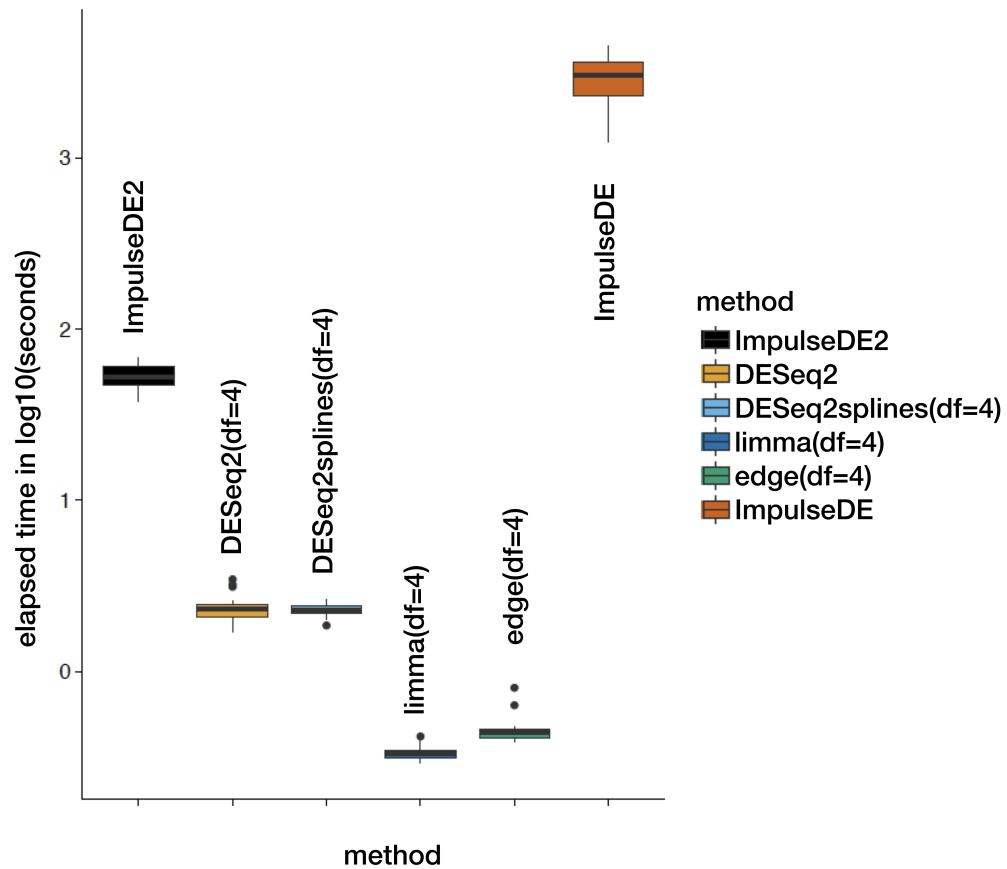


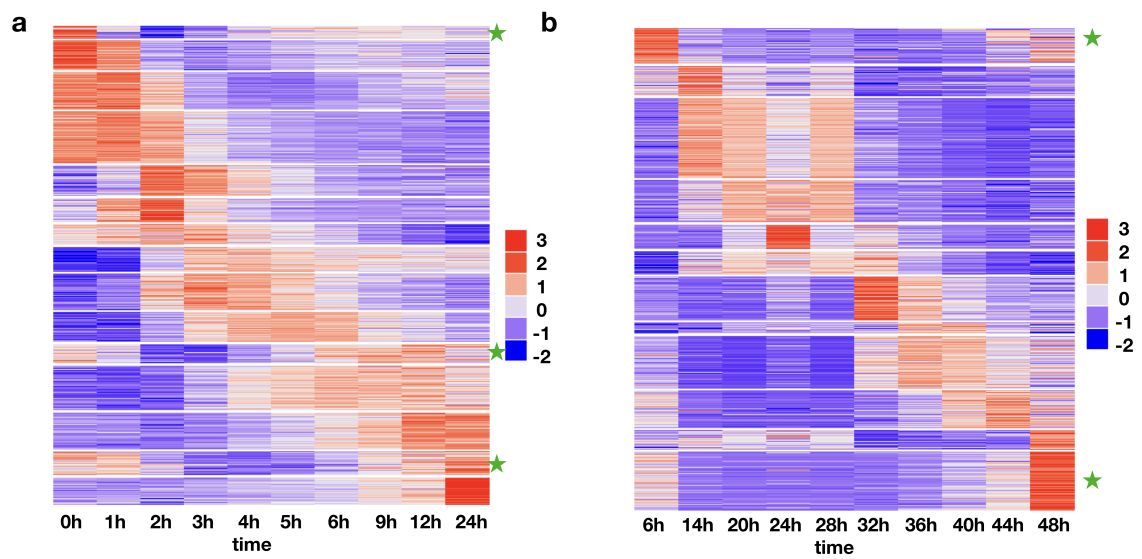
Supplementary Figures: Impulse model-based differential expression analysis of time course sequencing data

