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

David S. Fischer[a,b,d], Fabian J. Theis[a,c], and Nir Yosef[d,e,1]

[a]Institute of Computational Biology, Helmholtz Centre Munich
[b]TUM School of Life Sciences Weihenstephan, Technical University of Munich
[c]Department of Mathematics, Technical University of Munich
[d]Department of Electrical Engineering & Copmuter Science and Center for Computational Biology, University of California, Berkeley CA
[e]Ragon Institute of MGH, MIT& Harvard, Cambridge MA
[1]To whom correspondence should be addressed. E-mail: niryosef@berkeley.edu

March 3, 2017

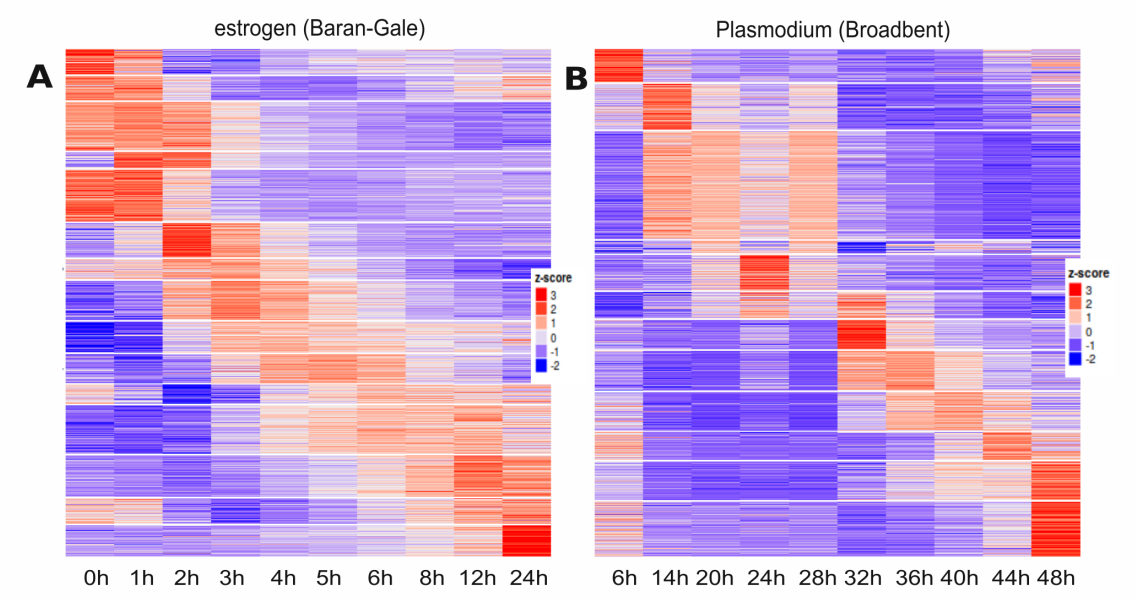

Figure 1: **The impulse model is descriptive of global transcriptome dynamics: Additional data sets. A** RNA-seq of the estrogen response of a human breast cancer cell line ("estrogen (Baran-Gale)"). **B** RNA-seq of long-noncoding RNAs of Plasmodium falciparum during infection("Plasmodium (Broadbent)"). Refer to the Supplementary Methods for further details on the heatmaps.

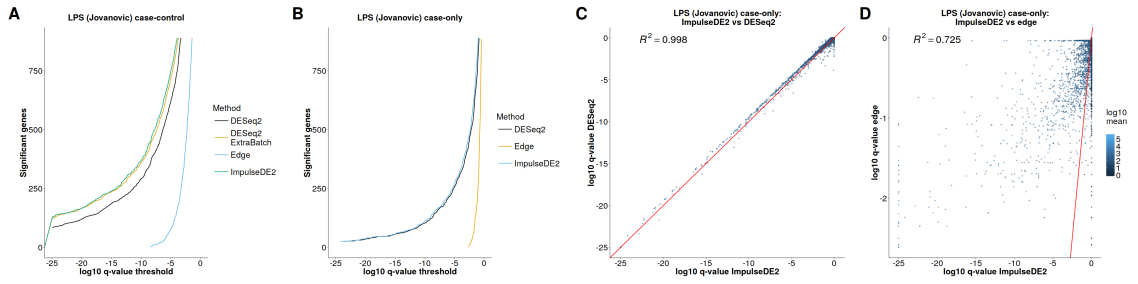

Figure 2: **ImpulseDE2 outperforms edge and performs similar to DESeq2 on LPS (Jovanovic) case-only data. A,B** Number of significantly differentially expressed (**A**: case-only, **B**: case-control) genes as a function of the significance threshold. **C,D** Correlation plots of the inferred differential expression (case-only) Benjamini-Hochberg corrected p-values for all genes between ImpulseDE2 and DESeq2 (**C**) and edge (**D**). R^2 shown are Pearson correlations coefficients.

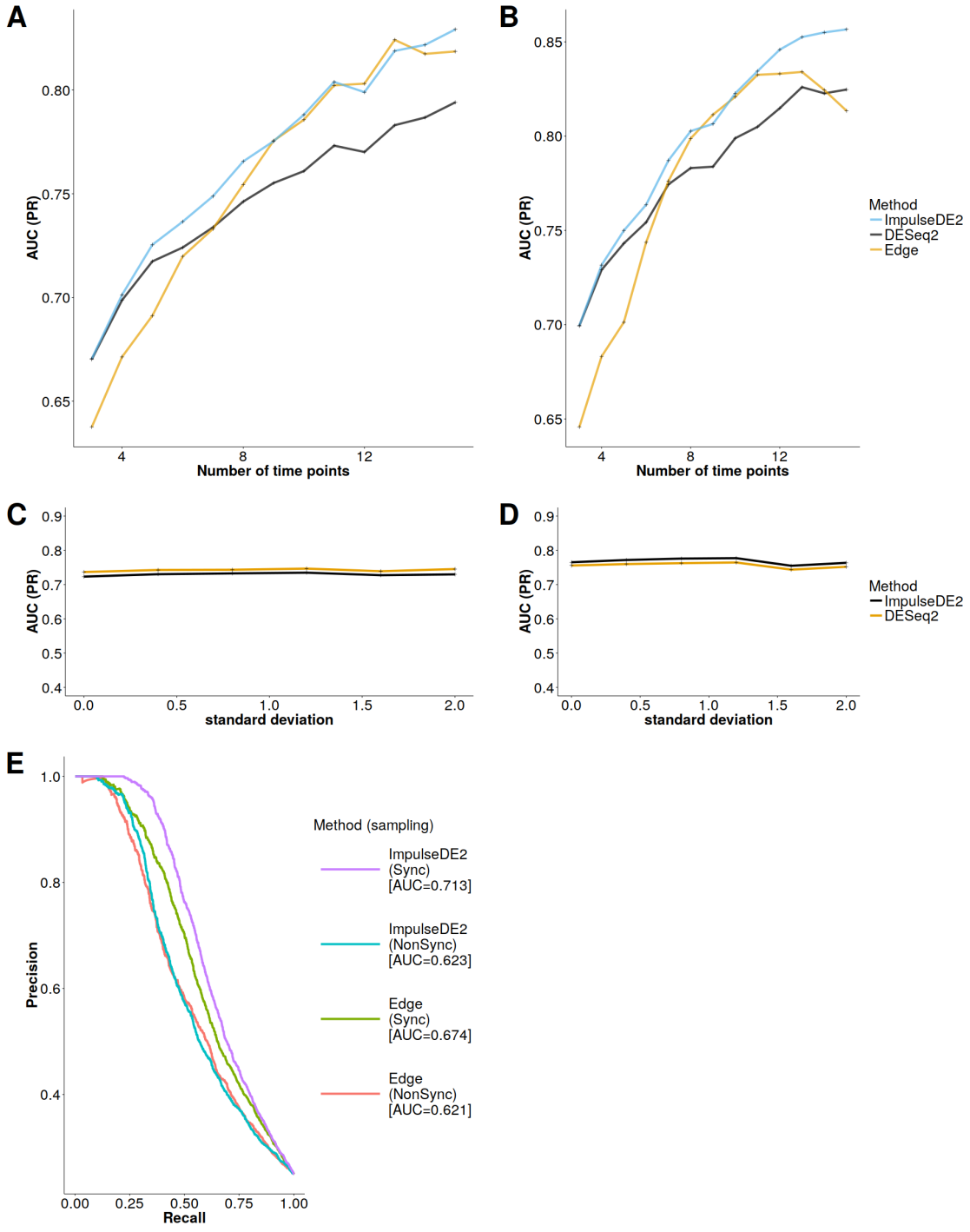

Figure 3: **ImpulseDE2 performance on simulated data.** AUC: area under the curve, PR: precision-recall. **A,B** AUC of PR curve for case-only (**A**) and case-control (**B**), no batch effects simulation with two replicates and varying number of time points. **C,D** AUC of ROC curve for case-only (**C**) and case-control (**D**) analysis with batch effects, the strength of the batch effects is quantified by the standard deviation of the normal distribution form which the batch factor is drawn. **E** PR curves of ImpulseDE2 and edge for the case-control scenario with shared time points between samples of both conditions (Sync) and no shared time points (NonSync).

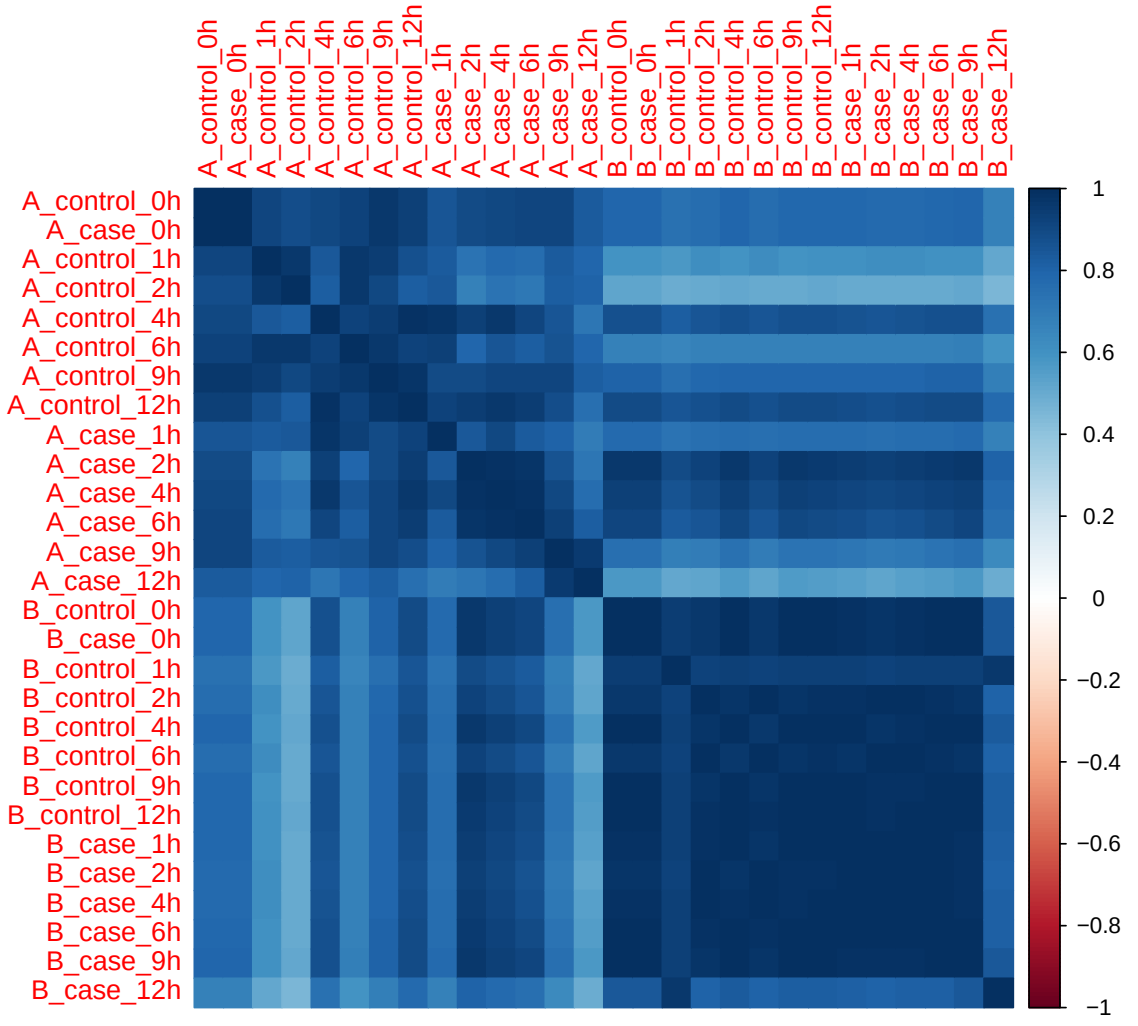

Figure 4: **Correlation structure and batch structure of LPS (Jovanovic) data set.** Batch: A, B. These groups of samples were handled together, note that they overlap case and control condition. Condition: "case" (case condition with LPS) and "ctrl" (control without LPS treatment). Time point 0h was only sampled in the control condition and was replicated for the case condition in each batch for analysis. Shown are Pearson correlation coefficients.

