**Summary**

In Quezada-Ramírez et al. (2023), GFP was transferred into hematopoietic stem and progenitor cells (HSPCs) at genomic locations for evaluation as Genomic Safe Harbors (GSHs). Bulk RNAseq data was generated from control and transfected HSPCs for the established GSH AAVS1 as well as five novel GSH sites as identified from endogenous parvoviral elements (EPVs). The expressions of the GFP is quantified from control and transfected HSPCs.

**Methods and Results**

A reference was generated obtaining coding DNA (cDNA) from human genome version 38 (hg38): <https://ftp.ensembl.org/pub/release-112/fasta/homo_sapiens/cdna/Homo_sapiens.GRCh38.cdna.all.fa.gz>. The DNA coding sequence of human-optimized GFP was appended to the transcript files (Yang, Cheng, and Kain 1996), and an index prepared for use with Salmon (Patro et al. 2017). Transcripts were quantified in terms of Transcripts Per Million (TPM) by using the *quant* option of Salmon. The following is an example command:

*srun -p med -t 72:00:00 --mem=32G --nodes=1 --ntasks-per-node=1 --cpus-per-task=8 \*

*salmon quant -i data/salmon/human\_index -l A \*

*-1 data/donors-1-2/AAVS1-1\_R1\_001.fastq.gz \*

*-2 data/donors-1-2/AAVS1-1\_R2\_001.fastq.gz \*

*-p 8 ==validateMappings -o outputs/301/AAVS1-Donor.1*

Subsequently, the output was consolidated using basic command line utilities and imported into R (R Development Core Team 2024) for creation of a summary (Table 1).

**Table 1** Transcripts Per Million (TPM) values calculated from six donors, with controls and six edited loci.