David S. Fischer, Fabian J. Theis, Nir Yosef

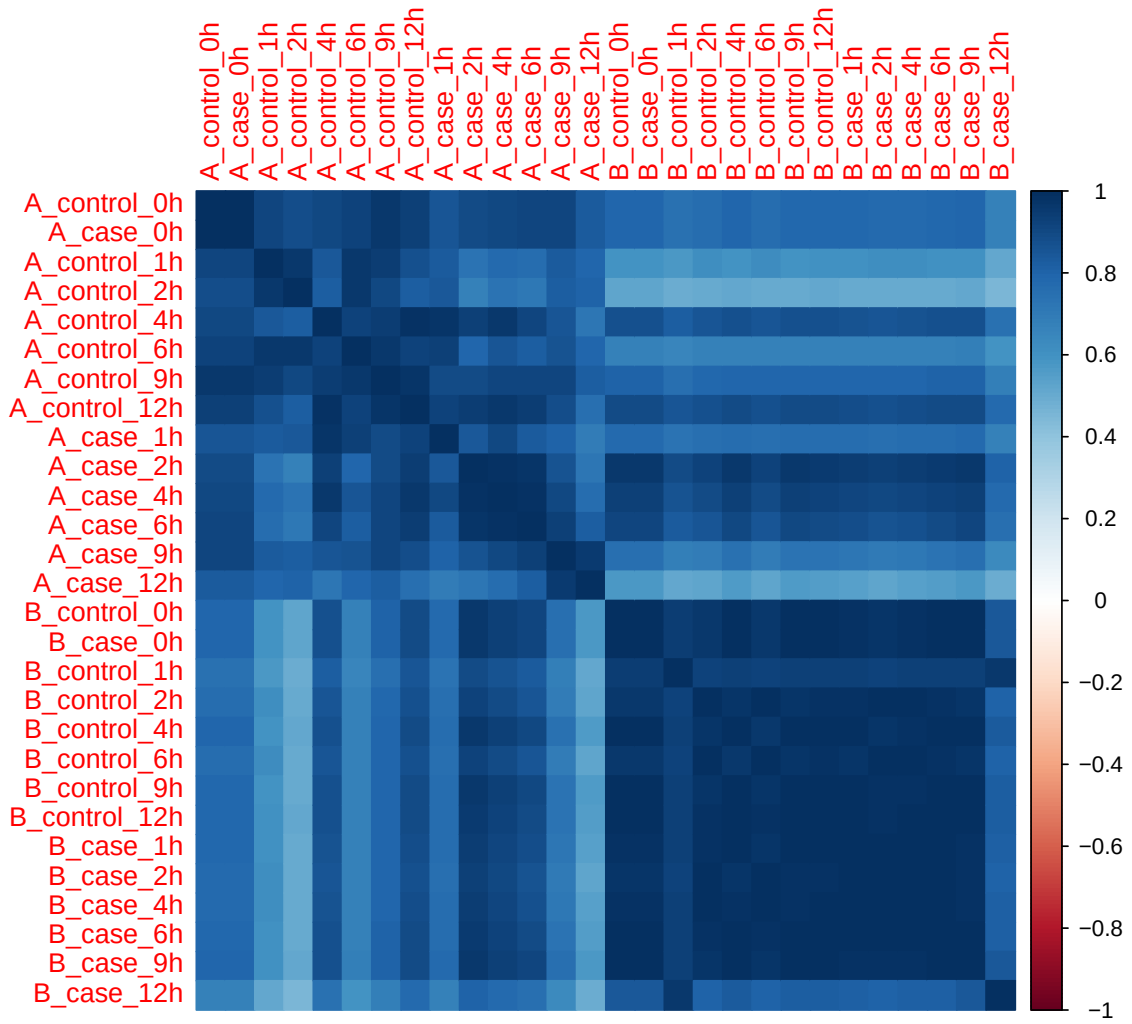
June 20, 2018



Supplementary Figure 1: **Run time comparison of all differential expression methods considered.** Shown are boxplots of the time in seconds required to run each method on the simulated data presented in Fig. 2a. The time shown is elapsed clock time, not net CPU time. ImpulseDE2, DESeq2, DESeq2splines and ImpulseDE were parallelized with 10 cores, edge and limma were run without parallelization.

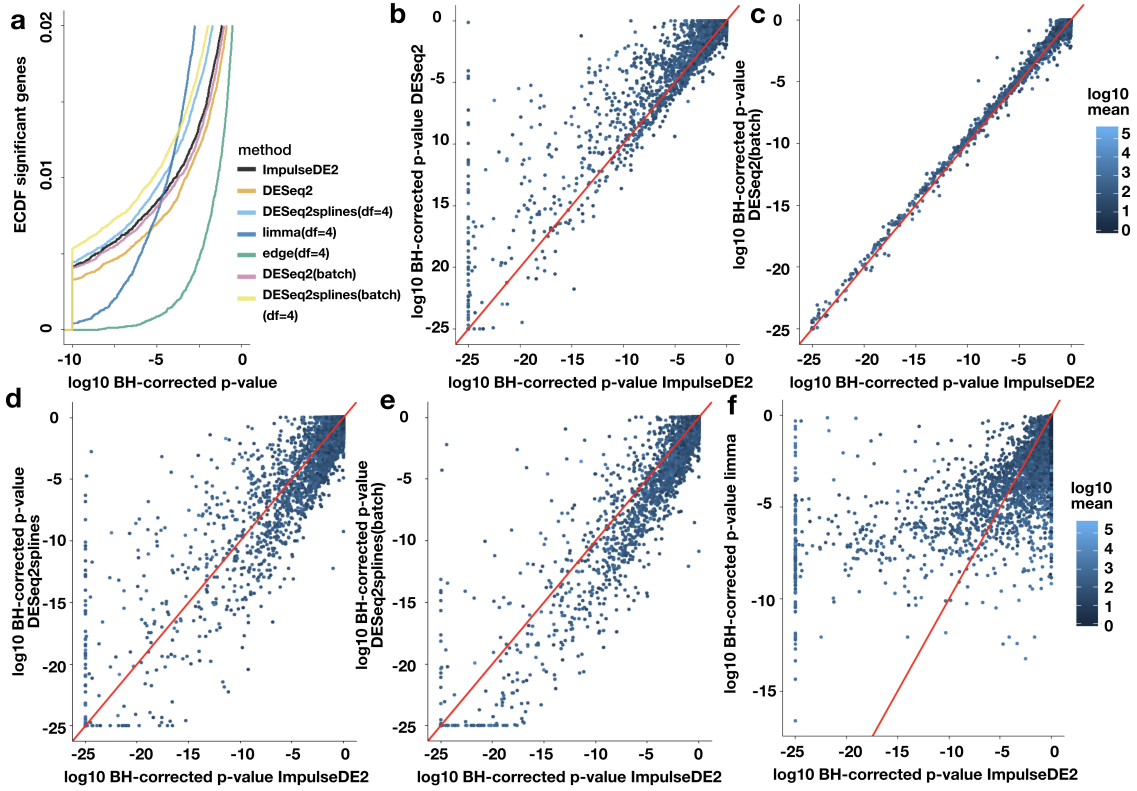


Supplementary Figure 2: **The impulse model is descriptive of global transcriptome dynamics: Additional data sets.** Heatmaps of z-scores of library depth normalized mean counts per time point of differentially expressed genes selected with DESeq2. Green stars indicate clusters that can be modelled with the valley model. **a** RNA-seq of the estrogen response of a human breast cancer cell line ("estrogen (Baran-Gale)"). **b** RNA-seq of long-noncoding RNAs of *Plasmodium falciparum* during infection("Plasmodium (Broadbent)").

