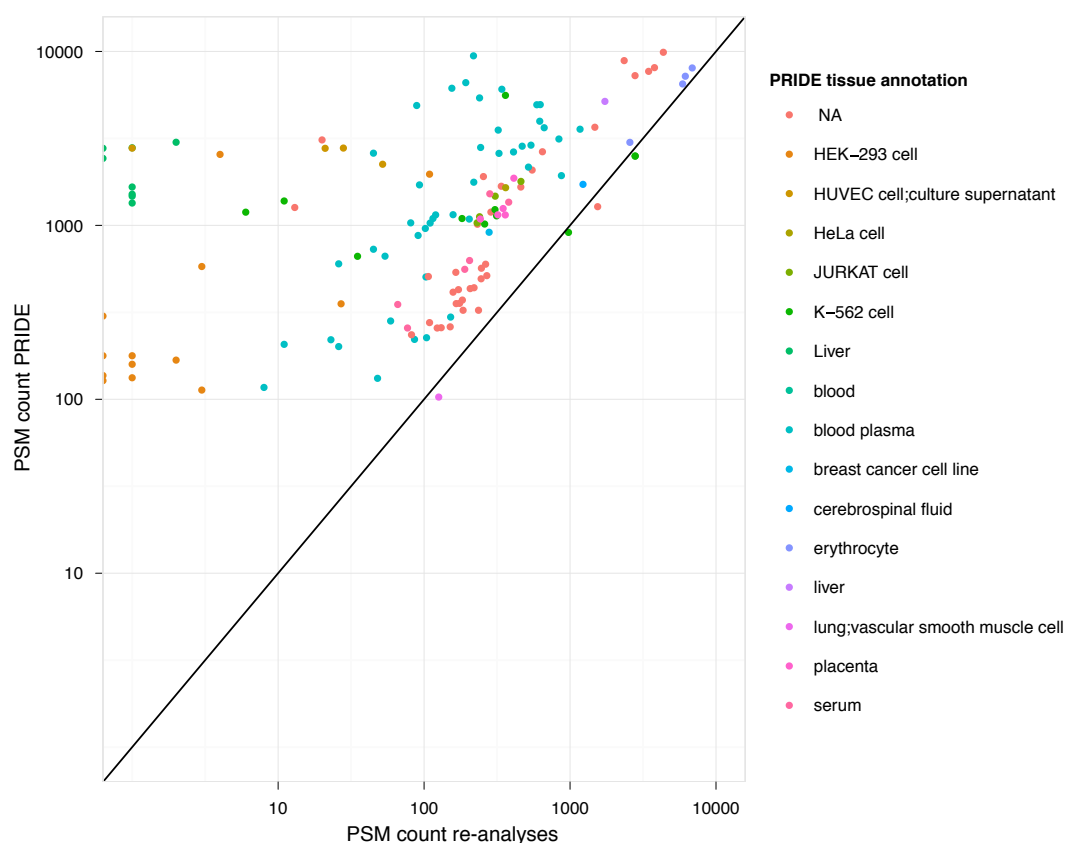
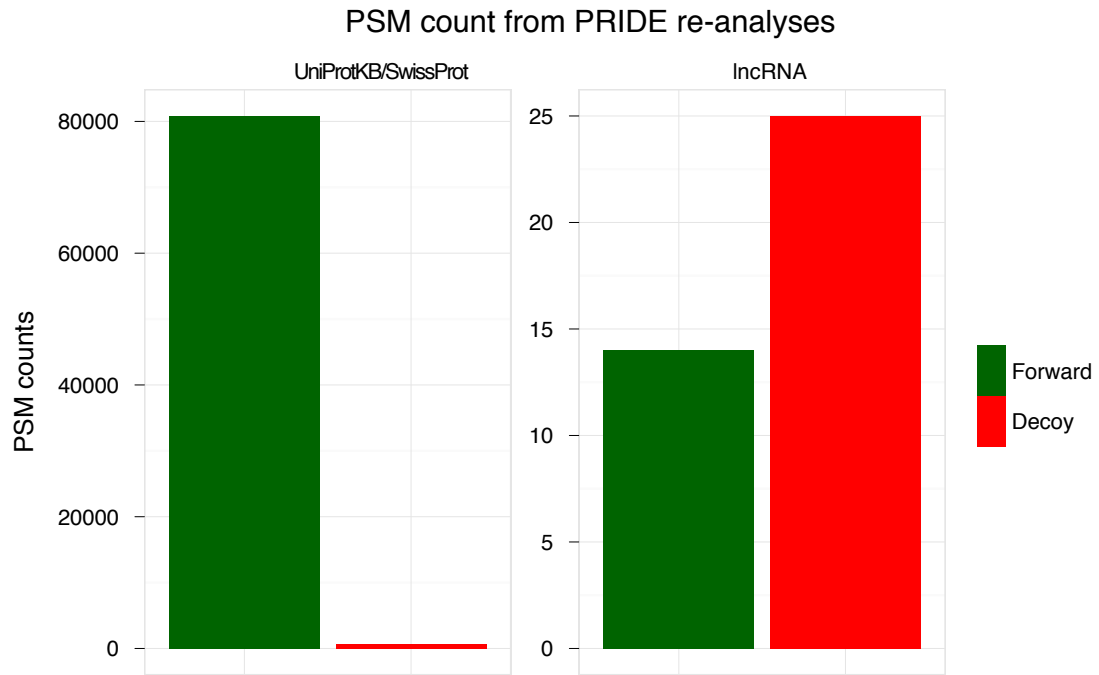


