Submitter Name: Jesse Niehaus
PI Name: Grégory Scherrer

Submitted Email: jesse_niehaus@med.unc.edu
PI Email: gregory_scherrer@med.unc.edu

# Accurately quantifying *Oprm1* requires an updated gene annotation.

Jesse K. Niehaus[1], Hannah C. Nourie[1], Alexander Barnette[1], Hongkui Zeng[2], Grégory Scherrer[1]

[1]Department of Cell Biology and Physiology, UNC-Chapel Hill
[2]Allen Institute for Brain Science.

Opioids produce analgesia and dangerous side effects by binding to the mu-opioid receptor (MOR; encoded by *Oprm1*) in different brain regions. Single-cell RNA-sequencing (scRNA-seq) can characterize the heterogeneity of cells expressing *Oprm1* (*Oprm1+*), identify marker genes to distinguish *Oprm1+* cell types, and reveal putative, cell type-specific, drug targets in different *Oprm1+* neurons. While generating a brain-wide atlas of *Oprm1+* cells, we discovered a ~20-fold difference in the quantification of *Oprm1+* cell proportions based on whether reads (*i.e.*, the sequenced fragments from the scRNA-seq library) aligning to introns were included. We determined that intronic *Oprm1* reads primarily originate from an erroneously shortened 3'UTR annotation in the canonical MOR-1 isoform and from alternative *Oprm1* splice variants. Bioinformatic and histological analyses revealed that intronic reads pertaining to the 3'UTR likely originate from *Oprm1* RNAs. These reads would only be accurately quantified if intronic reads were included during alignment. Conversely, we determined that intronic reads from alternative *Oprm1* isoforms are likely generated from technical artifacts. In this case, intronic read inclusion generates false positives, accounting for most *Oprm1* reads in cortical neurons and virtually all *Oprm1* reads in non-neuronal cells such as microglia. By generating a refined gene annotation, we retained false negative and removed false positive *Oprm1* reads, which collectively increased *Oprm1+* cell quantification concordance between sequencing and histological approaches. These findings extend to other neuronal genes, which may alter how neurons in reward pathways are classified by transcriptomics or directly affect quantification for RNAs related to drugs of abuse.