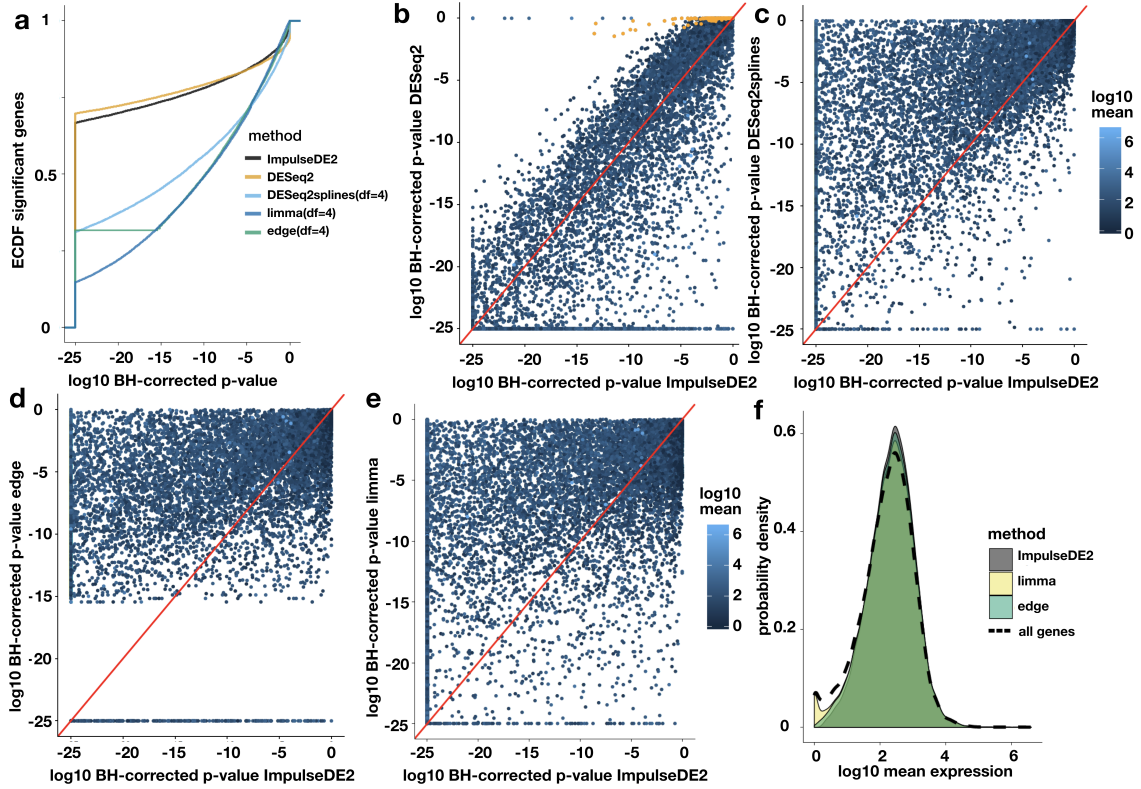


Supplementary Figure 3: **Correlation structure and batch effects of LPS (Jovanovic) data set.** Batch: A, B.

These groups of samples were handled together, note that they overlap case and control condition. Condition: "case" (case condition with LPS) and "ctrl" (control without LPS treatment). Time point 0h was only sampled in the control condition and was replicated for the case condition in each batch for analysis. Shown are Pearson correlation coefficients of the untransformed count vectors.

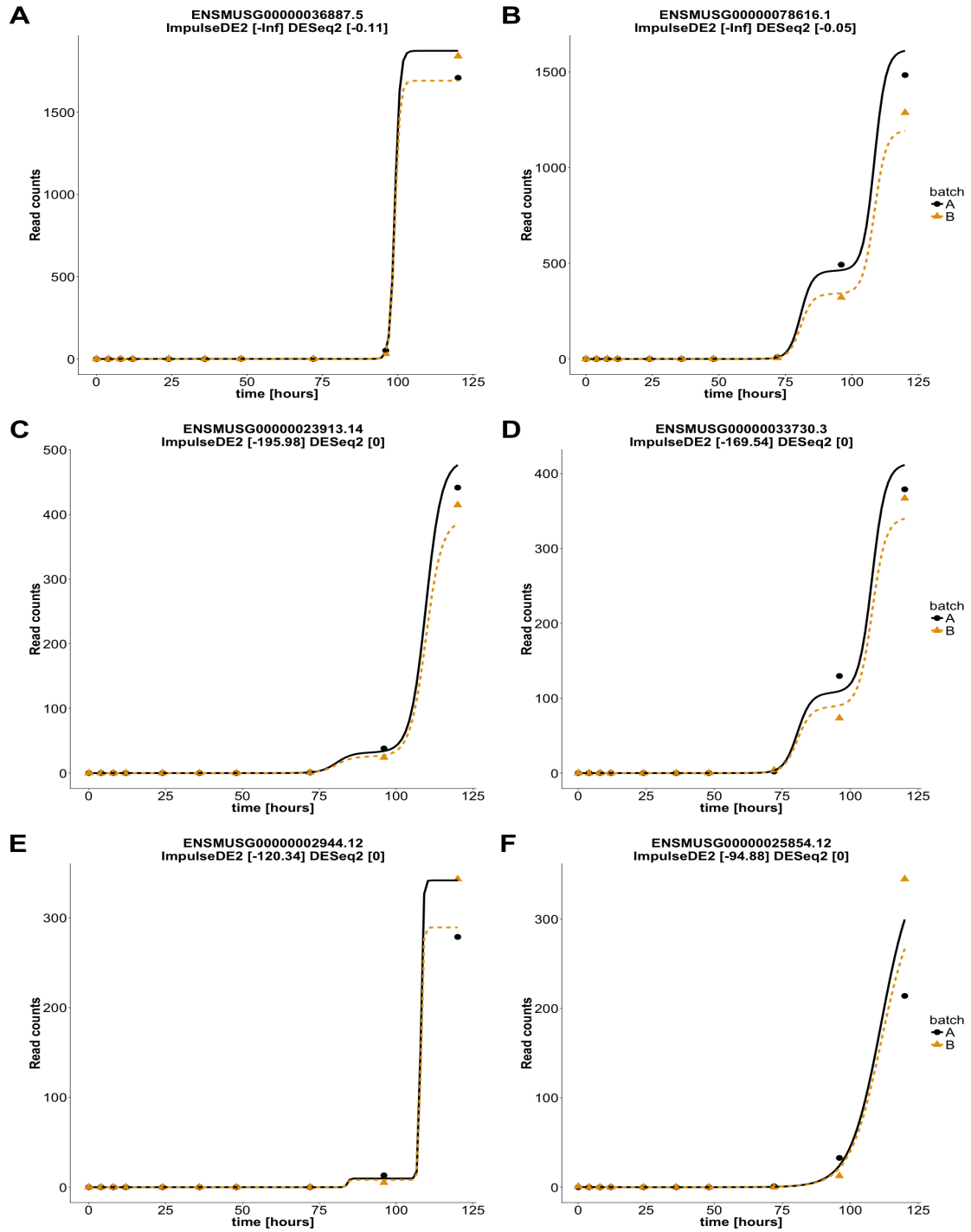


Supplementary Figure 4: **Comparison of ImpulseDE2, limma and DESeq2 on LPS (Jovanovic) case-control data.** DESeq2 and DESeq2splines: standard batch settings, DESeq2(batch) and DESeq2splines(batch): condition-wise batch correction, limma(df=4): limma with spline model with four degrees of freedom. **a** Fraction of significantly differentially expressed genes as a function of the significance threshold by method (case-control analysis). **b,c,d,e,f** Scatter plots of the inferred differential expression (case-control) Benjamini-Hochberg (BH) corrected p-values for all genes between ImpulseDE2 and the reference methods: DESeq2 (**b**), DESeq2(batch) (**c**), DESeq2splines (**d**) and DESeq2splines(batch) (**e**) and limma (**f**). The UpSetR plot for this data set is supplied in Supp. Fig. 11.