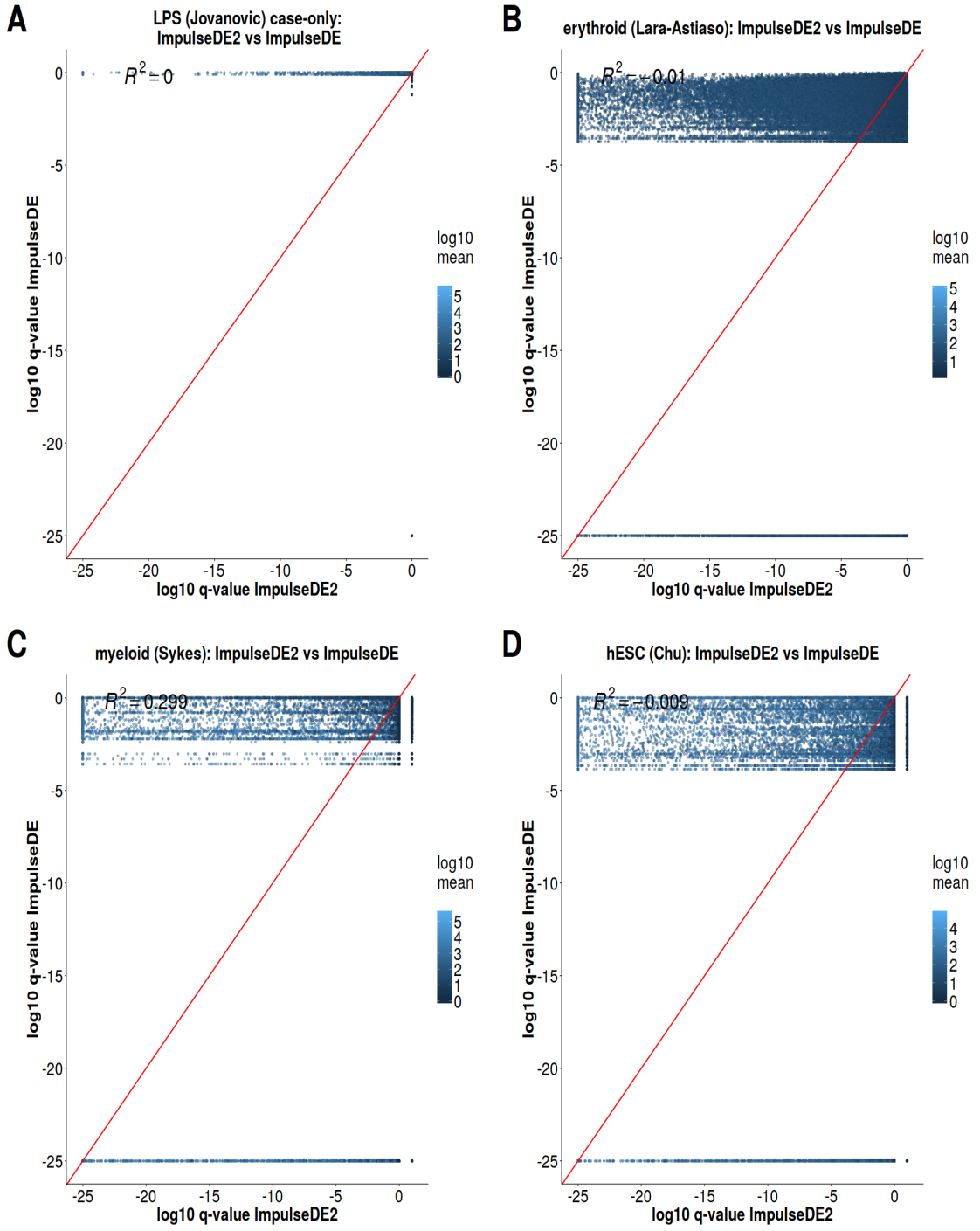

Figure 5: **Q-value correlation plots: ImpulseDE2 vs ImpulseDE.** Correlation plot of the inferred differential expression q-values for all genes between ImpulseDE2 and ImpulseDE: **A** LPS (Jovanovic) case-only, **B** myeloid (Sykes), **C** hESC (Chu), **D** erythroid ChIP (Lara-Astiaso).

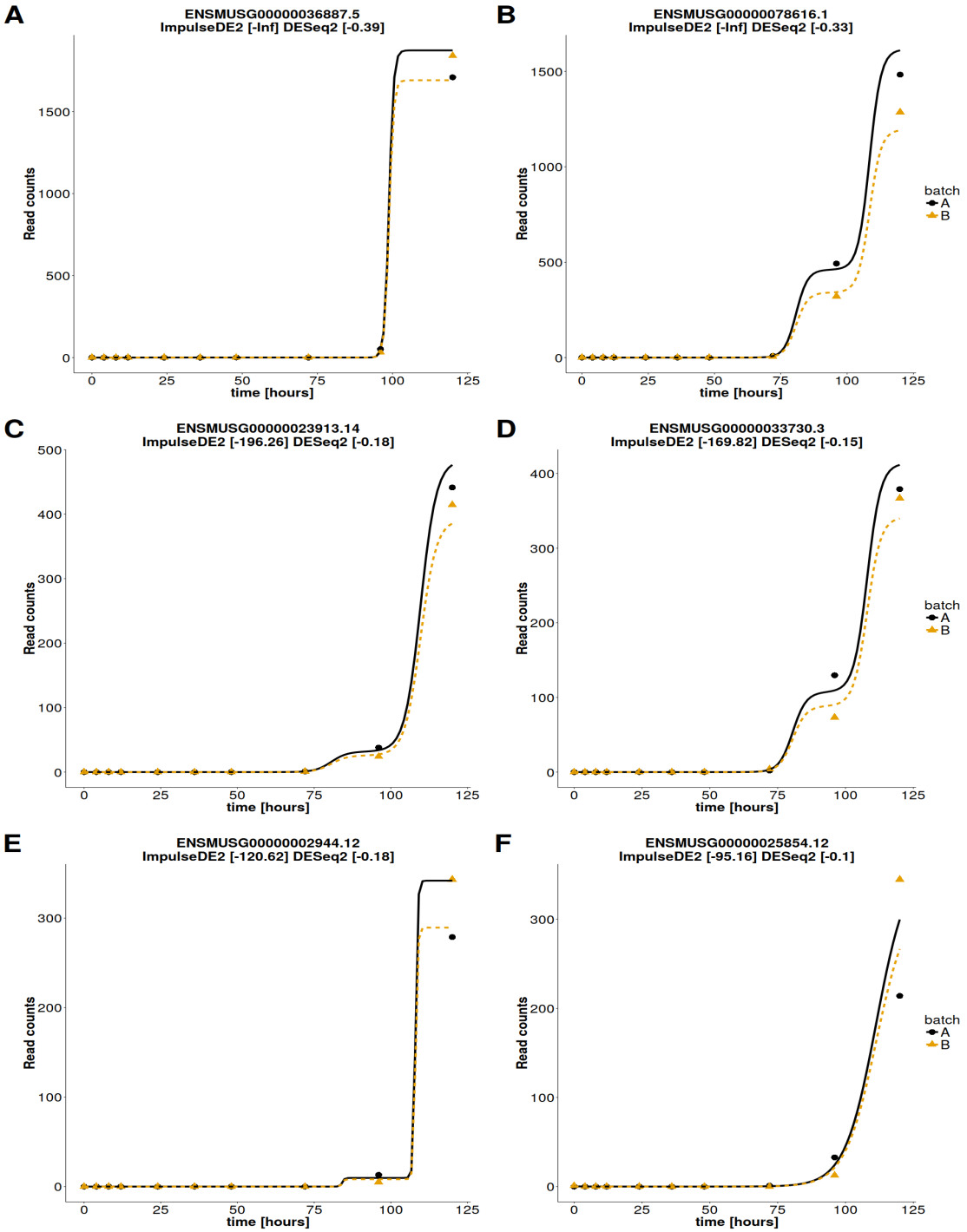

Figure 6: **Sykes et al. RNA-seq (case-only): ImpulseDE2 versus DEseq2: Examples gene with lower p-value assigned by ImpulseDE2 than DESeq2.** Title: Method[FDR-corrected log10 p-value]. DESeq2 is run with one batch factor per batch (standard setting). Shown are read counts normalized by size factors. The shown genes are the genes with the lowest ratio (equivalent to the highest difference) of Benjamini-Hochberg false discovery rate-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).

