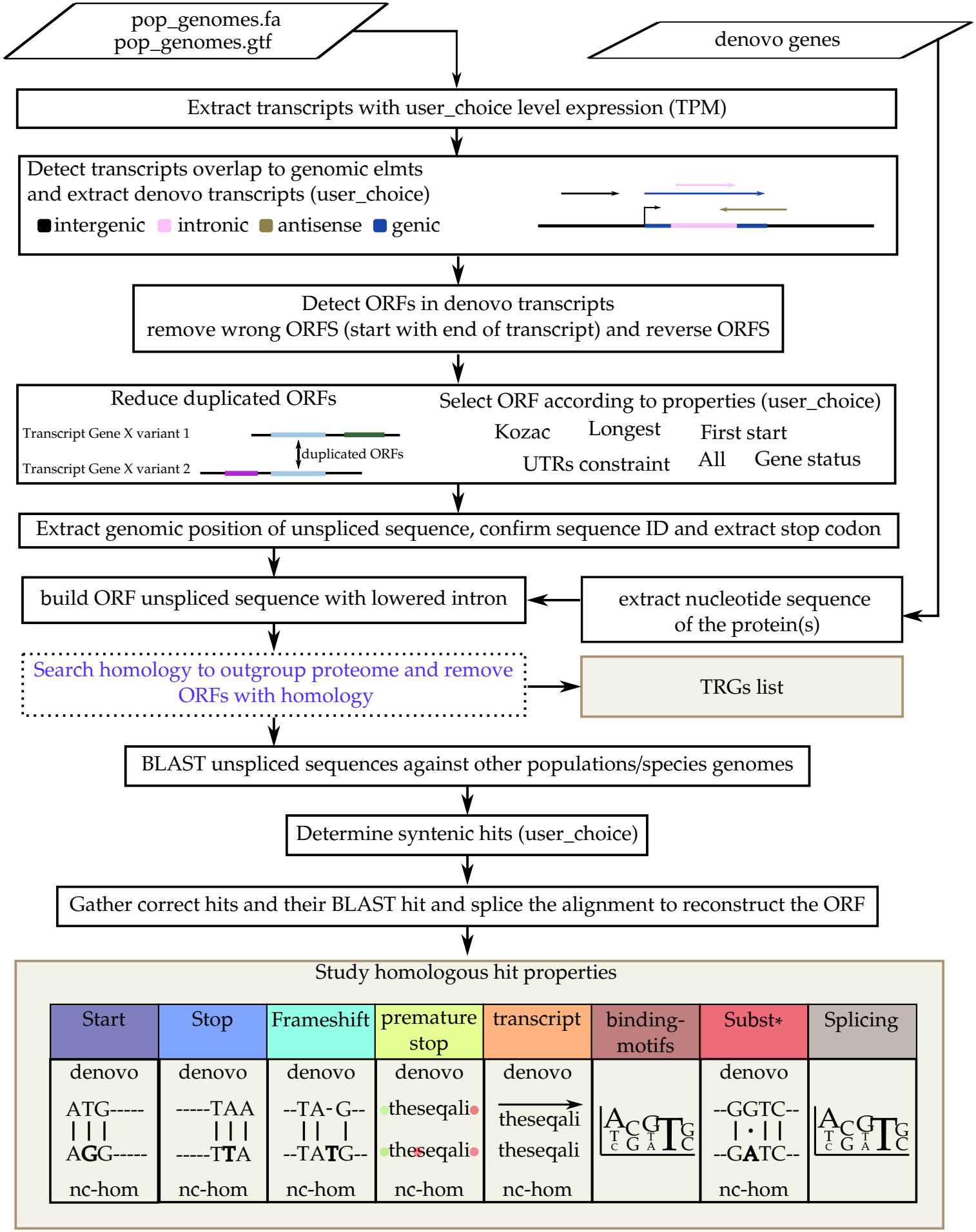
pop_genomes.fa
pop_genomes.gtf
denovo genes
Extract transcripts with user_choice level expression (TPM)
Detect transcripts overlap to genomic elmts
and extract denovo transcripts (user_choice)
intergenic
intronic
antisense
genic
Detect ORFs in denovo transcripts
remove wrong ORFS (start with end of transcript) and reverse ORFS
Reduce duplicated ORFs
Transcript Gene X variant 1
duplicated ORFs
Transcript Gene X variant 2
Select ORF according to properties (user_choice)
Kozac
Longest
First start
UTRs constraint
All
Gene status
Extract genomic position of unspliced sequence, confirm sequence ID and extract stop codon
build ORF unspliced sequence with lowered intron
extract nucleotide sequence
of the protein(s)
Search homology to outgroup proteome and remove
ORFs with homology
TRGs list
BLAST unspliced sequences against other populations/species genomes
Determine syntenic hits (user_choice)
Gather correct hits and their BLAST hit and splice the alignment to reconstruct the ORF
Study homologous hit properties
Start
Stop
Frameshift
premature stop
transcript
binding-motifs
Subst*
Splicing
denovo
ATG-----
AGG-----
nc-hom
denovo
-----TAA
-----TTA
nc-hom
denovo
--TA-G--
--TATG--
nc-hom
denovo
theseqali
theseqali
nc-hom
denovo
theseqali
theseqali
nc-hom
denovo
--GGTC--
--GATC--
nc-hom