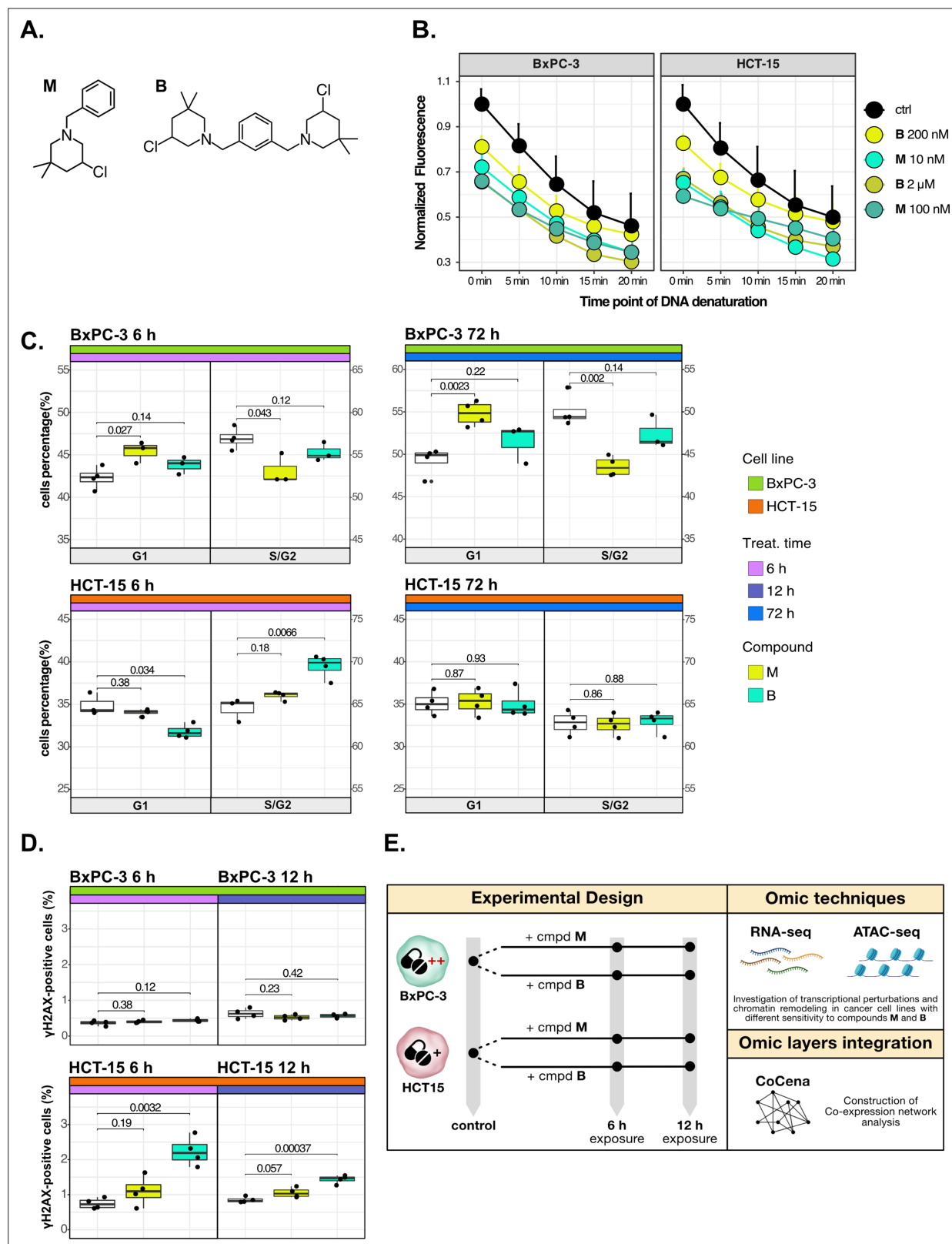


---

## Figures and figure supplements

Decoding mechanism of action and sensitivity to drug candidates from integrated transcriptome and chromatin state

