**Quantification of NPC transcriptomes with eXpress**

In addition to RSEM, we used eXpress (version 1.5.1) [1] to estimate transcript abundances for the complete RefSeq transcriptome and two the Event Analysis reduced references (100% events detected at APN > 0, ≥75% events detected at APN ≥ 5) for the mouse neural progenitor cell data. Reads were first aligned to transcript sequences using BWA-MEM version 0.7.12 [2] using default parameters. Output SAM files were then analyzed using eXpress to estimate abundances of transcripts in each reference transcriptome. Mean fragment length of 80bp and with a standard deviation of 50bp was set for eXpress, with all other parameters left at their default settings.

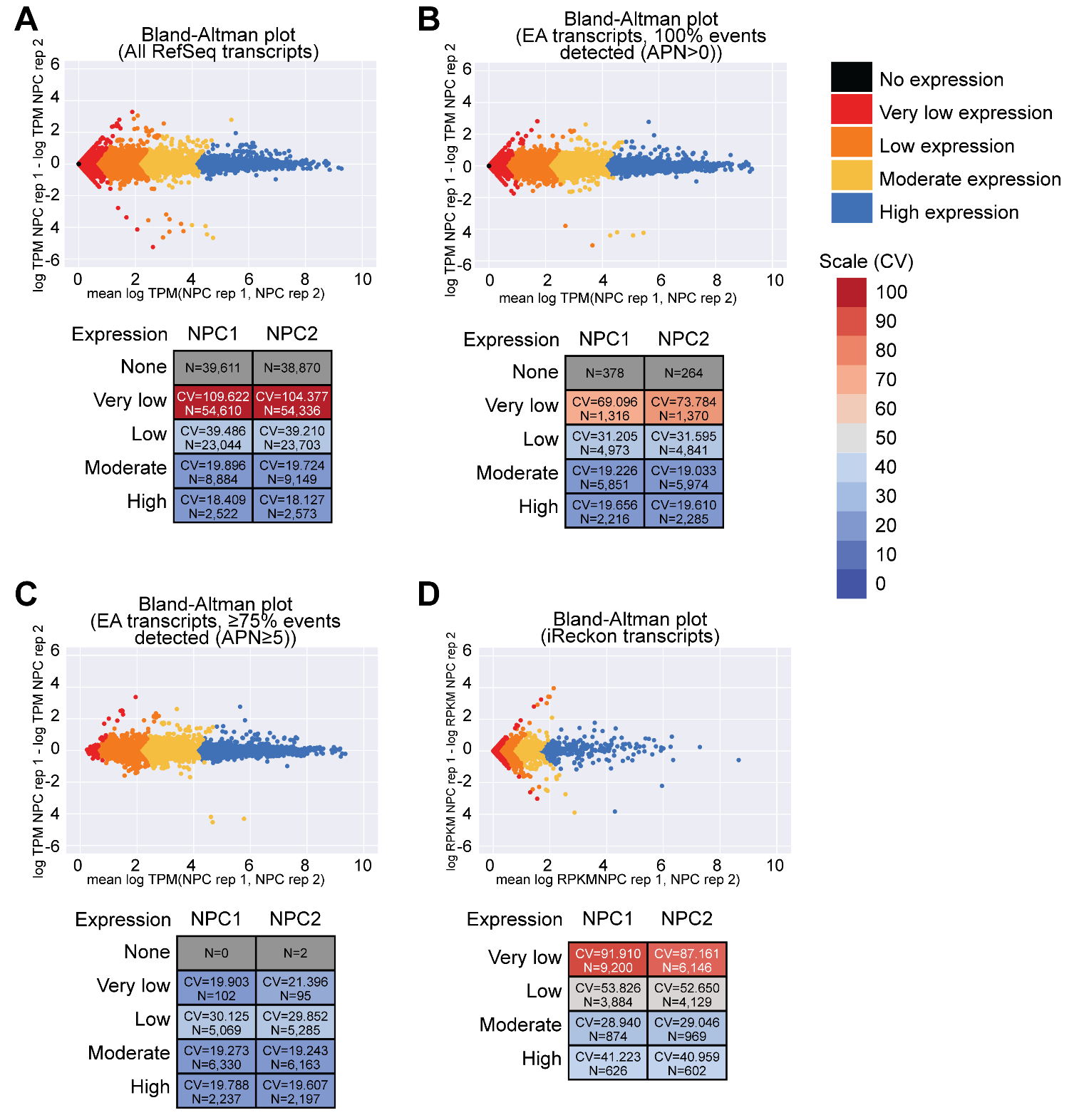
To estimate concordance of transcript estimates between replicates, we first calculated the simple agreement between replicates as the proportion of transcripts detected in both replicates. Transcripts were then binned based on their TPM estimates (transcripts per kilobase million) estimates: “no expression” (TPM=0 in both replicates), “very low expression” (minimum log-TPM: 0 - 0.5), “low expression” (minimum log-TPM: 0.5 - 2), “moderate expression” (minimum log-TPM: 2 – 4) and “high expression” (minimum log-TPM: 4 or greater). This approximately corresponds to no expression, 25th percentile of expression (“very low expression”), 25th to 75th percentile of expression (“low expression”), 75th to 90th percentile of expression (“moderate expression”), and 90th to 100th percentile of expression (“high expression”). The coefficient of variation (CV) for each TPM bin was then calculated.