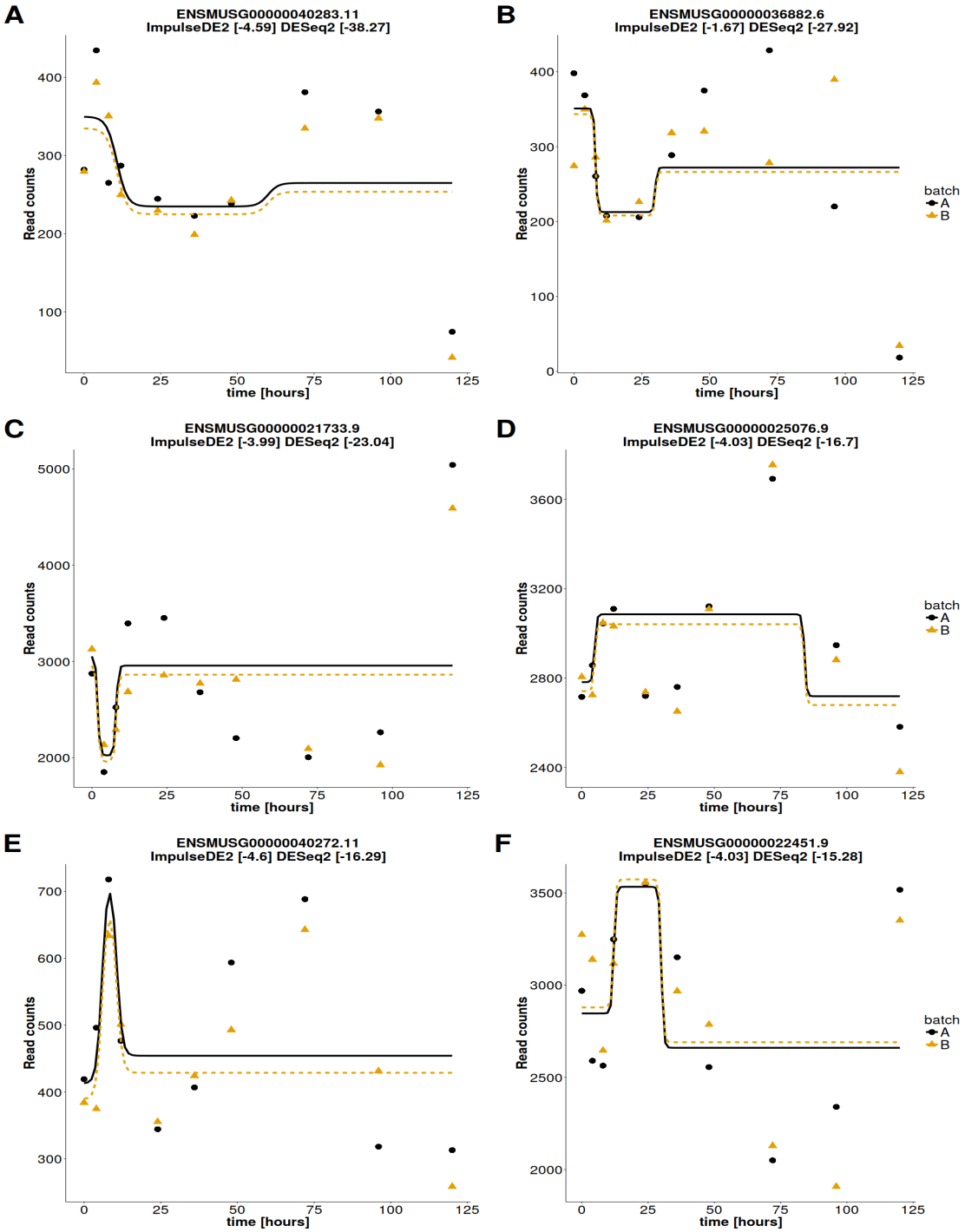

Figure 7: **Sykes et al. RNA-seq (case-only): ImpulseDE2 versus DEseq2: Examples gene with lower p-value assigned by DEseq2 than ImpulseDE2.** Title: Method[FDR-corrected log10 p-value]. DESeq2 is run with one batch factor per batch (standard setting). Shown are read counts normalized by size factors and impulse model fits. The shown genes are the genes with the lowest ratio (equivalent to the highest difference) of Benjamini-Hochberg false discovery rate-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).

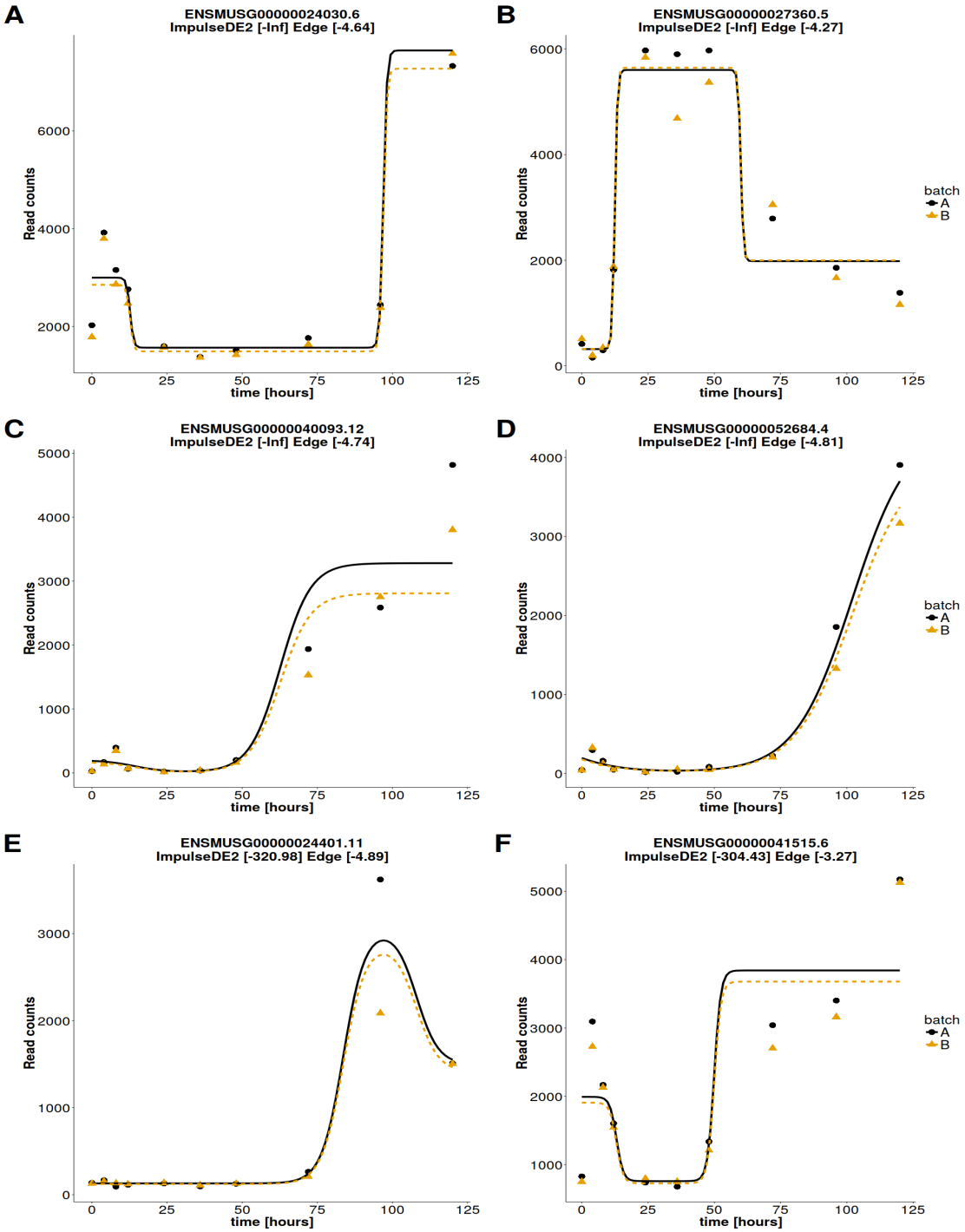

Figure 8: **Sykes et al. RNA-seq (case-only): ImpulseDE2 versus edge: Examples gene with lower p-value assigned by ImpulseDE2 than edge.** Title: Method[FDR-corrected log10 p-value]. Edge is run with one batch factor per batch (standard setting). Shown are read counts normalized by size factors and impulse model fits. The shown genes are the genes with the lowest ratio (equivalent to the highest difference) of Benjamini-Hochberg false discovery rate-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).

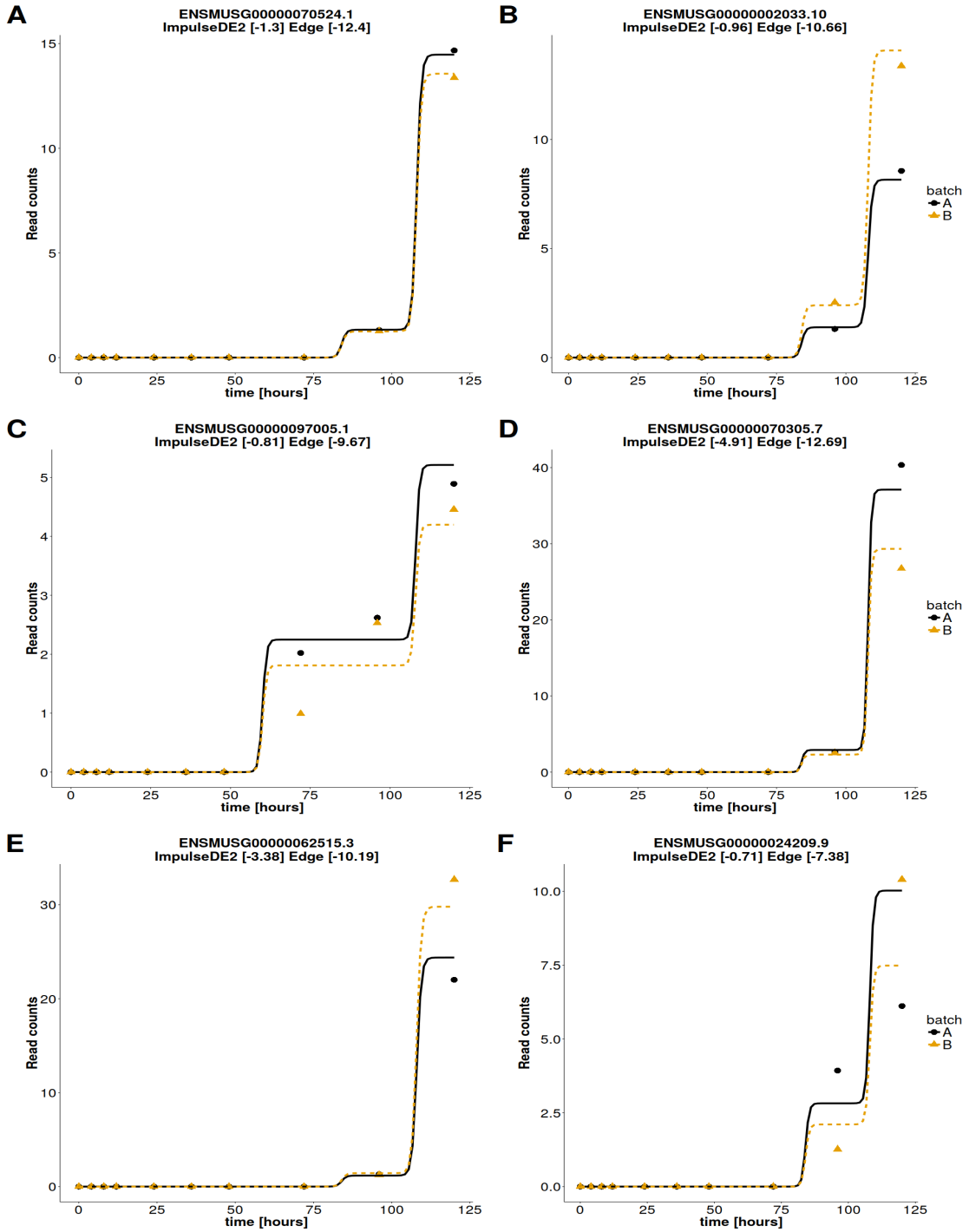

Figure 9: **Sykes et al. RNA-seq (case-only): ImpulseDE2 versus edge: Examples gene with lower p-value assigned by edge than ImpulseDE2.** Title: Method[FDR-corrected log10 p-value]. Edge is run with one batch factor per batch (standard setting). Shown are read counts normalized by size factors and impulse model fits. The shown genes are the genes with the lowest ratio (equivalent to the highest difference) of Benjamini-Hochberg false discovery rate-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).