**Caterina Carraro *et al***

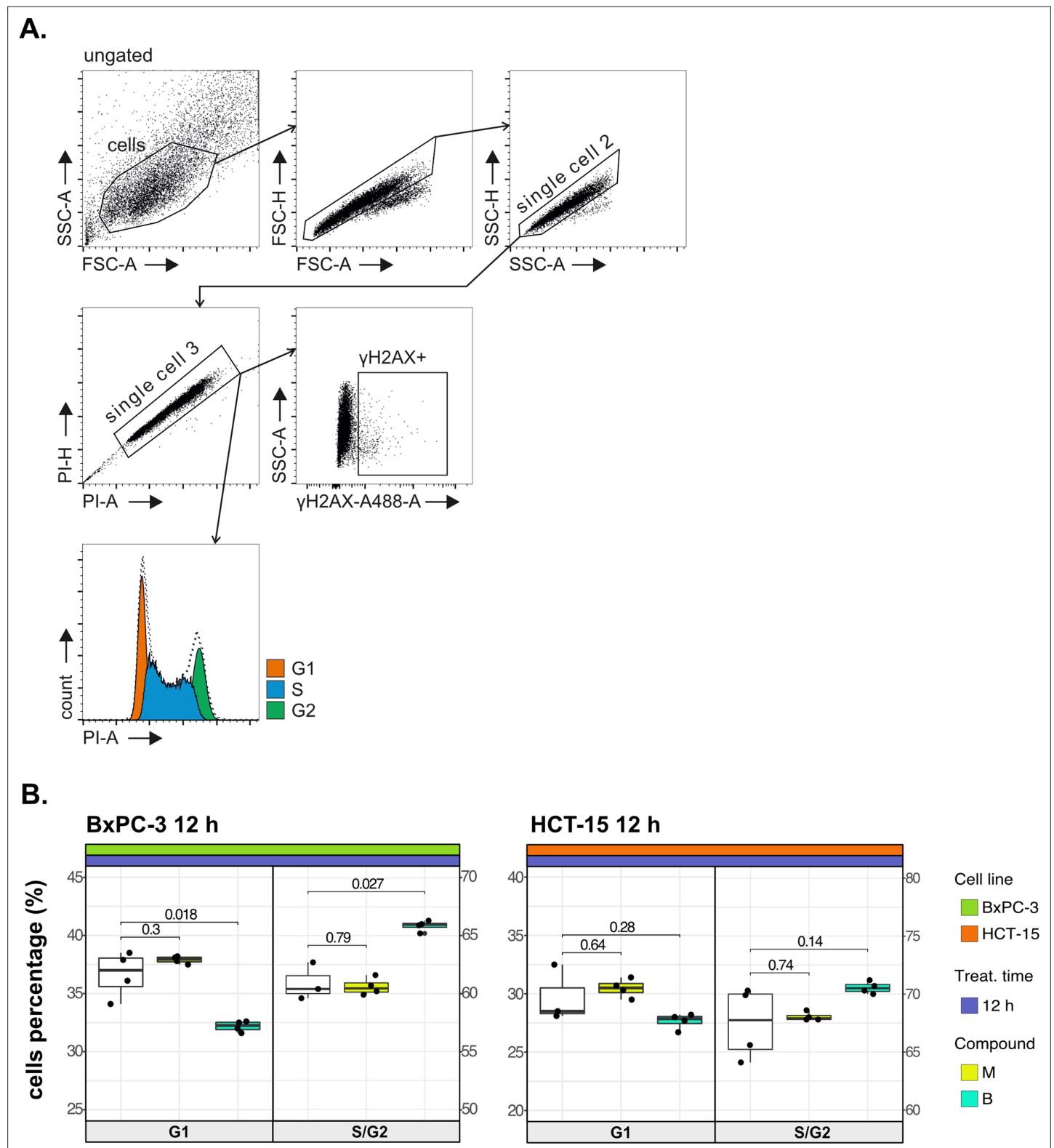


**Figure 1.** Cancer tropism of 3-chloropiperidines (3-CePs) is not explained by DNA damage. **(A)** Chemical structure of the analyzed 3-CePs (M=monofunctional and B=bifunctional). **(B)** Quantification of genomic DNA damage in BxPC-3 and HCT-15 cells treated with M (10 nM and 100 nM), B (200 nM and 2  $\mu$ M), or DMSO (dimethyl sulfoxide) 0.5% (ctrl) for 6 hr and analyzed by the fast micromethod single-strand-break assay: alkaline denaturation of DNA is followed in time up to 20 min by monitoring the fluorescence of the dsDNA-specific PicoGreen dye. **(C)** Cell cycle distribution

