Fig. 1. Overview of the Long-read RNA-seq Genome Annotation Assessment Project (LRGASP). a, Data produced for LRGASP consists of multiple species, multiple sample types, multiple library protocols, and multiple sequencing platforms for comparison. **b,** Distribution of read lengths, identify Q score, and sequencing depth (per biological replicate) for the WTC11 sample. **c**, LRGASP as an open research community effort for benchmarking and evaluating long-read RNA-seq approaches. **d,** Number of isoforms reported by each tool on different data types for the human WTC11 sample for Challenge 1. **e**, Median TPM value reported by each tool on different data types for the human WTC11 sample for Challenge 2. **f,** Number of isoforms reported by each tool on different data types for the mouse ES data for Challenge 3. **g,** Pairwise relative overlap of unique junction chains (UJCs) reported by each submission. The UJCs reported by a submission is used as a reference set for each row. The fraction of overlap of UJCs from the column submission is shown as a heatmap. For example, a submission that has a small, subset of many other UJCs from other submissions will have a high fraction shown in the rows, but low fraction by column for that submission. Data only shown for WTC11 submissions. **h,** Spearman correlation of TPM values between submissions to Challenge 2. **i,** Pairwise relative overlap of UJCs reported by each submission. The UJCs reported by a submission is used as a reference set for each row. The fraction of overlap of UJCs from the column submission is shown as a heatmap.