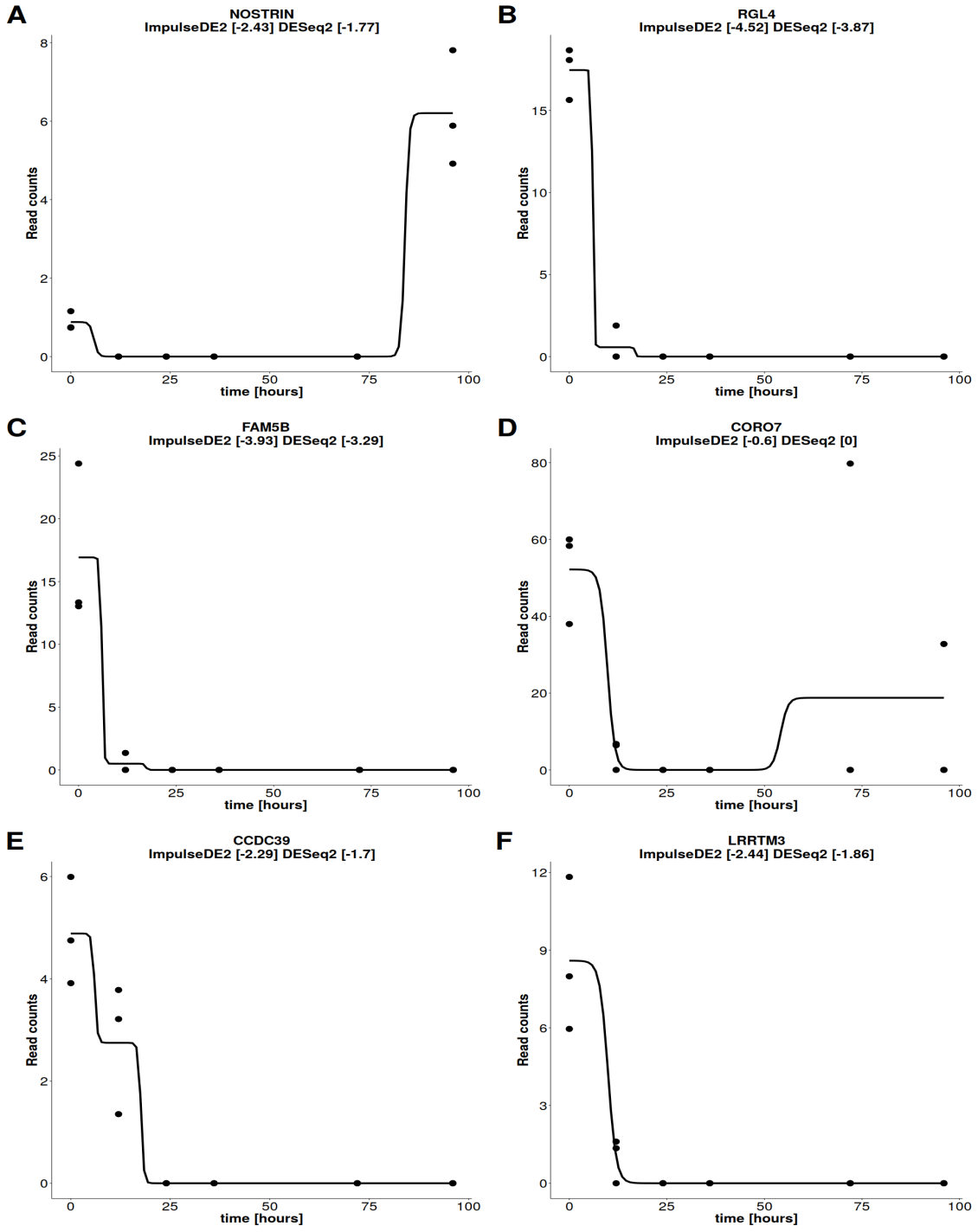

Figure 10: **Chu et al. RNA-seq (case-only): ImpulseDE2 versus DEseq2: Examples gene with lower p-value assigned by ImpulseDE2 than DESeq2.** Title: Method[FDR-corrected log10 p-value]. DESeq2 is run with one batch factor per batch (standard setting). Shown are read counts normalized by size factors and impulse model fits. The shown genes are the genes with the lowest ratio (equivalent to the highest difference) of Benjamini-Hochberg false discovery rate-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).

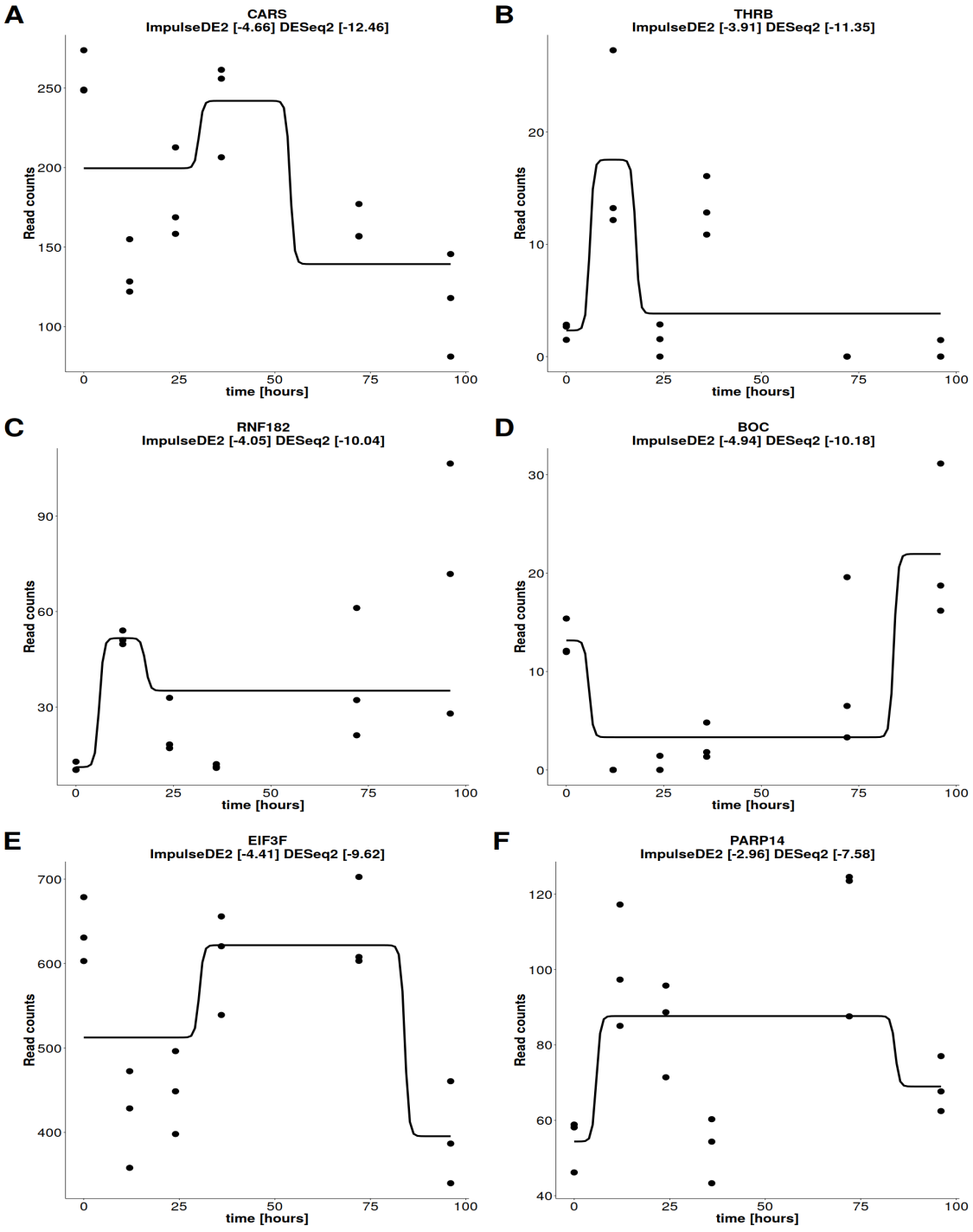

Figure 11: **Chu et al. RNA-seq (case-only): ImpulseDE2 versus DEseq2: Examples gene with lower p-value assigned by DEseq2 than ImpulseDE2.** Title: Method[FDR-corrected log10 p-value]. DESeq2 is run with one batch factor per batch (standard setting). Shown are read counts normalized by size factors and impulse model fits. The shown genes are the genes with the lowest ratio (equivalent to the highest difference) of Benjamini-Hochberg false discovery rate-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).

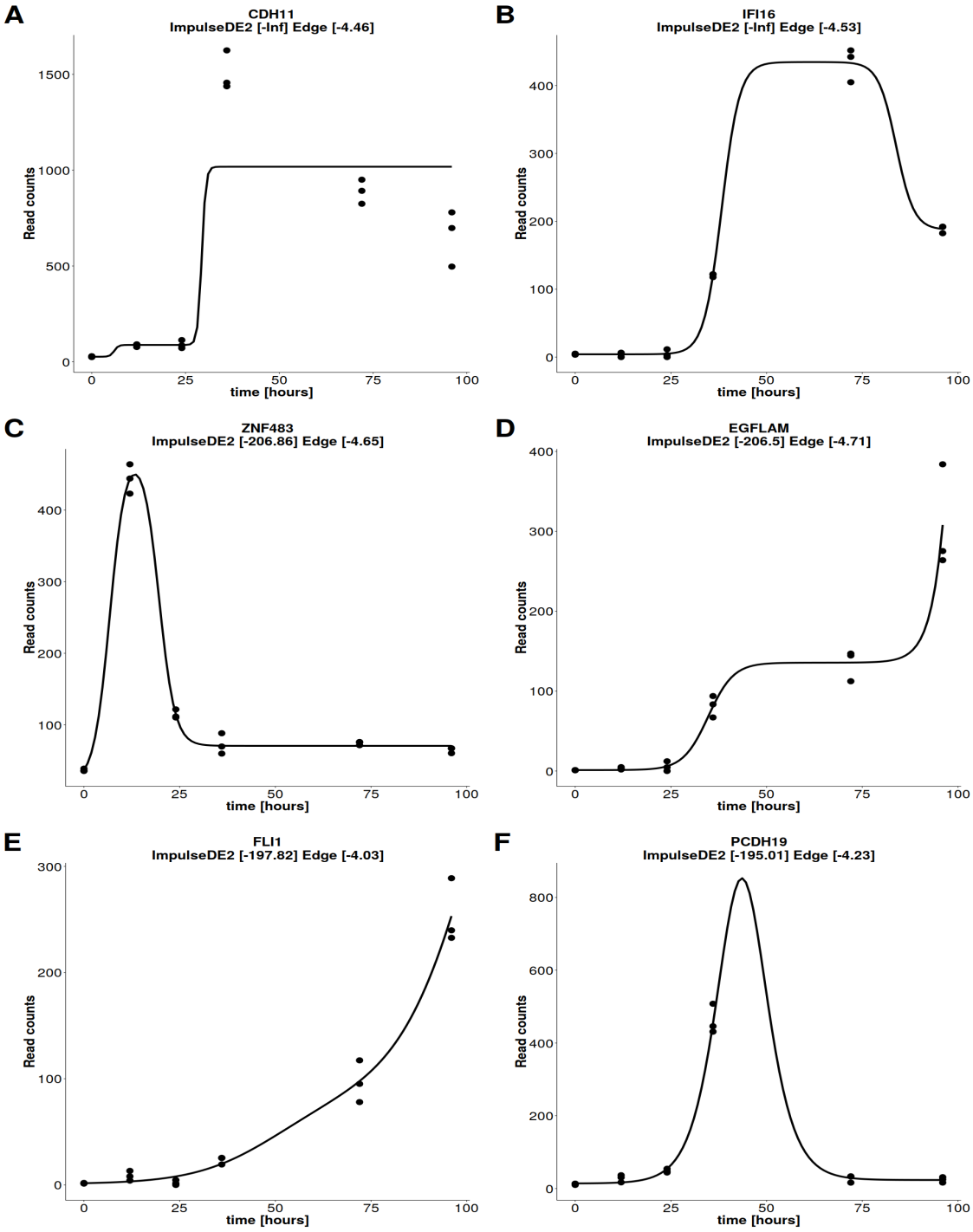

Figure 12: **Chu et al. RNA-seq (case-only): ImpulseDE2 versus edge: Examples gene with lower p-value assigned by ImpulseDE2 than edge.** Title: Method[FDR-corrected log10 p-value]. Edge is run with one batch factor per batch (standard setting). Shown are read counts normalized by size factors and impulse model fits. The shown genes are the genes with the lowest ratio (equivalent to the highest difference) of Benjamini-Hochberg false discovery rate-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).

