**SUPPLEMENTARY FIGURES**

**Supplementary Figure 1**: Non-N B6Eve coverage of GRCm38 chromosome assemblies showing that all the GRCm38 chromosomes (excluding chrY as we have female mouse) have ≥ 90% coverage in B6Eve assembly

**Supplementary Figure 2: a)** Figure showing IGV snapshot of B6Eve scaffold (#8912) partially filling a gap sized at 1,760 bp in chr2 (nt 172,624,657-172,626,416) with a deletion of 136 bp at the beginning of the gap **b)** Figure showing IGV snapshot of B6Eve scaffold (#2171) completely filling a gap of size 99bp in chr1 (183334907-183335006) by inserting 595bp at the beginning of the gap.​

**Supplementary Figure 3:** The gene *Traf5* has a 50kb assembly gap between the last three exons and the remainder of the transcript in GRCm38 (bottom panel). In the B6Eve assembly, this gap is closed to 2,430 bp. Additionally, IsoSeq shows that this gap overlaps an exon that was included as a new isoform in the CAT annotation. BLAT alignment of the isoform back to the reference shows the inserted sequence.

 **Supplementary Figure 4:** The gene *Slc26a6* has a 50kb assembly gap between the last exon and the remainder of the transcript in GRCm38 (bottom panel). In the B6Eve assembly, this gap is closed to 1,810bp (top panel). Additionally, IsoSeq shows that this gap overlaps an exon and removes an additional exon. Using the IsoSeq information CAT generated a new isoform that matches the IsoSeq alignment.

**SUPPLEMENTARY TABLES**

**Supplementary Table 1:** Summary PacBio and Illumina sequencing data.

**Supplementary Table 2:** Progress of Assembly at each step

**Supplementary Table 3:** Details for RefSeq transcript alignments to B6Eve and GRCm38

**Supplementary Table 4:** Coordinates of filled gaps with sequence in GRCm38 with B6Eve assembly

**Supplementary Table 5:** Locus of resolved variants in Exome, CC and Pedigree data, respectively

**Supplementary Table 6:** Coordinates of SVs common between Illumina and PacBio based analysis

**Supplementary Table 7a:** The complete distribution of classification of repeats in GRCm38

**Supplementary Table 7b**: The complete distribution of classification of repeats of B6Eve assembly

**Supplementary Table 7c:** The complete distribution of classification of repeats of Illumina GCA\_000185125

**Supplementary Table 7d:** The complete distribution of classification of repeats of Illumina GCA\_000185105

**Supplementary Table 8:** Coordinates of B6Eve splice site shifts relative to GRCm38, potential novel exons and loci

**SUPPLEMENTARY FILES**

**Supplementary File 2:** A file (FileS5\_LXEJ02\_contigs.tsv) with mapping of B6Eve scaffold names to GenBank accessions.

**Supplementary File 3:** Variant calls (in VCF file) made using the Illumina data of B6Eve to Pilon corrected B6Eve assembly

**Supplementary File 4:** Distributions of alternate allele frequency, genotype called, and total depth relationship for high quality variants called from alignment of B6Eve Illumina data to the B6Eve assembly. Data are binned by depth <=5 reads (1), >5 and <=10 reads (2), >10 reads and <=20 reads (3), >20 reads and <=30 reads (4) and >30 reads (5)

**Supplementary File 5:** High quality 227,523 calls lift over position in GRCm38 and annotation of these variants.

**Supplementary File 6:** Non-random distribution of SNP calls from alignment of B6Eve Illumina data to the B6Eve assembly.

**Supplementary File 7:** Common SNP variant callsfrom alignment of B6Eve Illumina data mapped to the B6Eve assembly and GRCm38.

**Supplementary File 8:** A detailed QUAST evaluation report (HTML format) reflecting the quality of the polished assembly

**Supplementary File 9:** Regions of the B6Eve assembly (BED file) remained unaligned to GRCm38 by a) Cactus based alignment using UCSC Comparative Annotation Toolkit b) NCBI assembly-assembly alignments and c) QUAST (minimap2 aligner)