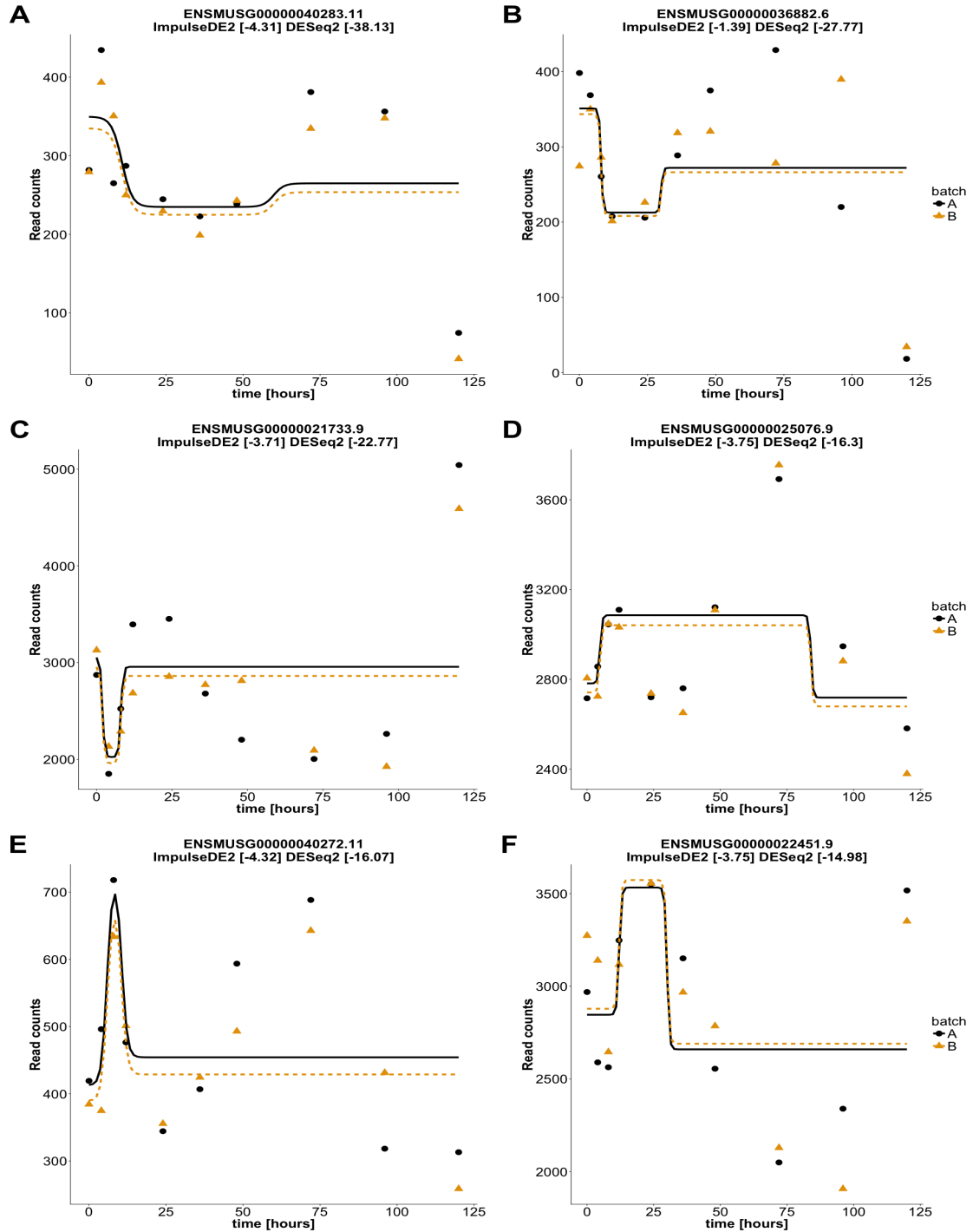


Supplementary Figure 5: **Comparison of ImpulseDE2, DESeq2, limma and edge on Drosophila (Graveley)**

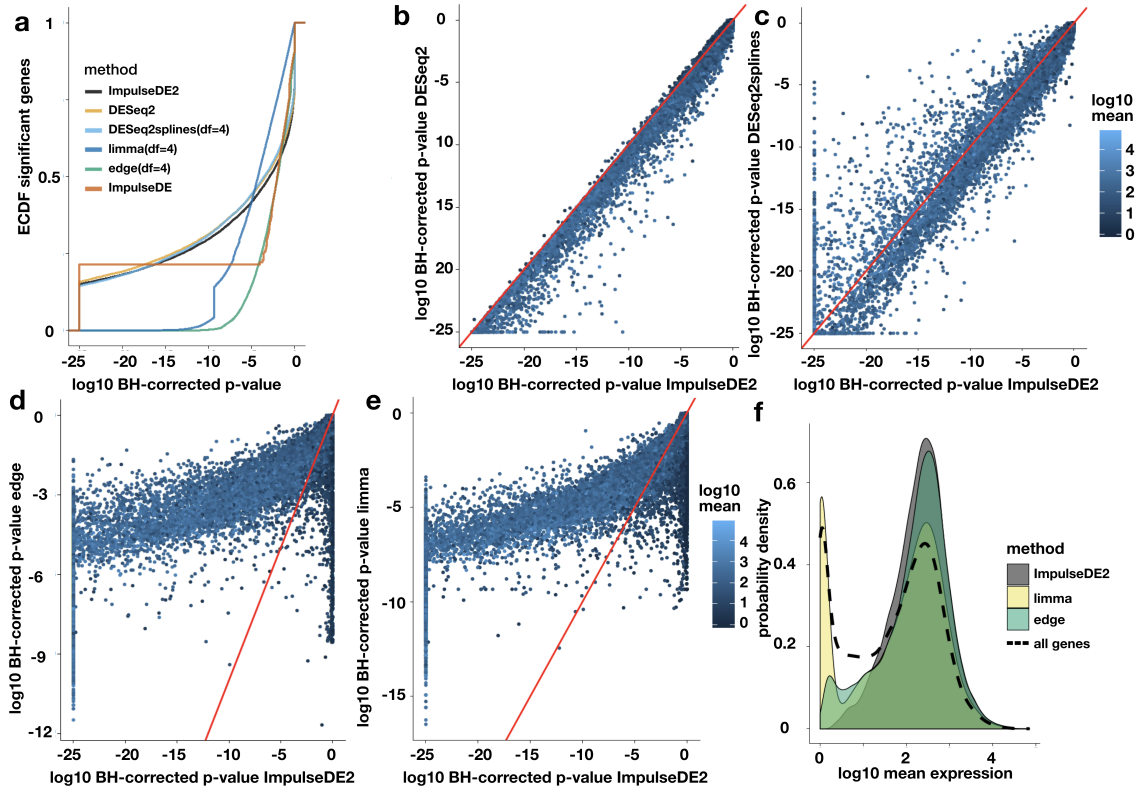
data set. **a** Fraction of significantly differentially expressed genes as a function of the significance threshold by method. **b,c,d,e** Correlation plot of the inferred differential expression Benjamini-Hochberg (BH) corrected p-values for all genes between ImpulseDE2 and DESeq2 (**b**), DESeq2splines (**c**), edge (**d**) and limma (**e**). Orange points correspond to genes for which ImpulseDE2 disabled DESeq2 dispersion outlier handling. **f** Kernel density estimate of density of the distribution of expression means of genes called differentially expressed at a q-value threshold of $1e - 2$ of ImpulseDE2, limma and edge. The mean expression distribution across all genes is shown as *all genes*. The UpSetR plot for this data set is supplied in Supp. Fig. 12.



