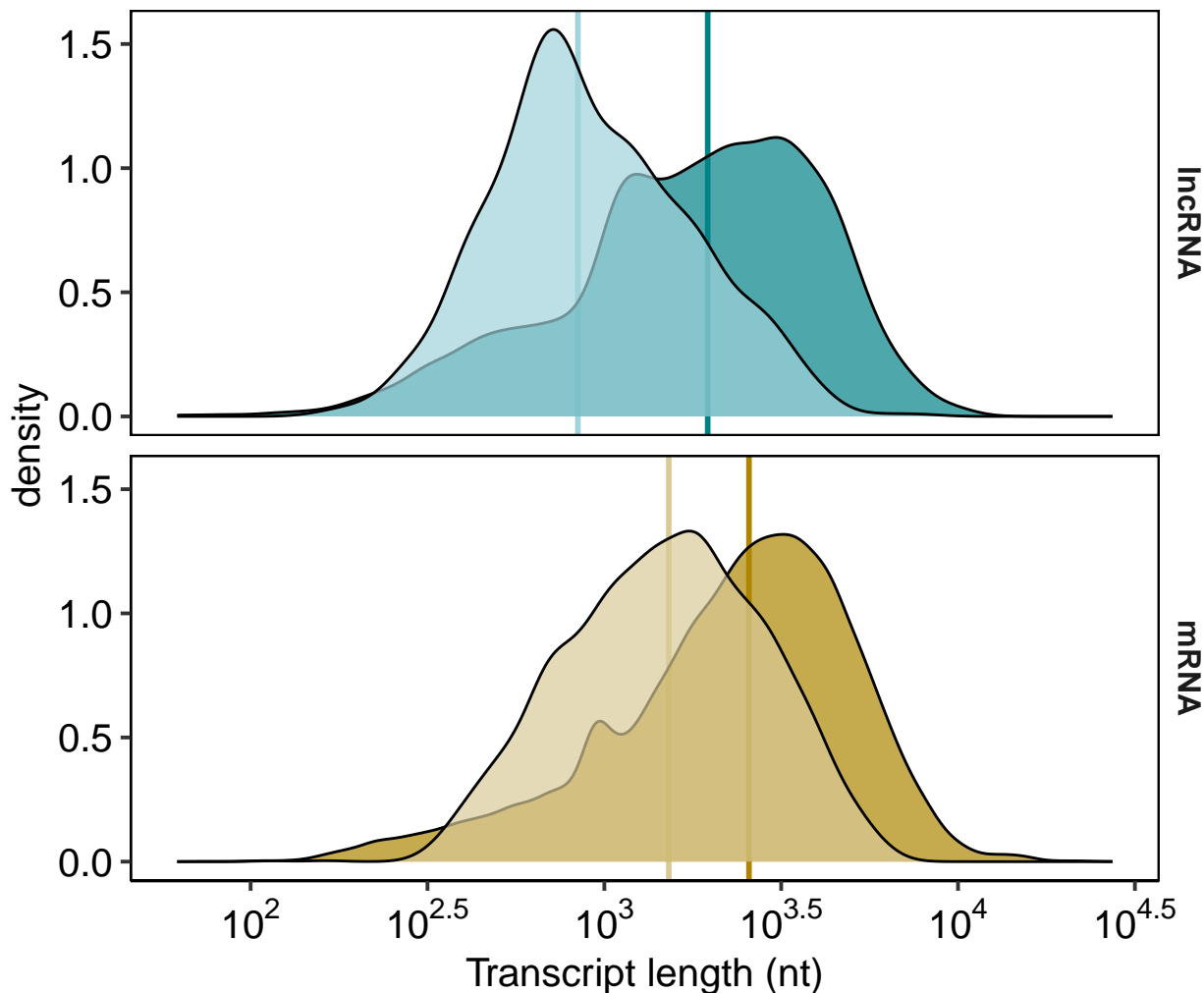
K=known, N=novel

K in K gene	N in N gene	N in K gene
N in K gene	K in K gene	N in N gene



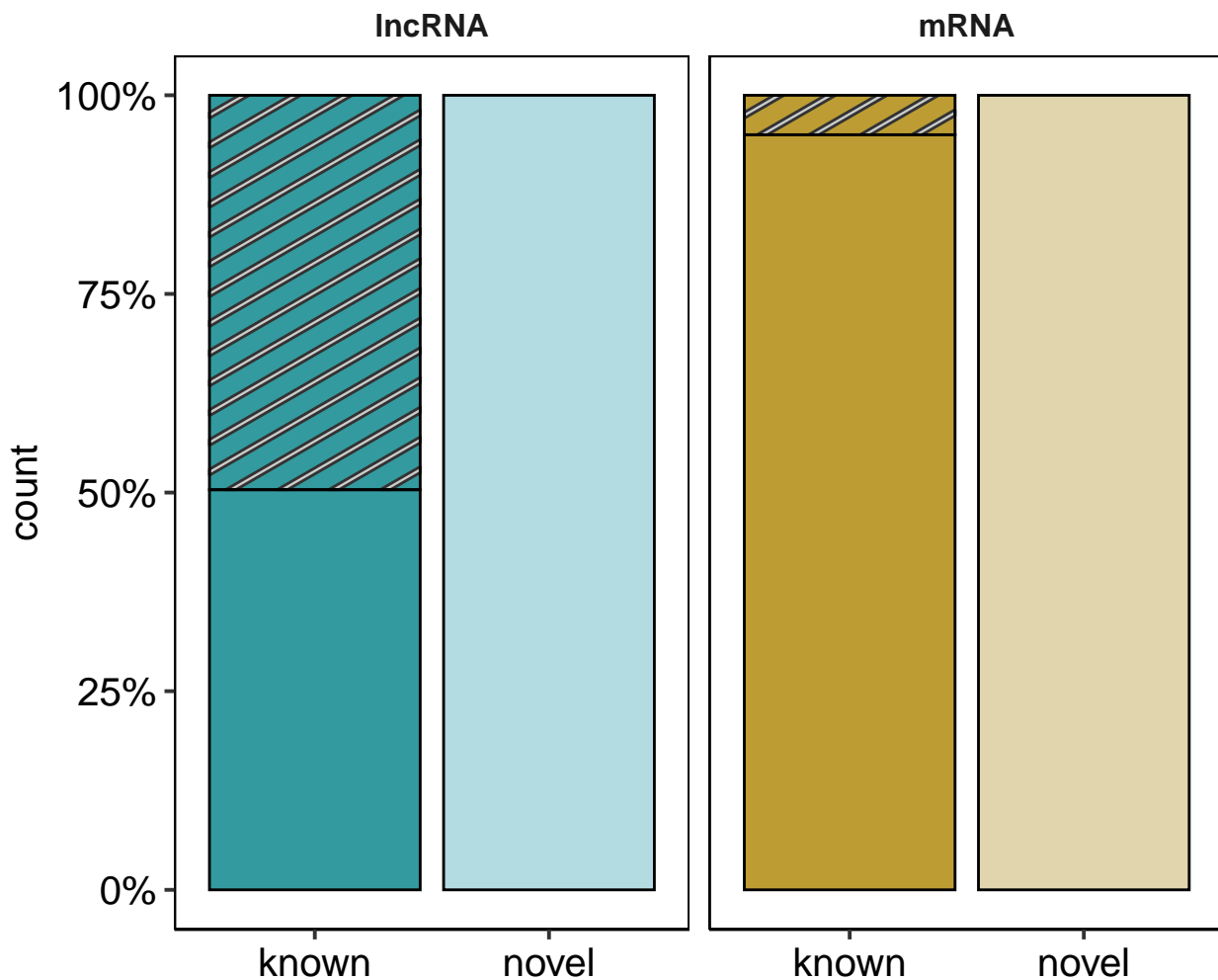