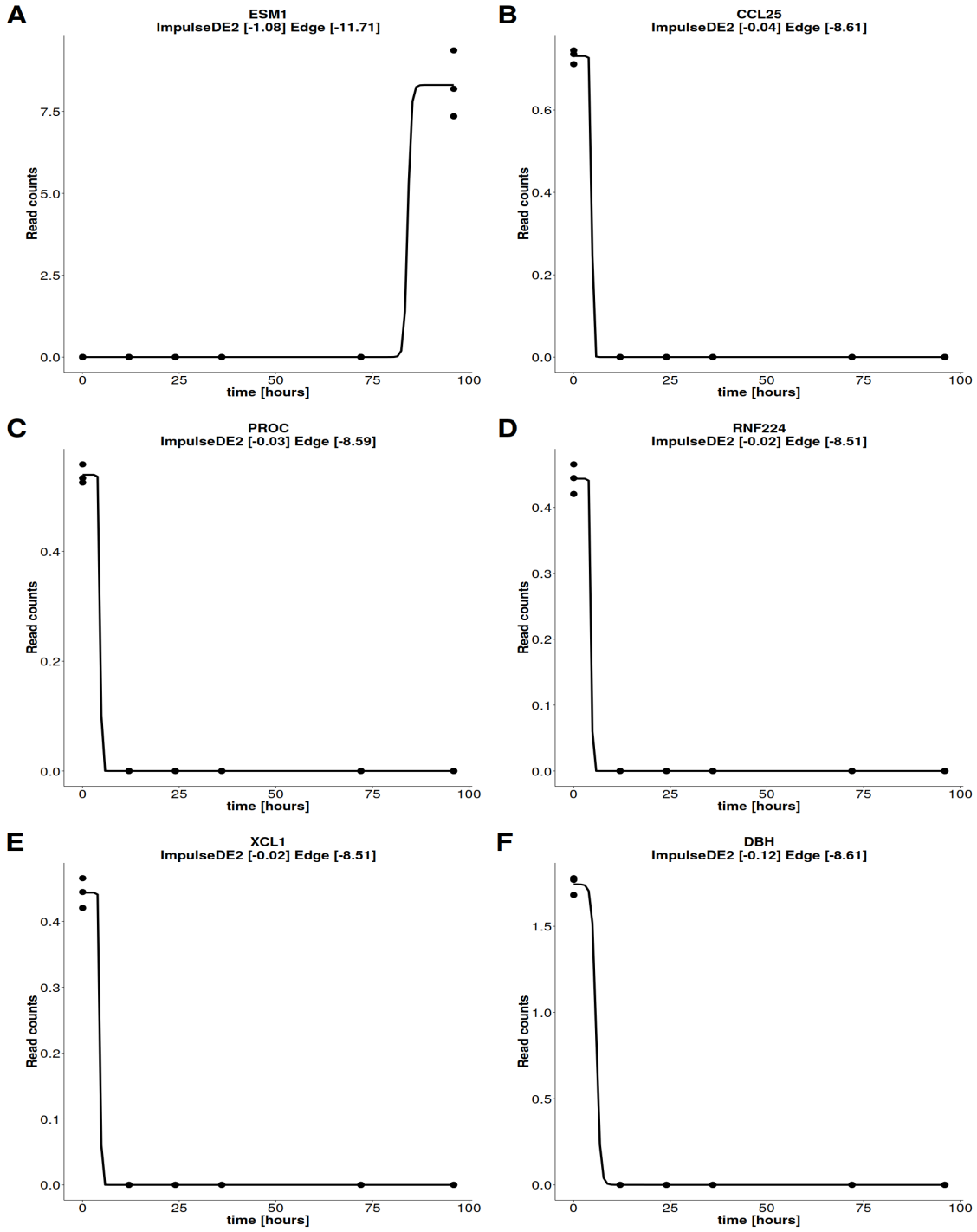

Figure 13: **Chu et al. RNA-seq (case-only): ImpulseDE2 versus edge: Examples gene with lower p-value assigned by edge than ImpulseDE2.** Title: Method[FDR-corrected log10 p-value]. Edge is run with one batch factor per batch (standard setting). Shown are read counts normalized by size factors and impulse model fits. The shown genes are the genes with the lowest ratio (equivalent to the highest difference) of Benjamini-Hochberg false discovery rate-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).

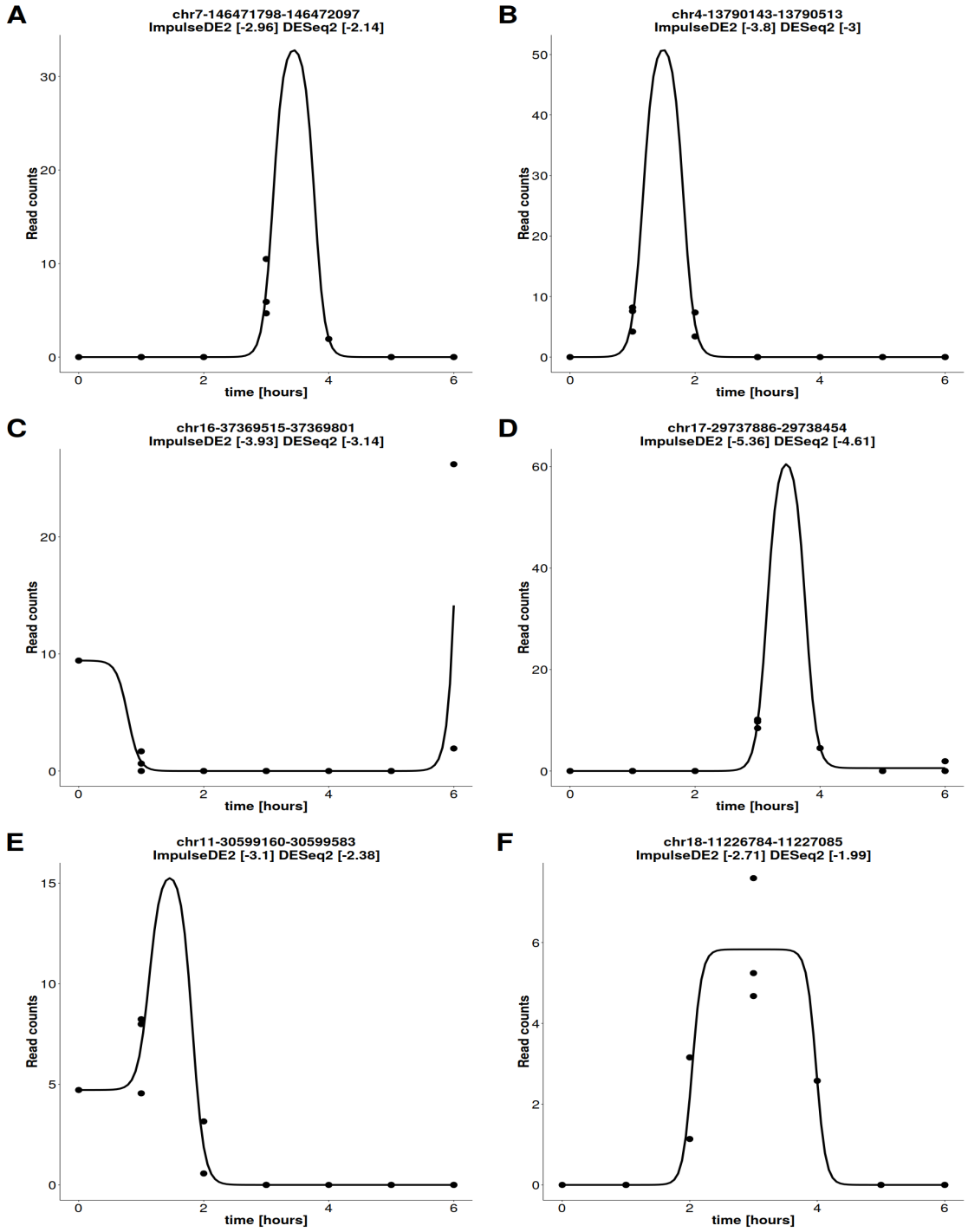

Figure 14: **Erythroid chromatin (Lara-Astiaso) (case-only): ImpulseDE2 versus DEseq2: Examples gene with lower p-value assigned by ImpulseDE2 than DESeq2.** Title: Method[FDR-corrected log10 p-value]. Shown are read counts normalized by size factors and impulse model fits. The shown genes are the genes with the lowest ratio (equivalent to the highest difference) of Benjamini-Hochberg false discovery rate-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).

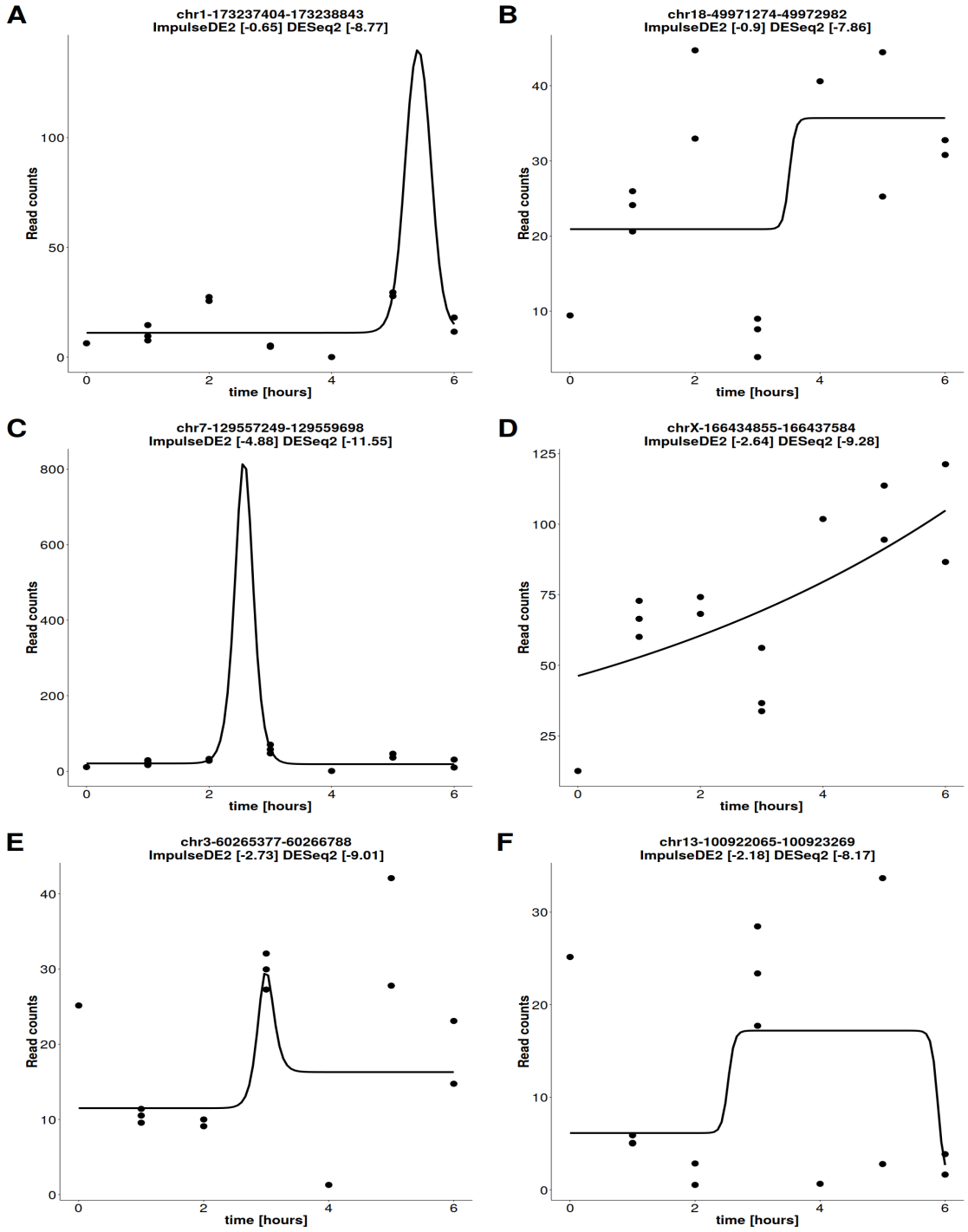

Figure 15: **Erythroid chromatin (Lara-Astiaso) (case-only): ImpulseDE2 versus DEseq2: Examples gene with lower p-value assigned by DEseq2 than ImpulseDE2.** Title: Method[FDR-corrected log10 p-value]. Shown are read counts normalized by size factors and impulse model fits. The shown genes are the genes with the lowest ratio (equivalent to the highest difference) of Benjamini-Hochberg false discovery rate-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).

