IsoTools – a flexible workflow for long-read transcriptome sequencing analysis
Supplementary Figures

Matthias Lienhard[1,*], Twan van den Beucken[2], Bernd Timmermann[3], Myriam Hochradel[3], Stefan Boerno[3], Florian Caiment[2], Martin Vingron[1], and Ralf Herwig[1,*]

[1]*Max Planck Institute for Molecular Genetics, Department of Computational Biology, Ihnestraße 63-73*
[2]*Maastricht University, Department of Toxicogenomics, Universiteitssingel 40, 6229ER Maastricht, NETHERLANDS*
[3]*Max Planck Institute for Molecular Genetics, Sequencing core unit, Ihnestraße 63-73, 14195 Berlin, GERMANY*
[*]*To whom correspondence should be addressed.*
ML: Tel: +4930 8413 1675; Email: lienhard@molgen.mpg.de
RH: Tel: +4930 8413 1126; Fax: +4930 8413 1152; Email: herwig@molgen.mpg.de

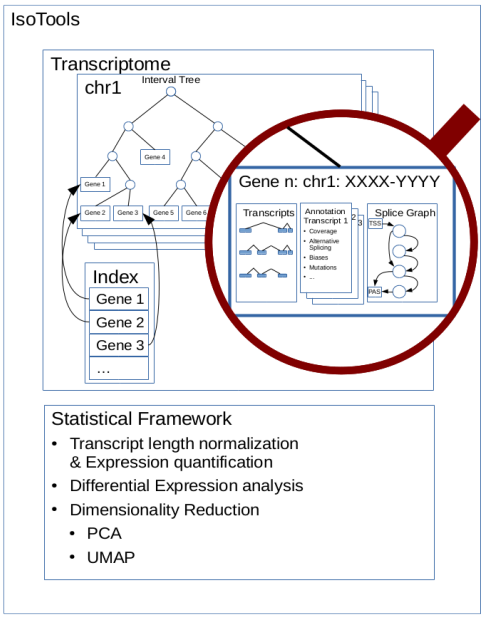

Figure S1: The internal data structure of the IsoTools framework

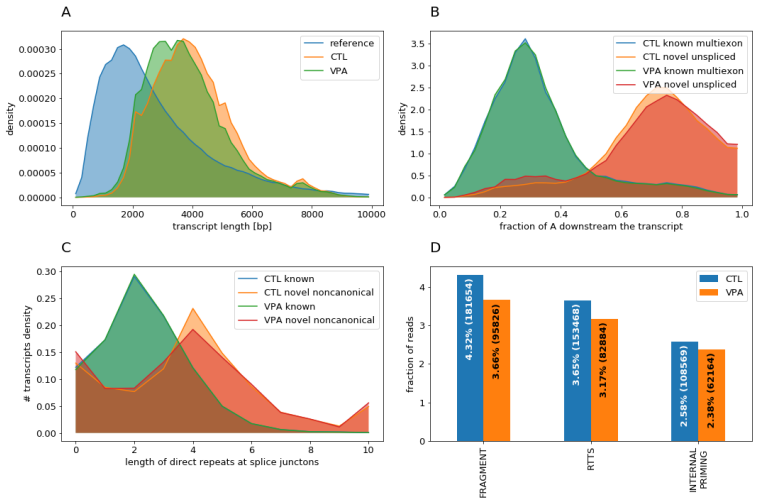

Figure S2: Quality control metrics. A) Read length distribution compared to level 1 GENCODE transcripts. B) A content downstream of novel unspliced transcripts and reference matching multi-exon transcripts. C) Direct repeat length at intron boundaries of GENCODE transcripts and novel non-canonical splice junctions. D) Fraction of reads affected by one of the three artifacts.

Figure S3: Prototypical examples for novel transcript classes. Reference transcripts are depicted in blue, and classified novel transcripts in green. Labels indicate the novelty class, with the Sqanti classification in brackets.

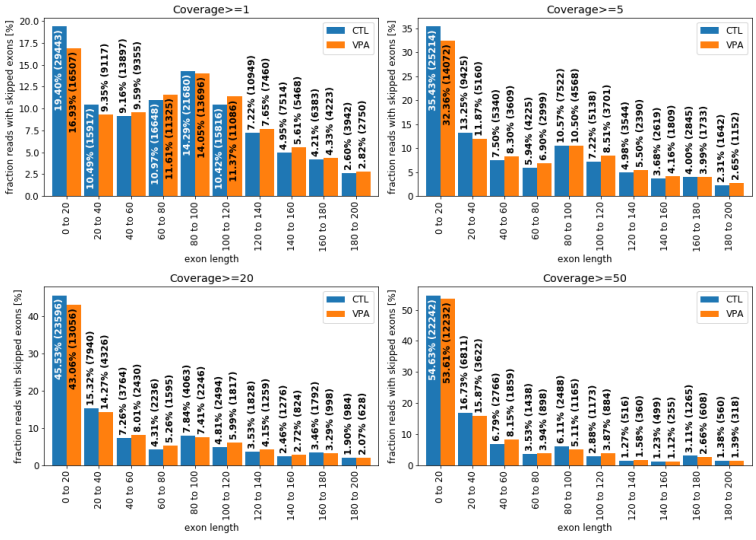

Figure S4: Histogram of novel exon skipping transcripts for different ranges of exon lengths. Short exons are not correctly aligned by the alignment tool, resulting in the detection of false skipped exons.