**Transcriptome assembly and quantification with iReckon**

The STAR aligner was used to align reads from the mouse neural progenitor cell data to the complete mouse mm10 genome. STAR references were generated using RefSeq mm10 GFF3 annotations, and a maximum junction overhang size of 40 bp either side of the junction was set for the creation of STAR's splice junction database, analogous to that of the junction catalog used by Event Analysis. STAR alignment parameters were also set to mimic those used by Event Analysis: reads were aligned to the mm10 genome using the alignment mode “EndToEnd” to force STAR to not soft-clip reads, minimum allowable read junction overhang was set to 16 bp, up to three mismatches were allowed per alignment, and multimapping alignments were disallowed.

Genome alignments and junctions identified using STAR where then used to assemble and estimate transcripts using iReckon, version 1.0.8 [3]. Annotations consisting of a 12-column BED file of transcripts (including genomic start positions and lengths of each exon) was used to guide transcript assembly. Alignments to transcript sequences were carried out using BWA-MEM (version 0.7.12) [2] as per the iReckon user guide, and all parameters were left at their default settings.

To estimate concordance of transcript estimates between replicates, we first calculated the simple agreement between replicates as the proportion of transcripts detected in both replicates. Transcripts observed in both samples were then binned based on the minimum RPKM (reads per kilobase million) estimate between NPC1 and NPC2: “no expression” (RPKM = 0 in both replicates), “very low expression” (minimum log-RPKM: 0 - 0.1), “low expression” (minimum log-RPKM: 0.1 - 0.6), “moderate expression” (minimum log-RPKM: 0.6 – 1.6) and “high expression” (minimum log-RPKM: 1.6 or higher). This approximately corresponds to no expression, 25th percentile of expression (“very low expression”), 25th to 75th percentile of expression (“low expression”), 75th to 90th percentile of expression (“moderate expression”), and 90th to 100th percentile of expression (“high expression”). Transcripts that were present in only one replicate were assigned to the RPKM bins described above. The coefficient of variation (CV) for each RPKM bin for each sample was then calculated.



**Supplementary Figure S3.1.** Replicate-to-replicate concordance of eXpress transcript estimates for RefSeq and Event Analysis transcripts. Bland-Altman plots and coefficients of variance (CVs) for (A) all RefSeq transcripts (128,631 transcripts); (B) Event Analysis (EA) transcripts with 100% of events detected (APN>0; N=14,734); and (C) EA transcripts with at least 75% of events detected (APN≥5; N=13,740). Transcripts are binned into no expression (log-TPM = 0; black), very low expression (0 < log-TPM < 0.5; red), low expression (0.5 ≤ log-TPM < 2; orange), moderate expression (2 ≤ log-TPM < 4; yellow), and high expression (log-TPM ≥ 4; blue) and the CVs for each bin and for each NPC replicate (NPC rep 1, NPC rep 2) are indicated. (D) Bland-Altman plot for all iReckon-assembled transcripts from both NPC replicates (N=2,462 transcripts), and CVs for all assembled transcripts in each sample (NPC rep 1=14,318 transcripts, NPC rep 2=11,582 transcripts). Transcripts are binned into no expression (log-RPKM = 0; black), very low expression (0 < log- RPKM < 0.1; red), low expression (0.1 ≤ log- RPKM < 0.6; orange), moderate expression (0.6 ≤ log-TPM < 1.6; yellow), and high expression (log-TPM ≥ 1.6; blue).

**References**

1. Roberts A, Pachter L: **Streaming fragment assignment for real-time analysis of sequencing experiments**. *Nature Methods* 2013, **10**(1):71-U99.

2. Li H: **Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM**. *arXic preprint arXiv:13033997* 2013.

3. Mezlini AM, Smith EJM, Fiume M, Buske O, Savich GL, Shah S, Aparicio S, Chiang DY, Goldenberg A, Brudno M: **iReckon: Simultaneous isoform discovery and abundance estimation from RNA-seq data**. *Genome Research* 2013, **23**(3):519-529.