Figure 1 continued on next page

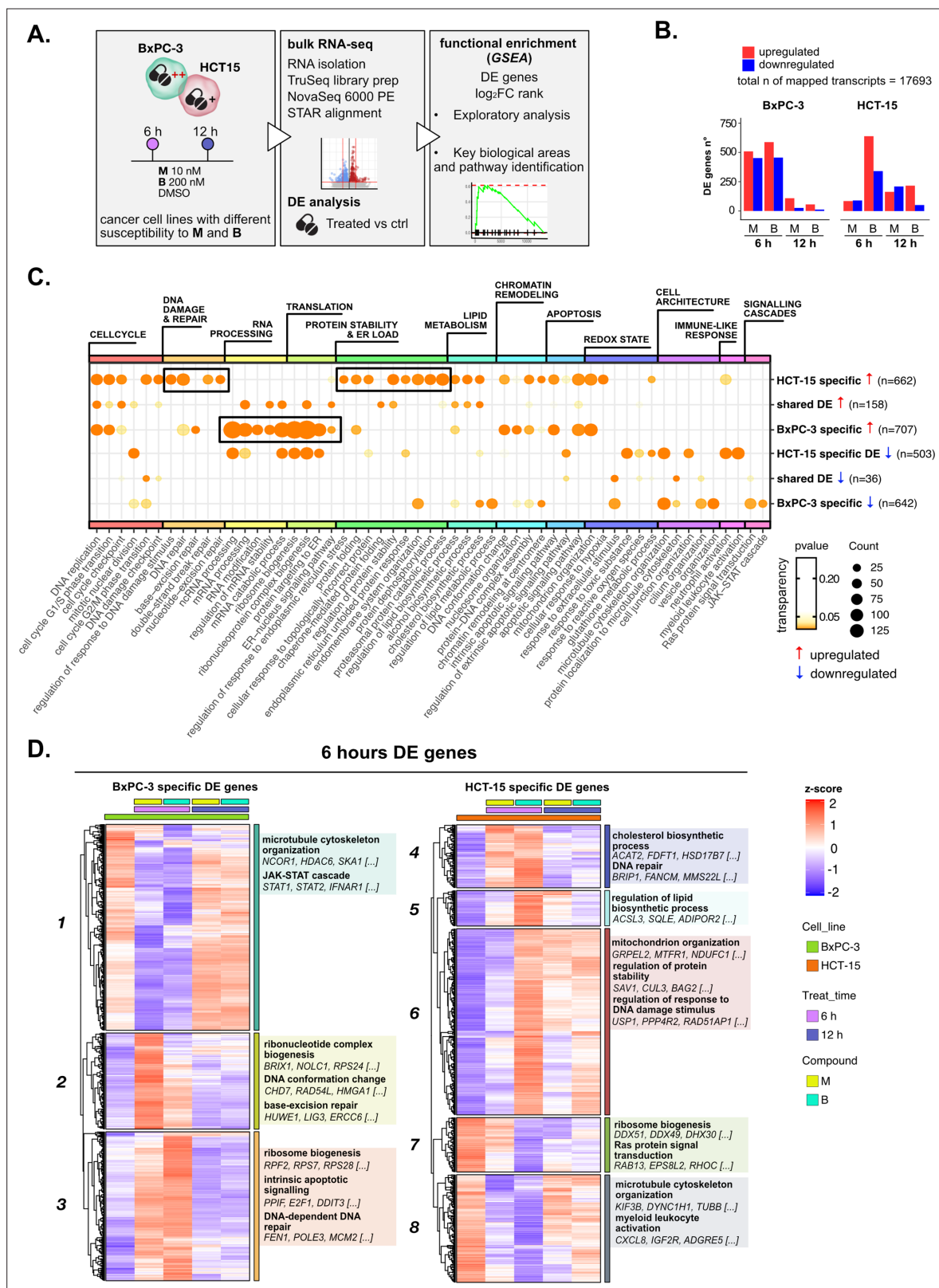
*Figure 1 continued*

(accumulation in G1 vs G2/S phases) of BxPC-3 and HCT-15 cells treated with M (10 nM), B (200 nM), or DMSO 0.5% for 6 hr and 72 hr analyzed by FACS (Fluorescence Activated Cell Sorting). At least three biological replicates were obtained per condition and unpaired two-tailed Student's *t*-test was performed to assess statistical significance ( $p < 0.05$ ). **(D)** Analysis of H2AX phosphorylation in BxPC-3 and HCT-15 cells treated with M (10 nM), B (200 nM), or DMSO 0.5% for 6 hr and 12 hr analyzed by FACS. At least three biological replicates were obtained per condition, and unpaired two-tailed Student's *t*-test was performed to assess statistical significance ( $p < 0.05$ ). **(E)** Schematic representation of the adopted omic-based approach.



