Title	ANNOgesic - A computational pipeline for RNA-Seq based transcriptome annotations of bacteria
Author	*Sung-Huan Yu*, Konrad U. Förstner, Jörg Vogel
Affiliation	Institute for Molecular Infection Biology (IMIB), University of Würzburg, Germany
Contact	konrad.foerstner@uni-wuerzburg.de
URL	https://github.com/Sung-Huan/ANNOgesic
License	ISC license

High-troughput RNA sequencing (RNA-Seq) has become a powerful tool to improve the transcriptome/genome annotations of organisms. This technology has helped to detect new transcripts including numerous ones of non-protein-conding genes which are hard to predict purely on the genome sequence. Still, the translation from RNA-Seq data into meaningful annotations is a labor intensive task and lacks streamlining. Here we present the open-source licensed (ISC license) command line tool *ANNOgesic* which provides several subcommands that assist in the RNA-Seq data based generation of high-resolution transcriptome annotations with a focus on bacterial species.

Depending on the specific task the tool requires different input files like the reference genome sequence, RNA-Seq read alignments and available annotations of the organism to study or closely related species. *ANNOgesic* searches for new loci and redefines gene boundaries of those and previously known ones based on transcript assemblies as well as on transcriptional start sites and terminator predictions. It integrates those findings with further information from public sources (like gene functions classifications from Gene Ontology) and produces high-resolution annotations in GFF3 format (Gene feature format). Additionally, it can group genes into operons and suboperons, detect circular RNAs, Single Nucleotide Variation (SNV) as well as processing sites and performs target predictions for newly found sRNAs.

Taken together, *ANNOgesic* offers several functionalities that improve the quality and increase the speed of the transcriptome annotation process significantly.