Transcript length distribution

Source lncRNA known lncRNA novel mRNA known mRNA novel







Proportion of mono versus multi exonic transcripts

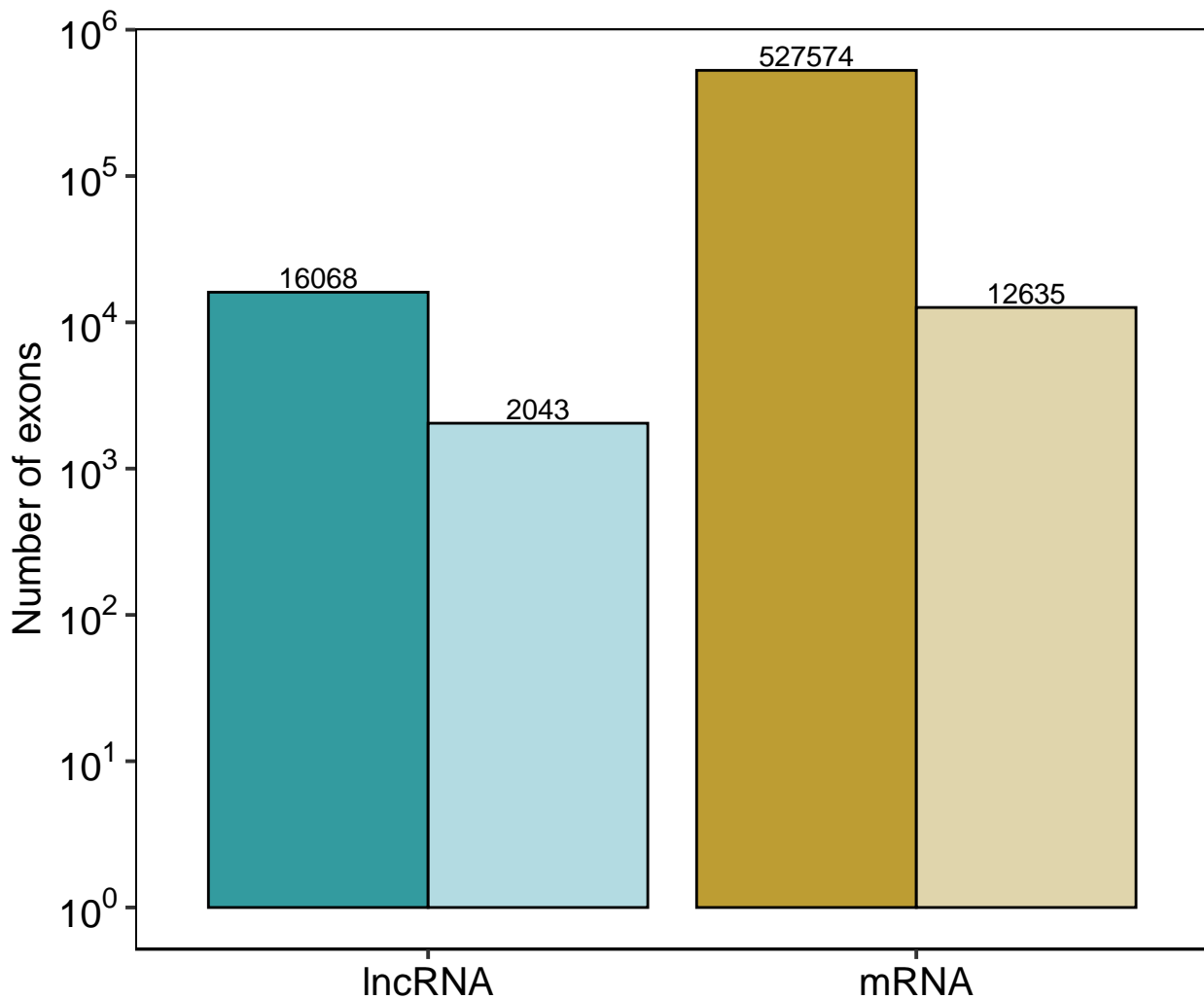
Exons 1 2+



Exon Characterization

Number of exons

Source  lncRNA known  lncRNA novel  mRNA known  mRNA novel



Exon length distribution

Source lncRNA known lncRNA novel mRNA known mRNA novel