Supplementary Figure 6: Sykes et al. RNA-seq (case-only): ImpulseDE2 versus DESeq2: Examples gene with lower p-value assigned by ImpulseDE2 than DESeq2. Title: Method[Benjamini-Hochberg corrected log10 p-value]. Shown are read counts normalized by size factors and impulse model fits. The shown genes are the genes with the lowest ratio (equivalent to the highest difference) of Benjamini-Hochberg corrected p-values between the two compared methods out of the genes which were assigned a corrected p-value higher than a threshold of $1e-5$ by the method which was in the denominator of the comparison (the method which assigned the higher corrected p-value for the shown genes). All genes shown here were labelled as dispersion outliers by DESeq2 with the standard settings.

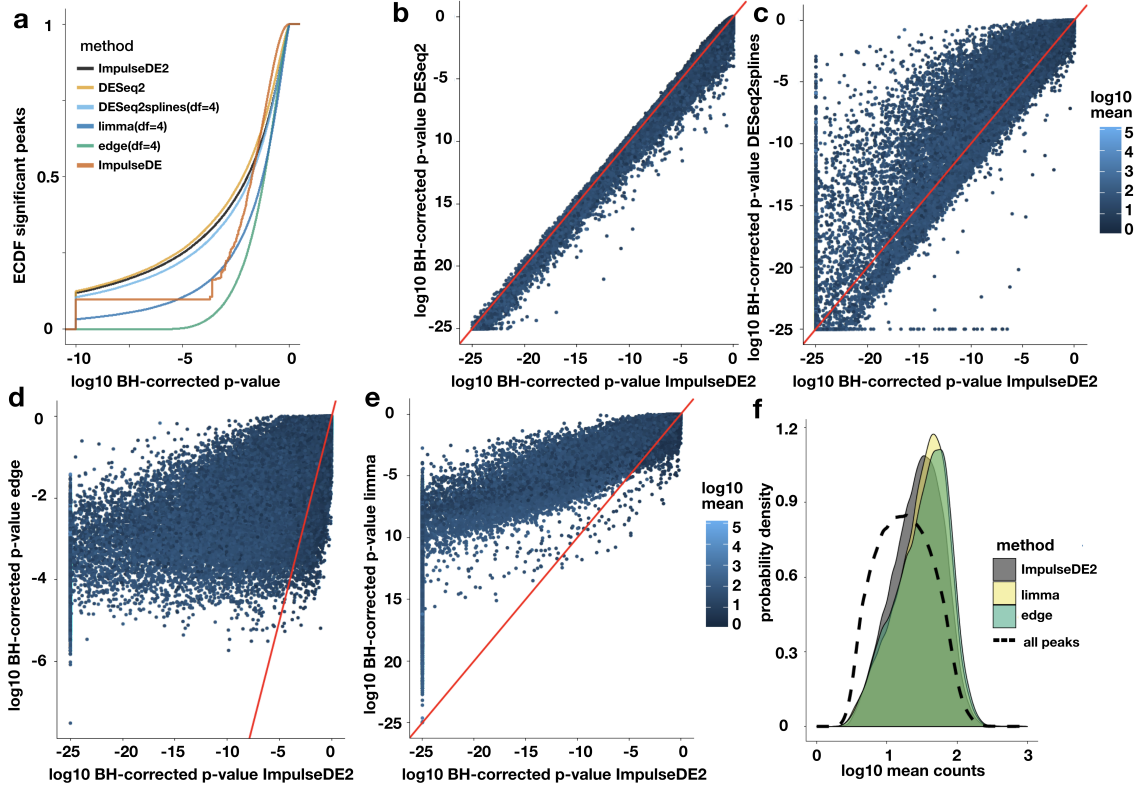


Supplementary Figure 7: Sykes et al. RNA-seq (case-only): ImpulseDE2 versus DESeq2: Examples gene with lower p-value assigned by DESeq2 than ImpulseDE2. Title: Method[Benjamini-Hochberg corrected log₁₀ p-value]. Shown are read counts normalized by size factors and impulse model fits. The shown genes are the genes with the lowest ratio (equivalent to the highest difference) of Benjamini-Hochberg corrected p-values between the two compared methods out of the genes which were assigned a corrected p-value higher than a threshold of $1e-5$ by the method which was in the denominator of the comparison (the method which assigned the higher corrected p-value for the shown genes).