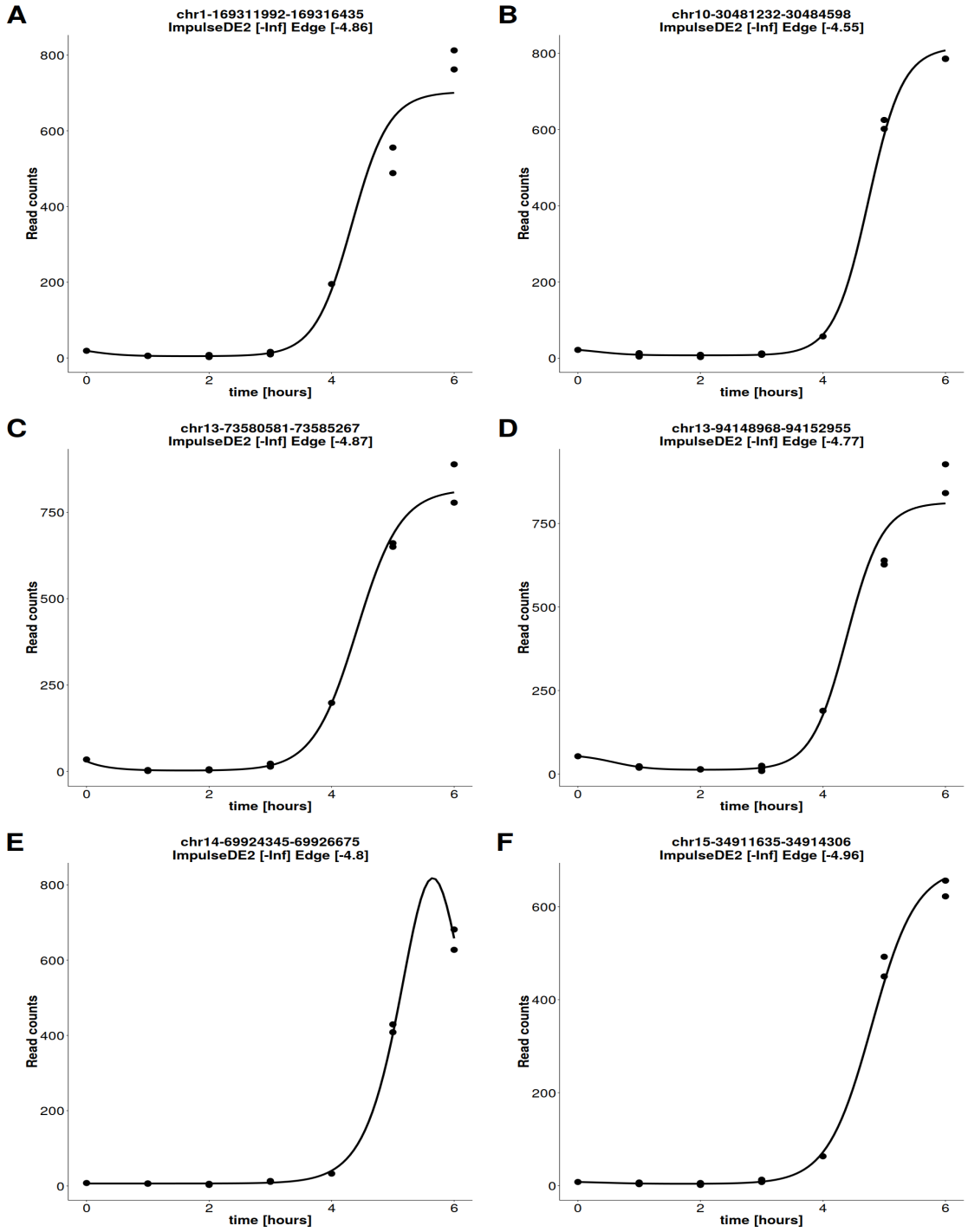

Figure 16: **Erythroid chromatin (Lara-Astiaso) (case-only): ImpulseDE2 versus edge: Examples gene with lower p-value assigned by ImpulseDE2 than edge.** Title: Method[FDR-corrected log10 p-value]. Shown are read counts normalized by size factors and impulse model fits. The shown genes are the genes with the lowest ratio (equivalent to the highest difference) of Benjamini-Hochberg false discovery rate-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).

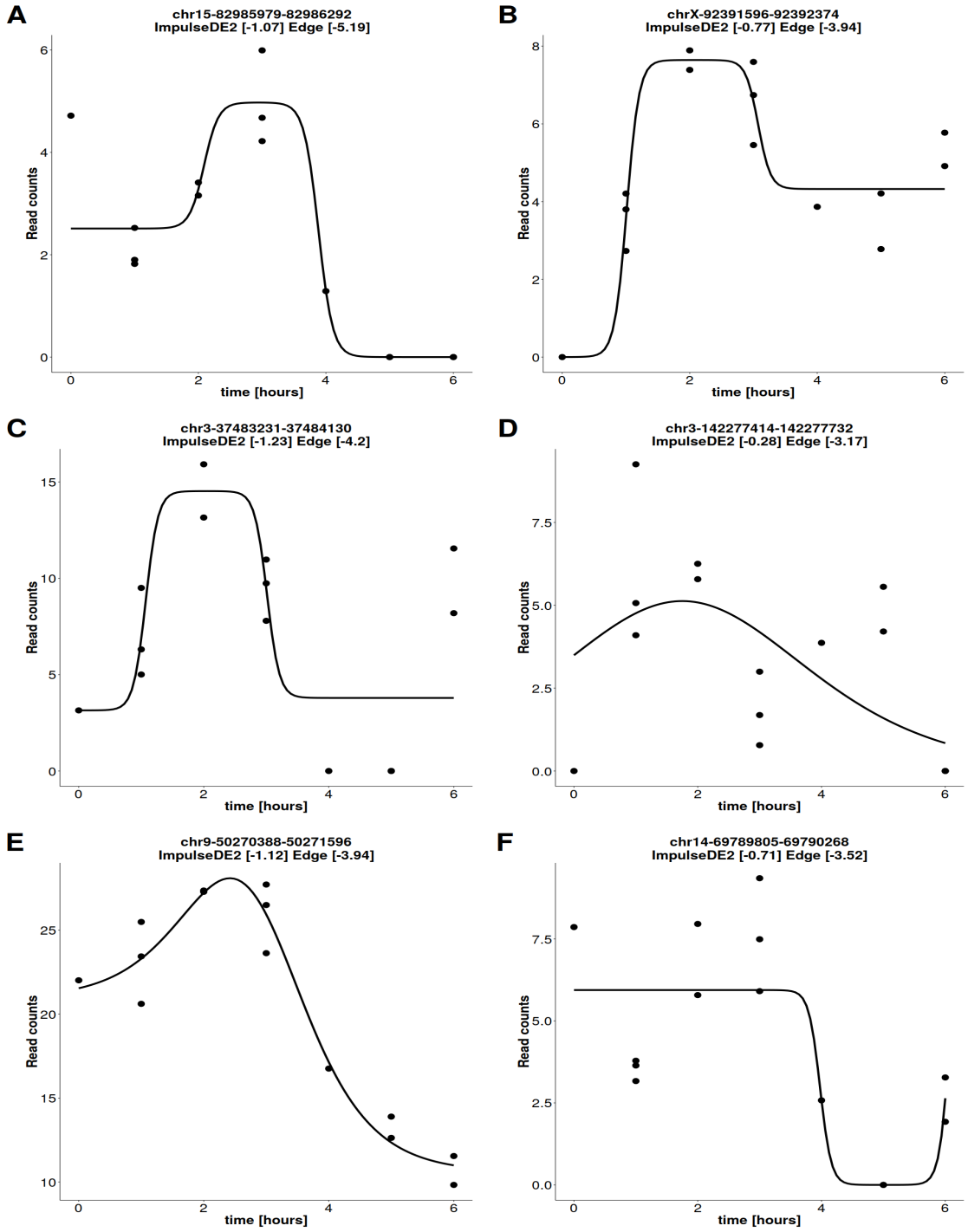

Figure 17: **Erythroid chromatin (Lara-Astiaso) (case-only): ImpulseDE2 versus edge: Examples gene with lower p-value assigned by edge than ImpulseDE2.** Title: Method[FDR-corrected log10 p-value]. Shown are read counts normalized by size factors and impulse model fits. The shown genes are the genes with the lowest ratio (equivalent to the highest difference) of Benjamini-Hochberg false discovery rate-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).