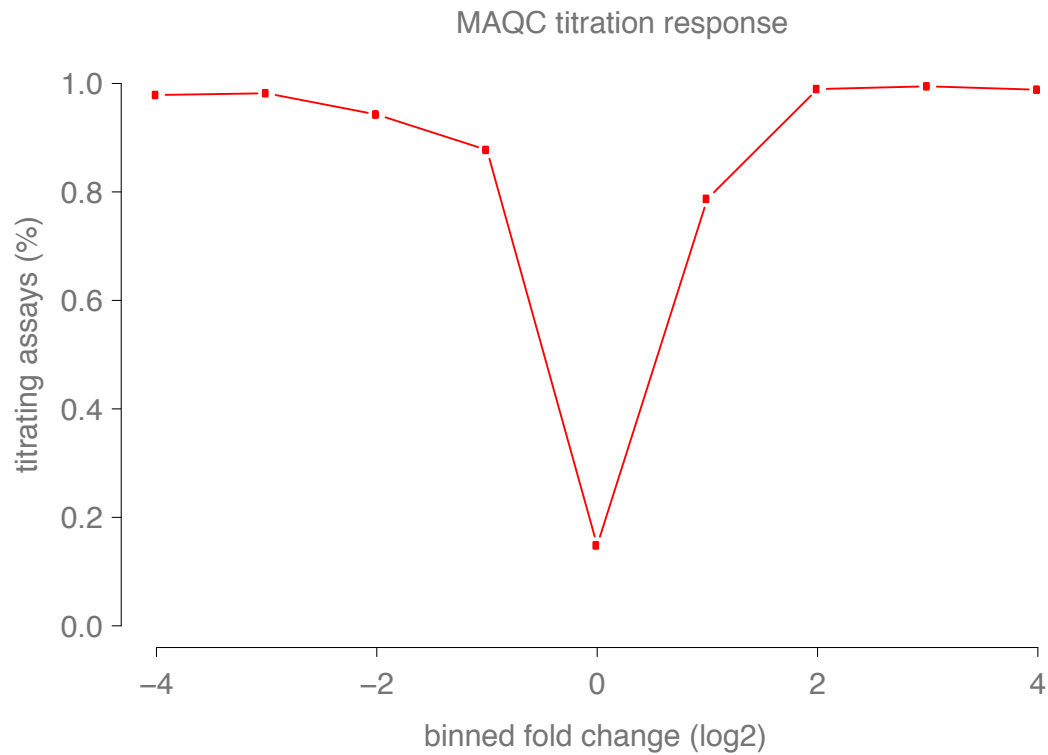
## Supplemental Figures



Supplemental Figure 1. Scatterplot of re-analysis of 149 PRIDE experiments. The X-axis and the Y-axis show the number of peptide-to-spectrum matches (PSMs) *per* experiment by our automated re-analysis and as deposited in PRIDE, respectively. For most experiments, our analysis yields roughly half the amount of PSMs annotated in PRIDE. One reason for this is that our approach applies a stringent 1% FDR cutoff, while such stringency is not required when depositing an experiment into PRIDE. Furthermore, our protein sequence database is considerably larger than UniProtKB/Swiss-Prot since it contains a translated version of Incipedia. This inherently leads to larger e-values, and thus less PSMs in our stringent results set.



Supplemental Figure 2. PSM counts after re-analysis of 149 PRIDE experiments with an FDR limit of 1%. PSMs from UniProtKB/Swiss-Prot are called 'false' (left bar chart) and from lncRNA translations are dubbed 'true' (right bar chart). Decoy hits, indicative of the amount of false positives, are given in red, while normal hits are given in green. Note that while the left bar chart with UniprotKB/Swiss-Prot hits shows an expected FDR of 1%, the right bar chart with PSMs from lncipedia translations shows a much larger FDR of 166%.



Supplemental Figure 3. MAQC titration response of lncRNA probes. lncRNA expression was measured for samples A (Universal human reference RNA, Agilent Technologies), B (Human brain total RNA, Ambion), C (25% A + 75% B) and D (75% A + 25% B). The percentage of lncRNA probes that follow the monotonic titration response (Y-axis) is plotted in function of the binned log<sub>2</sub>-fold change (X-axis) between samples A and B. Titration response was calculated according to Shippy et al., Nature Biotechnology, 2006.

## **Supplemental Methods**

Data is read from the PRIDE database after filtering applicable experiments by taxonomy, number of spectra and consistent taxonomic origin of the reported proteins. The data is then analyzed to detect applicable search engine settings, notably the precursor and fragment ion mass tolerances as well as the (variable) modifications to consider. Allowed missed cleavages are set to 1. PeptideShaker is run in automatic mode to filter the proposed peptide-to-spectrum matches hits at the 1% false discovery rate as calculated through the decoy database searching built-in to SearchGUI.