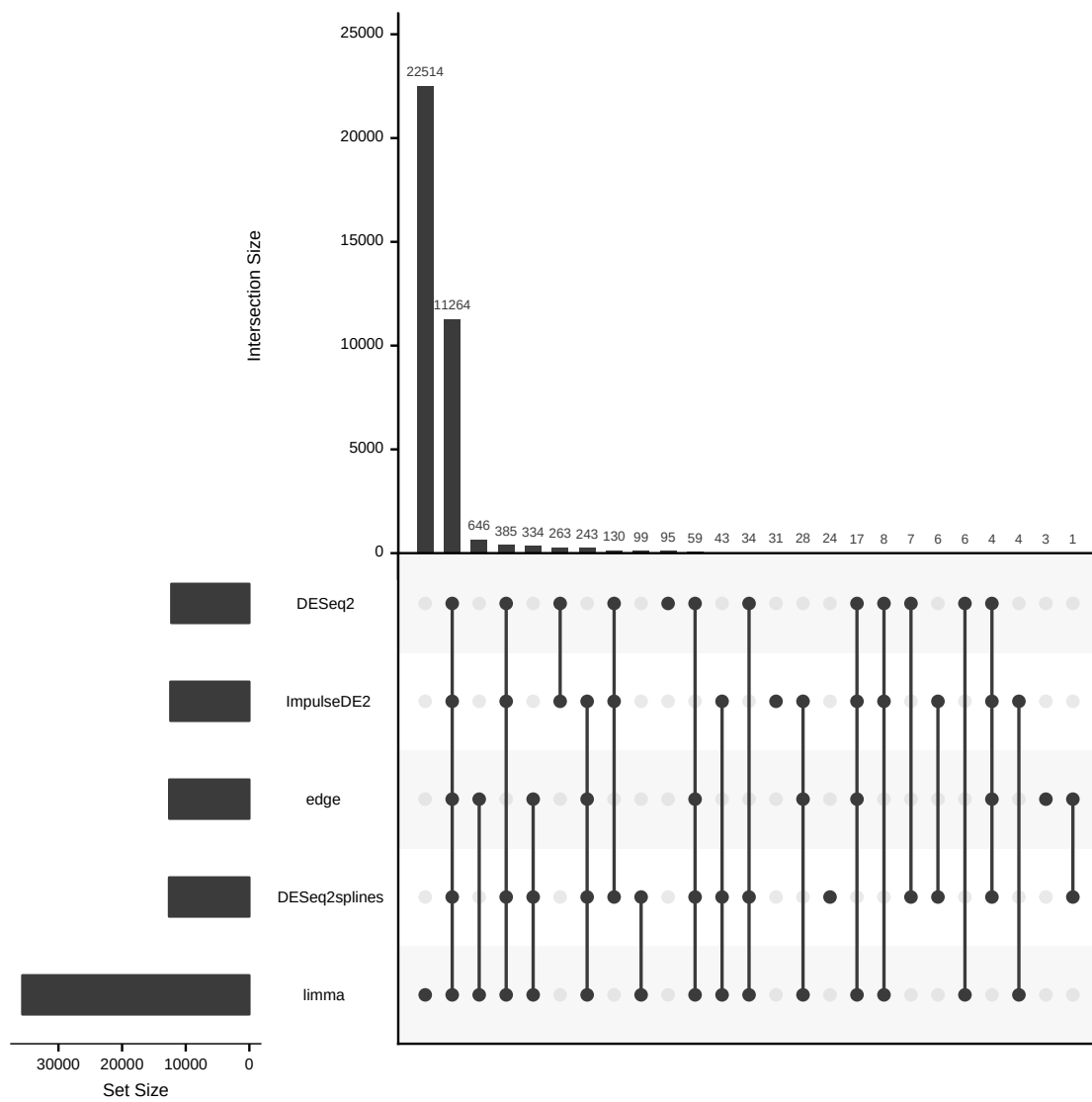


Supplementary Figure 8: Comparison of ImpulseDE2, DESeq2, limma and edge on hESC (Chu) data set.

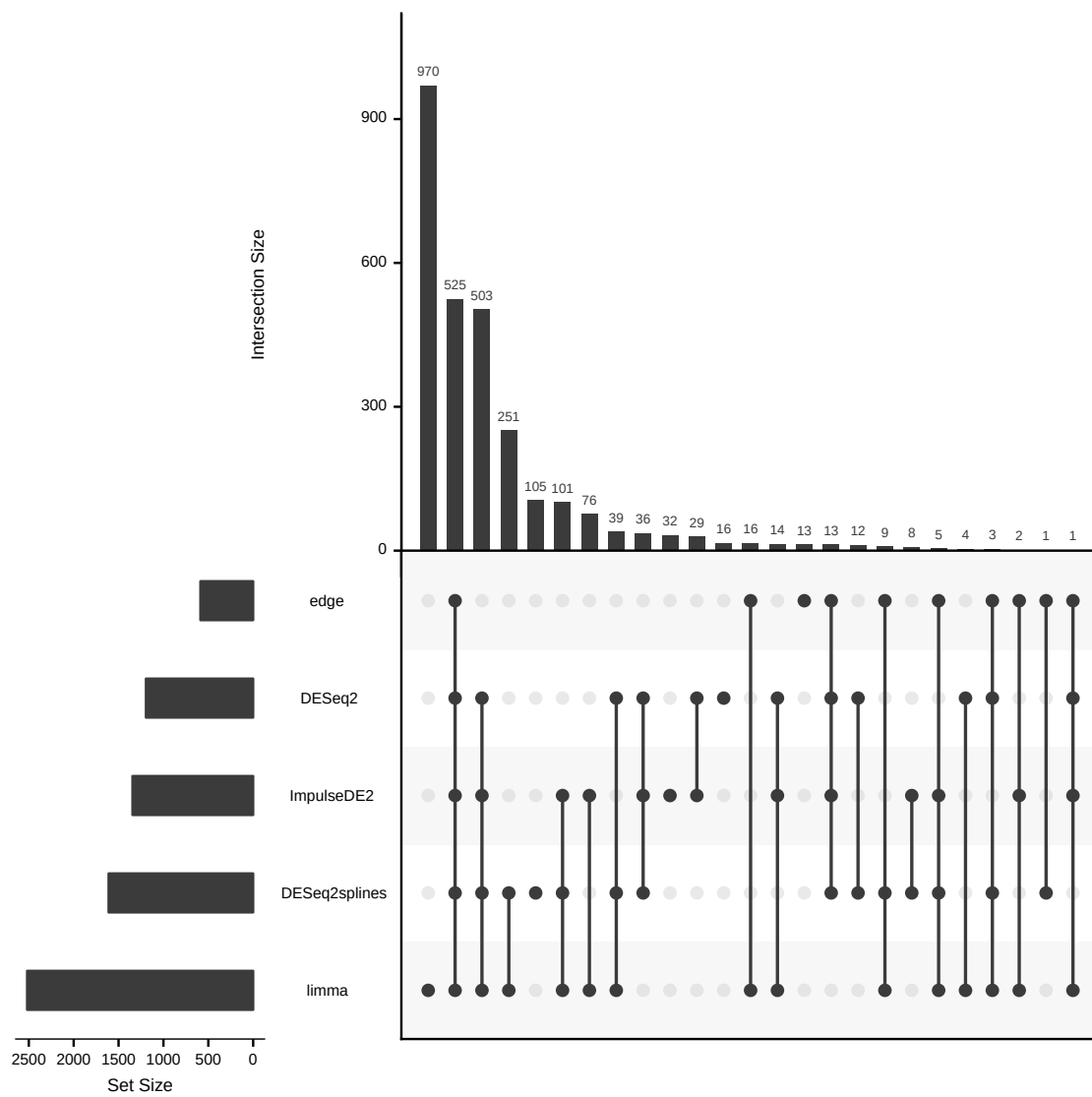
Orange points correspond to genes for which ImpulseDE2 disabled DESeq2 dispersion outlier handling. **a** Fraction of significantly differentially expressed genes as a function of the significance threshold by method. **b,c,d,e** Correlation plot of the inferred differential expression Benjamini-Hochberg (BH) corrected p-values for all genes between ImpulseDE2 and DESeq2 (**b**), DESeq2splines (**c**), edge (**d**) and limma (**e**). **f** Kernel density estimate of density of the distribution of expression means of genes called differentially expressed at a q-value threshold of $1e-2$ of ImpulseDE2, limma and edge. The mean expression distribution across all genes is shown as *all genes*. The UpSetR plot for this data set is supplied in Supp. Fig. 13.