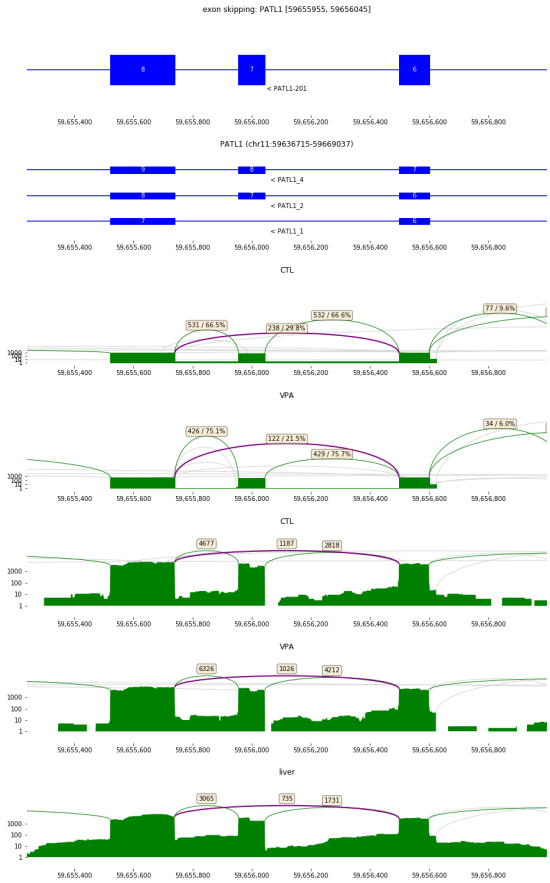

Figure S5: The novel isoform of PATL1 skipps exon 7 of the GENCODE annotation (top track). Relevant isoforms detected by IsoSeq are depicted in the second track. The following two Sashimi plots depict the IsoSeq coverage for the hepatocytes (CTL and VPA), and the bottom three Sashimi plots Illumina RNA-Seq coverage. The exon skipping junction is depicted in purple, other highly covered junctions in green, and poorly covered junctions in grey.

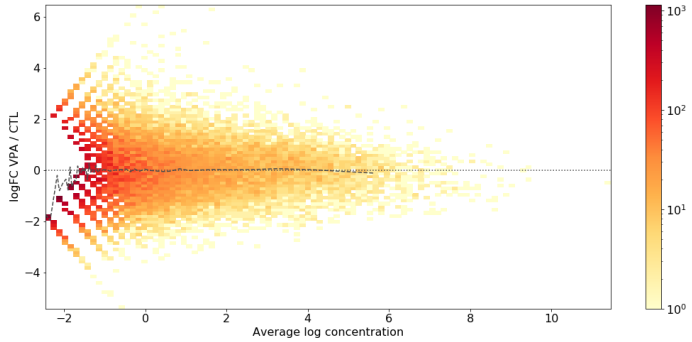

Figure S6: MA plot of upper quantile normalized IsoSeq transcript read counts in VPA vs CTL treated hepatocytes. Dashed line represent the binned average of logFCs, within equally sized bins of 500 transcripts.

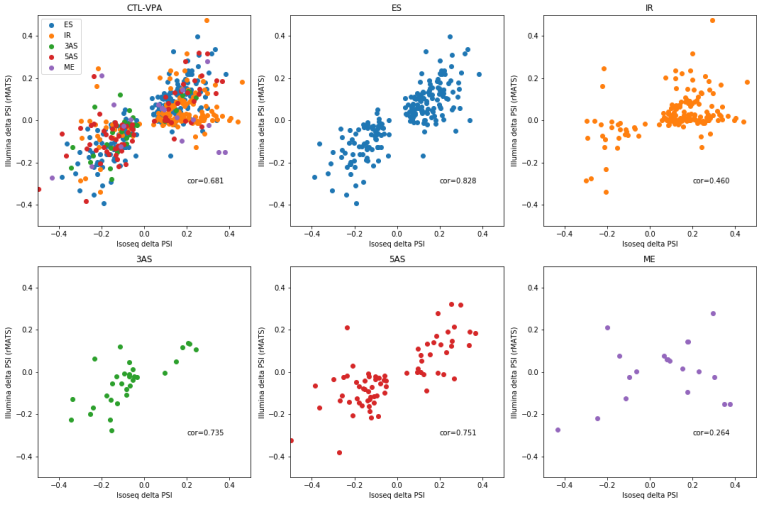

Figure S7: Comparison PSI differences between VPA treated and control samples for differential ASE (two-proportions z-test $FDR < .01$), quantified by IsoTools from LRTS and rMATS from short read RNAseq, for all events and the splicing event classes individually.

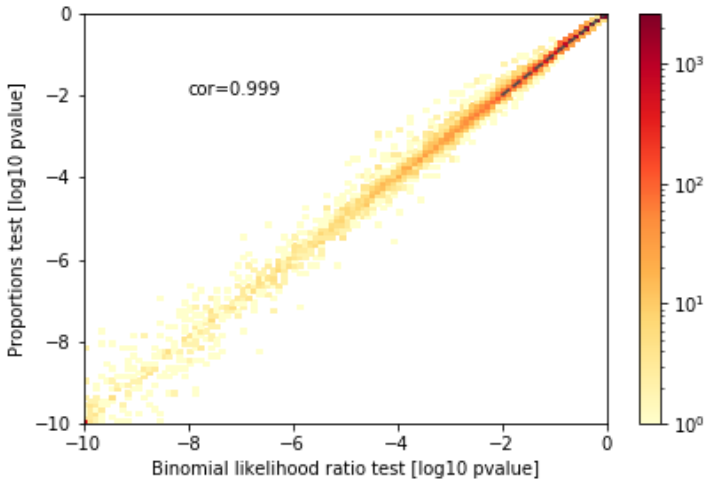

Figure S8: Comparison of binomial likelihood ratio test and two proportions z-test, for VPA vs. CTL samples.