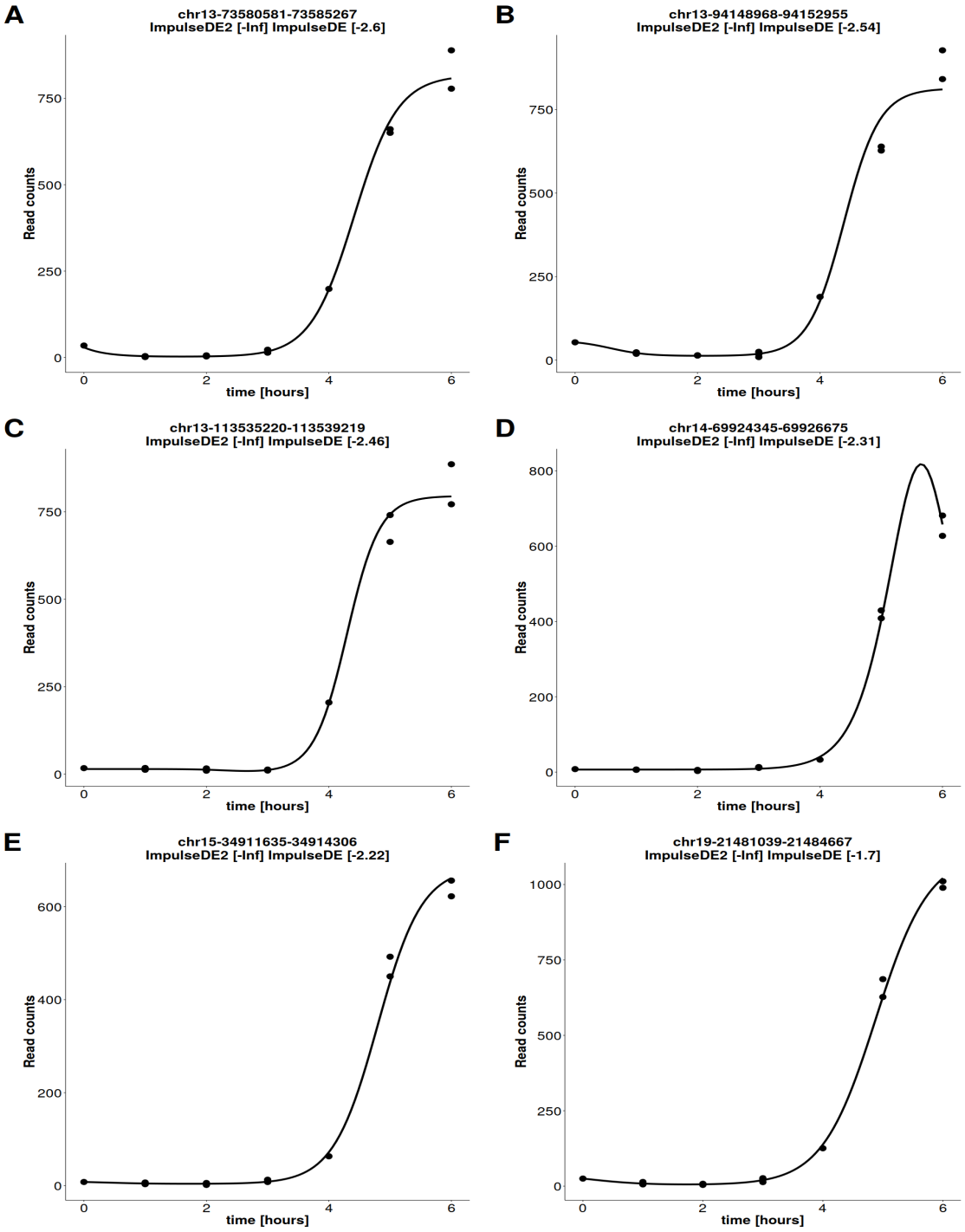

Figure 18: **Erythroid chromatin (Lara-Astiaso) (case-only): ImpulseDE2 versus ImpulseDE: Examples gene with lower p-value assigned by ImpulseDE2 than ImpulseDE.** Title: Method[FDR-corrected log10 p-value]. Shown are read counts normalized by size factors and impulse model fits. The shown genes are the genes with the lowest ratio (equivalent to the highest difference) of Benjamini-Hochberg false discovery rate-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).

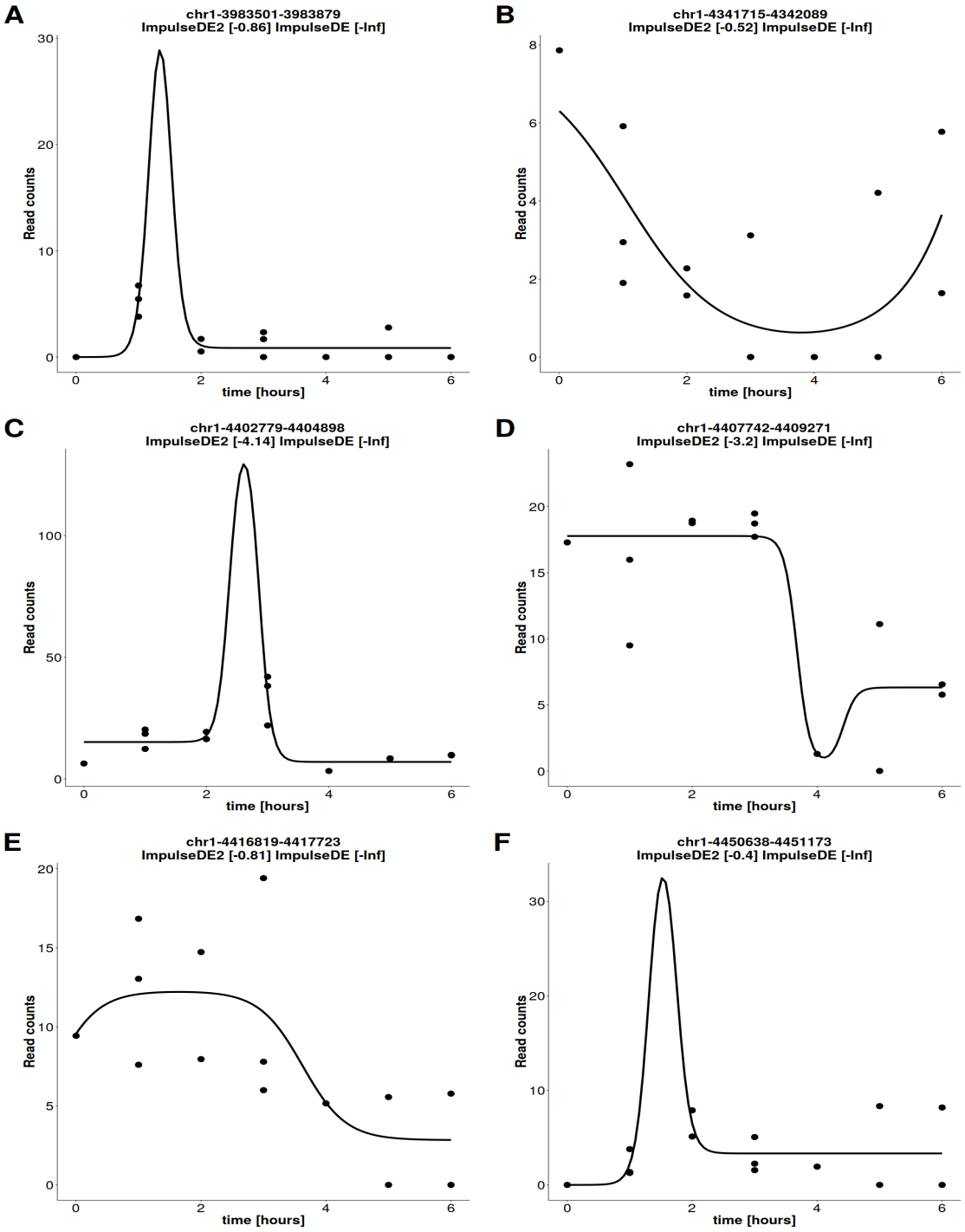

Figure 19: **Erythroid chromatin (Lara-Astiaso) (case-only): ImpulseDE2 versus ImpulseDE: Examples gene with lower p-value assigned by ImpulseDE than ImpulseDE2.** Title: Method[FDR-corrected log10 p-value]. Shown are read counts normalized by size factors and impulse model fits. The shown genes are the genes with the lowest ratio (equivalent to the highest difference) of Benjamini-Hochberg false discovery rate-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).

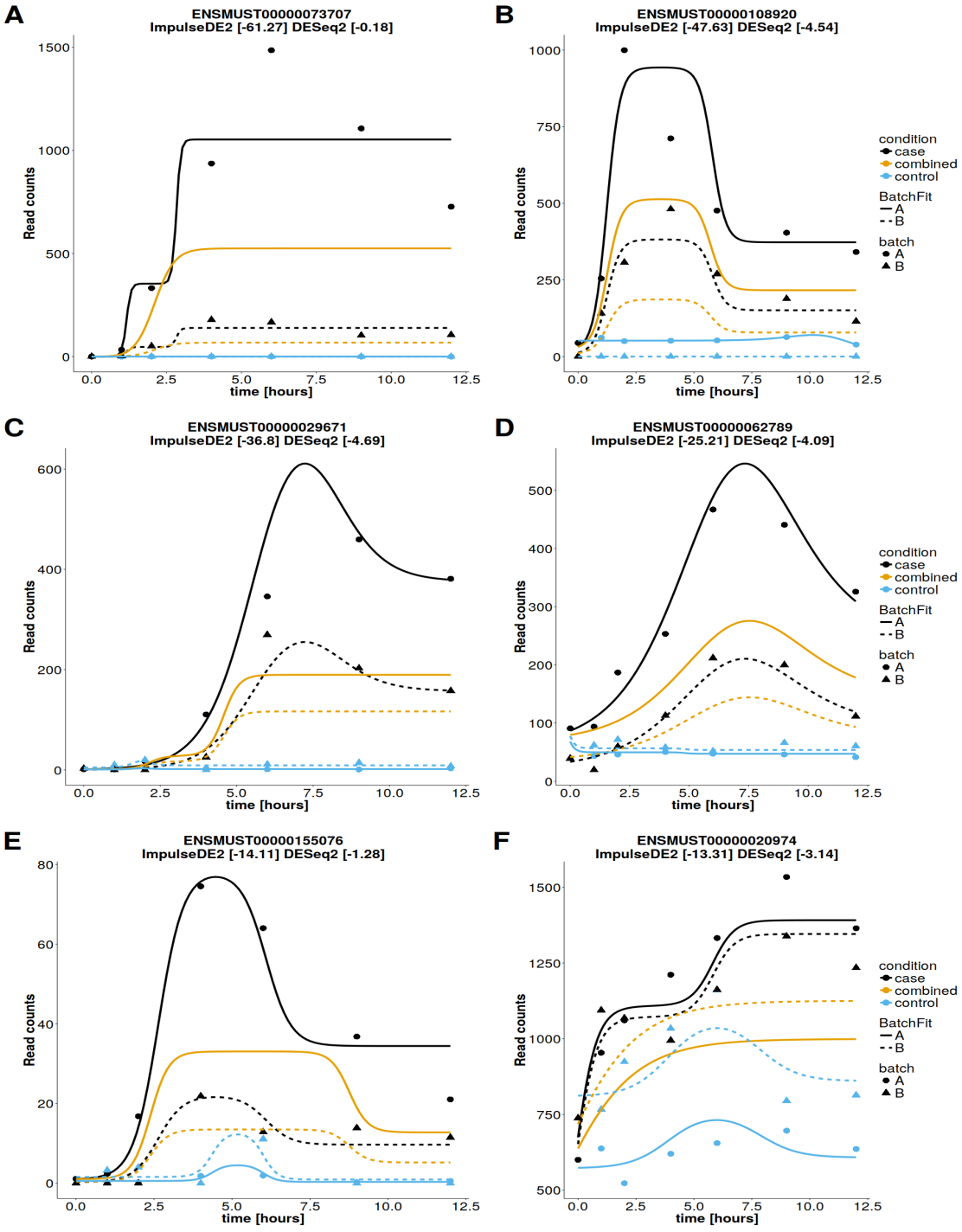

Figure 20: **Jovanovic et al. RNA-seq (case-control): ImpulseDE2 versus DEseq2: Examples gene with lower p-value assigned by ImpulseDE2 than DESeq2.** Title: Method[FDR-corrected log10 p-value]. DESeq2 is run with one batch factor per batch (standard setting). Shown are read counts normalized by size factors and impulse model fits. The shown genes are the genes with the lowest ratio (equivalent to the highest difference) of Benjamini-Hochberg false discovery rate-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).