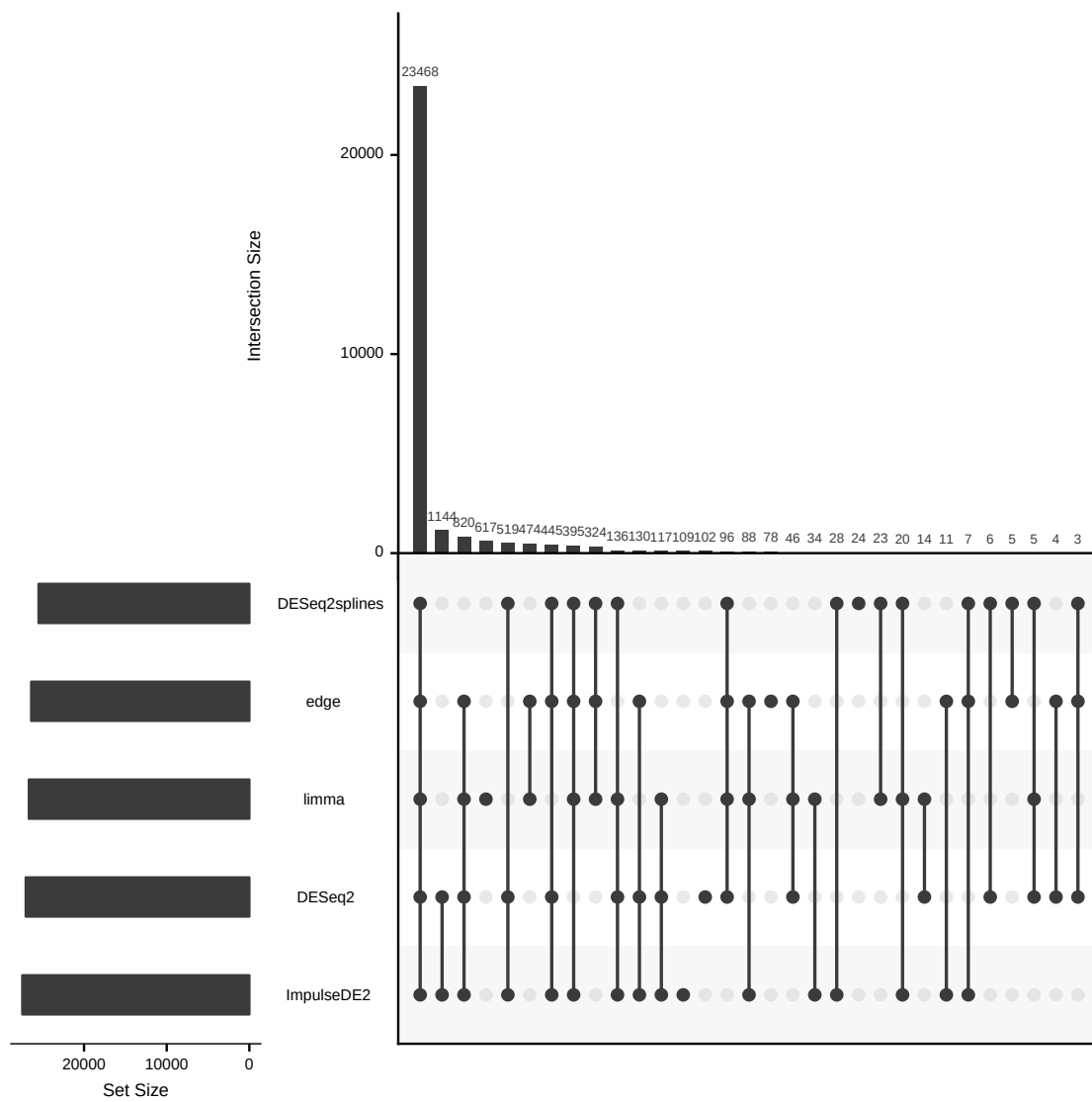
Supplementary Figure 9: **Comparison of ImpulseDE2, DESeq2, limma and edge on iChIP (Lara-Astiaso) data set.** **a** Fraction of significantly differentially expressed genes as a function of the significance threshold by method. **b,c,d,e** Correlation plot of the inferred differential expression Benjamini-Hochberg (BH) corrected p-values for all genes between ImpulseDE2 and DESeq2 (**b**), DESeq2splines (**c**), edge (**d**) and limma (**e**). **f** Kernel density estimate of density of the distribution of expression signal per peak called differentially expressed at a q-value threshold of $1e-2$ of ImpulseDE2, limma and edge. The mean signal distribution across all peaks is shown as *all peaks*. The UpSetR plot for this data set is supplied in Supp. Fig. 14.



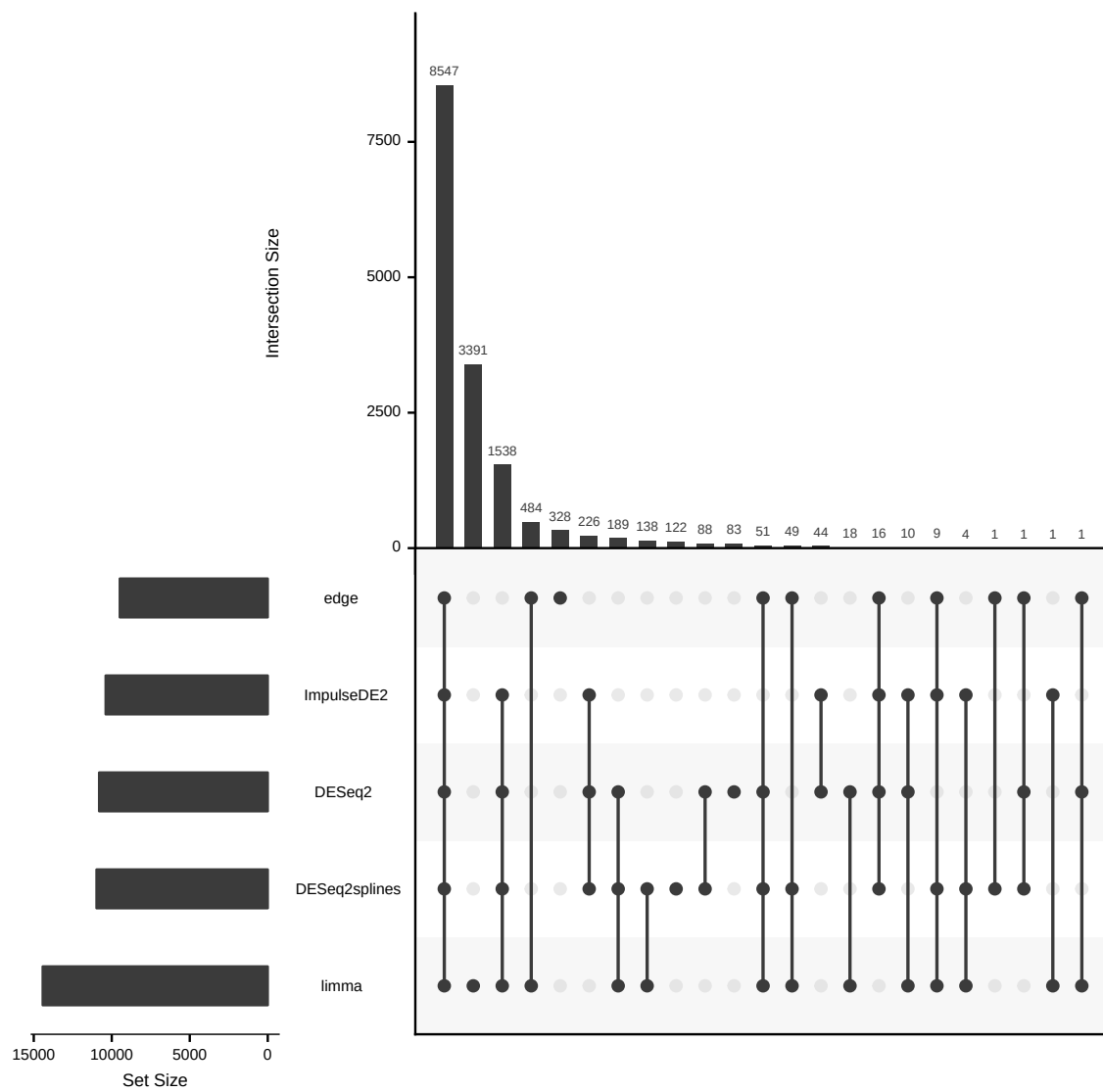
Supplementary Figure 10: UpSetR plots of sets of differentially expressed genes called at a q-value threshold of 0.01 for myeloid (Sykes).



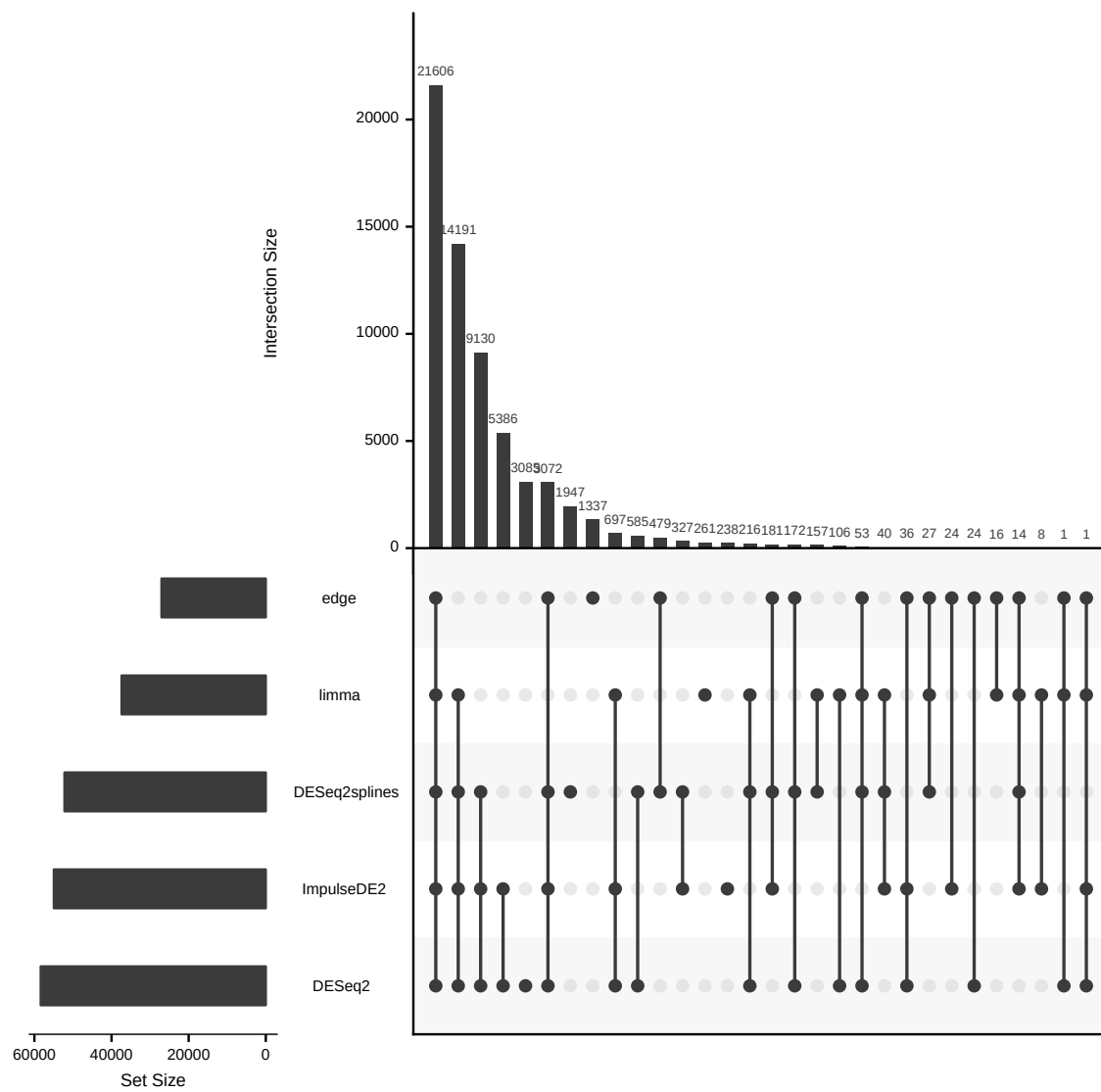
Supplementary Figure 11: UpSetR plots of sets of differentially expressed genes called at a q-value threshold of 0.01 for LPS (Jovanovic).



Supplementary Figure 12: UpSetR plots of sets of differentially expressed genes called at a q-value threshold of 0.01 for *Drosophila* (Graveley).



Supplementary Figure 13: UpSetR plots of sets of differentially expressed genes called at a q-value threshold of 0.01 for hESC (Chu).



Supplementary Figure 14: UpSetR plots of sets of differentially expressed genes called at a q-value threshold of 0.01 for erythroid (Lara-Astiaso).