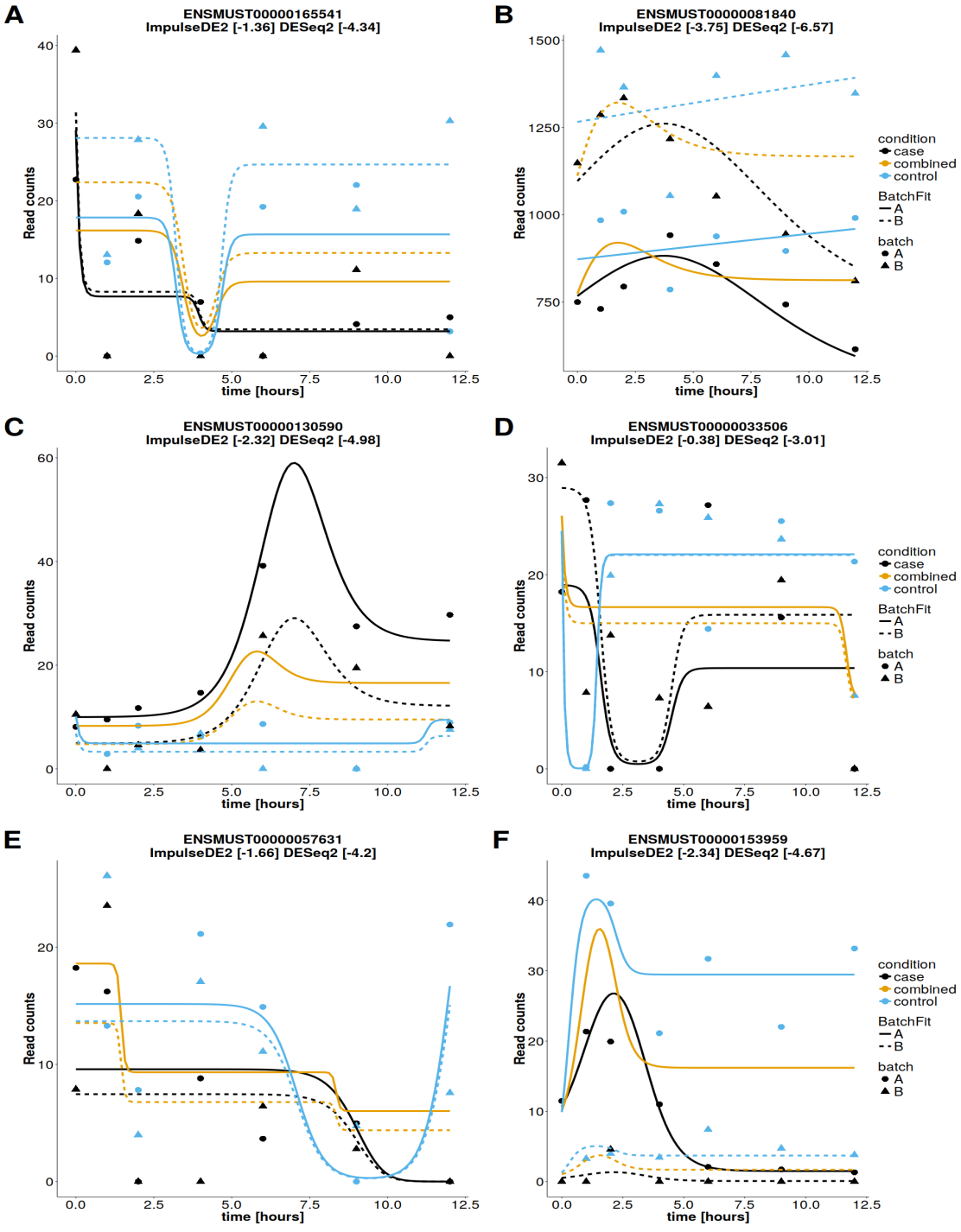

Figure 21: **Jovanovic et al. RNA-seq (case-control): ImpulseDE2 versus DEseq2: Examples gene with lower p-value assigned by DEseq2 than ImpulseDE2.** Title: Method[FDR-corrected log10 p-value]. DESeq2 is run with one batch factor per batch (standard setting). Shown are read counts normalized by size factors and impulse model fits. The shown genes are the genes with the lowest ratio (equivalent to the highest difference) of Benjamini-Hochberg false discovery rate-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).

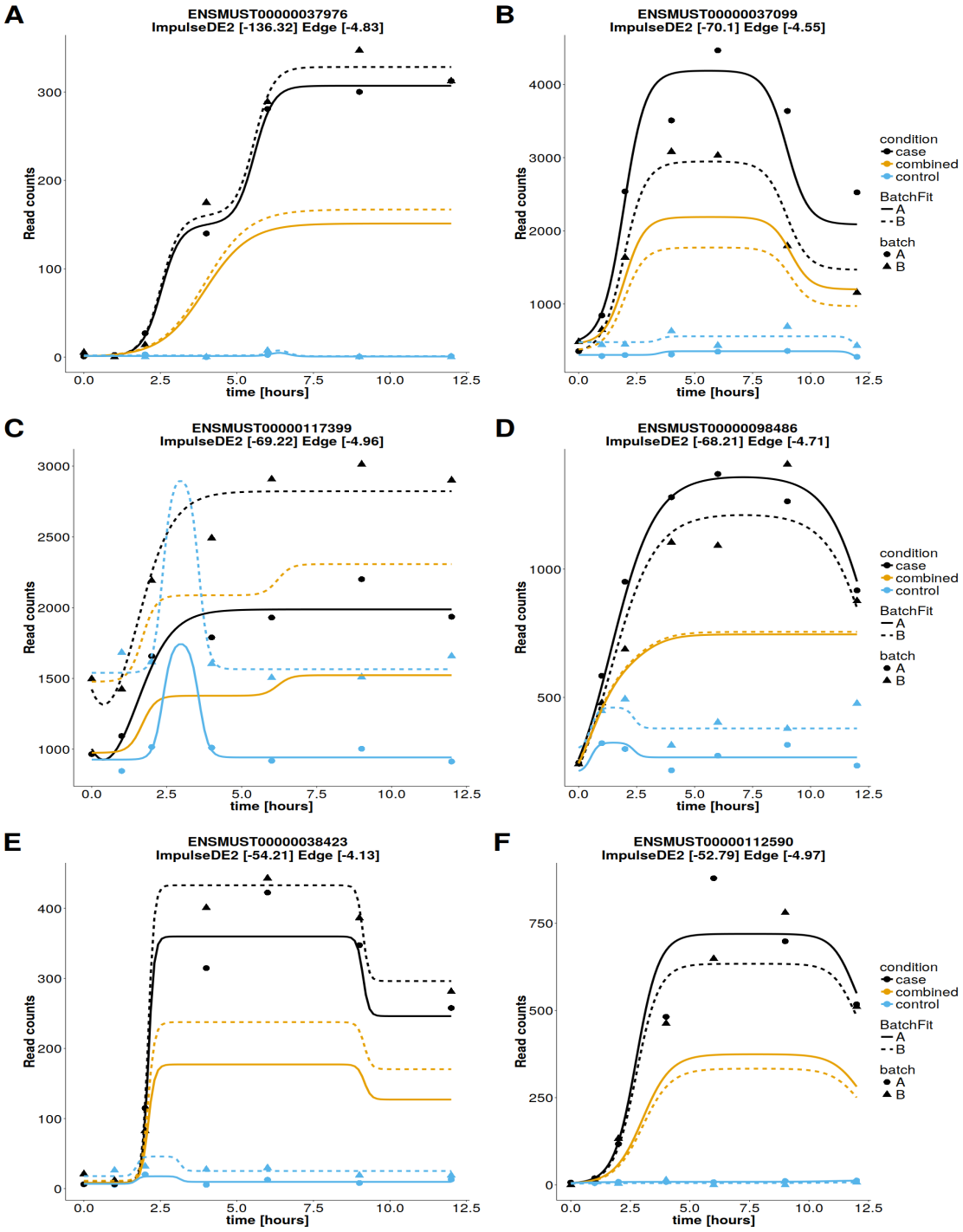

Figure 22: **Jovanovic et al. RNA-seq (case-control): ImpulseDE2 versus edge: Examples gene with lower p-value assigned by ImpulseDE2 than edge.** Title: Method[FDR-corrected log10 p-value]. Edge is run with one batch factor per batch (standard setting). Shown are read counts normalized by size factors and impulse model fits. The shown genes are the genes with the lowest ratio (equivalent to the highest difference) of Benjamini-Hochberg false discovery rate-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).

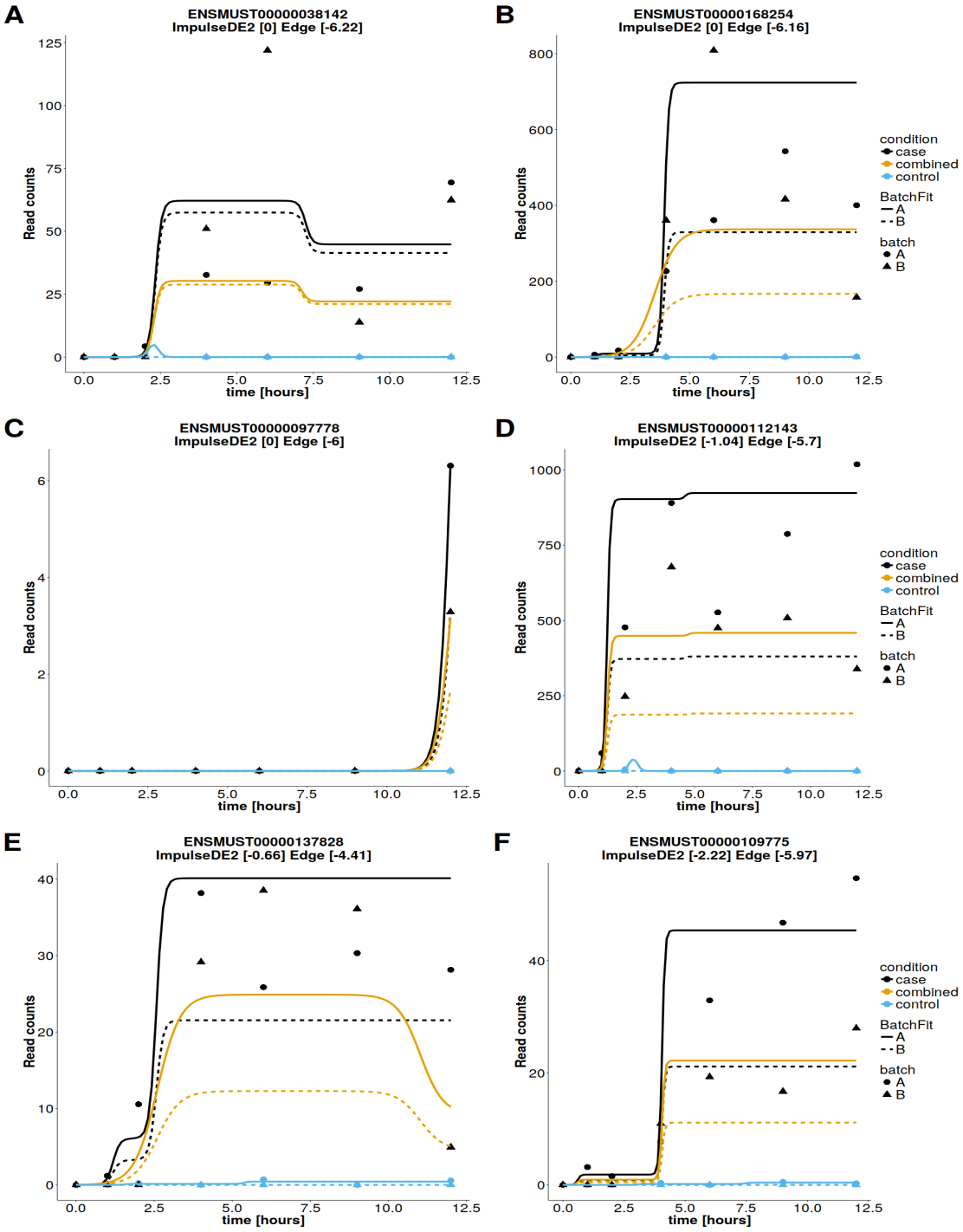

Figure 23: **Jovanovic et al. RNA-seq (case-control): ImpulseDE2 versus edge: Examples gene with lower p-value assigned by edge than ImpulseDE2.** Title: Method[FDR-corrected log10 p-value]. Edge is run with one batch factor per batch (standard setting). Shown are read counts normalized by size factors and impulse model fits. The shown genes are the genes with the lowest ratio (equivalent to the highest difference) of Benjamini-Hochberg false discovery rate-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).