**Figure 1—figure supplement 1.** Cancer tropism of 3-CePs is not explained by DNA damage—supporting data. **(A)** Flow cytometry gating strategy for the cell cycle analysis and  $\gamma$ H2AX induction. **(B)** Cell cycle distribution (accumulation in G1 vs G2/S phases) of BxPC-3 and HCT-15 cells treated with M (10 nM), B (200 nM), or DMSO 0.5% for 12 hr analyzed by FACS. At least three biological replicates were obtained per condition, and unpaired two-tailed Student's t-test was performed to assess statistical significance ( $p < 0.05$ ).



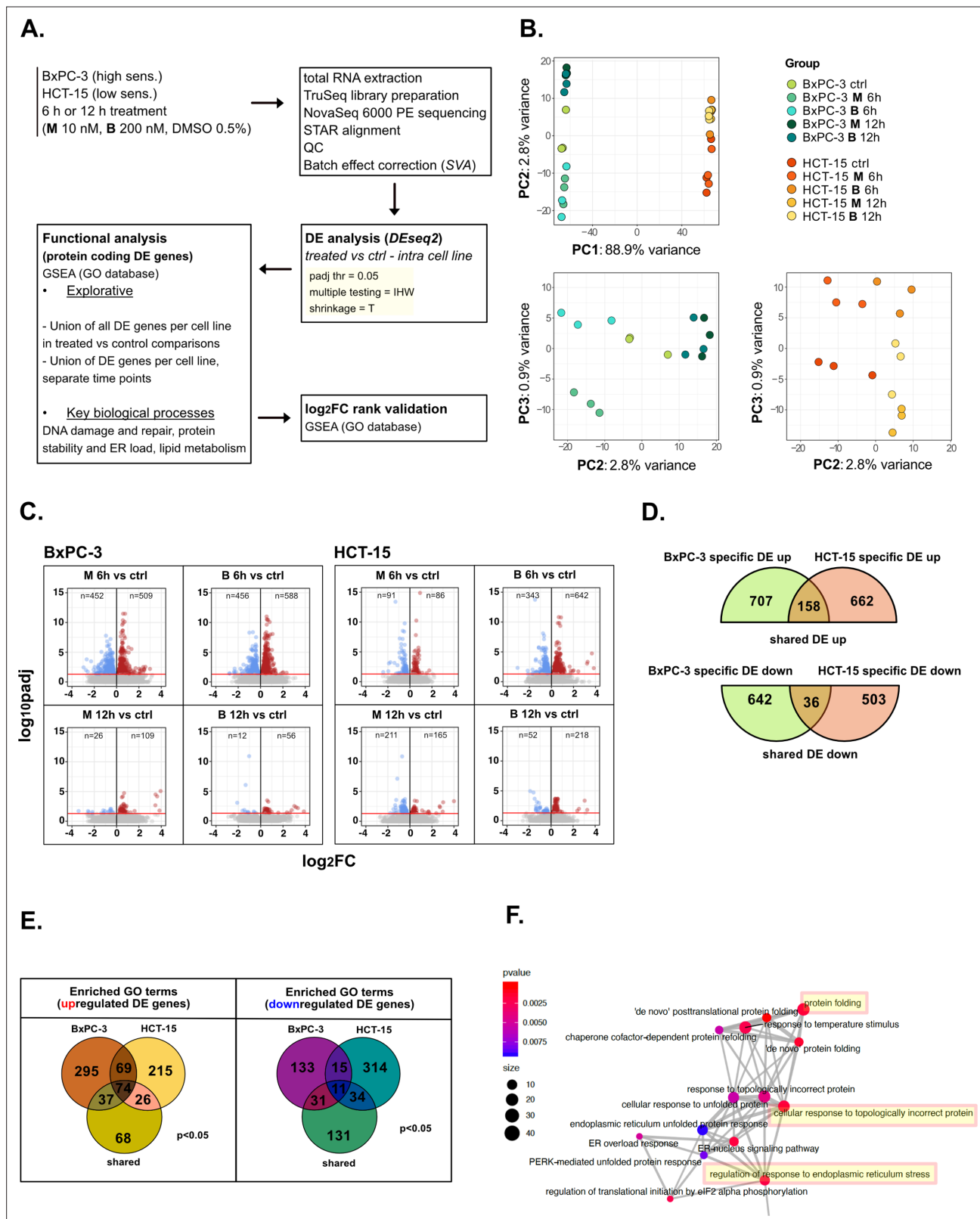


**Figure 2.** Treatments elicit cell-specific transcriptional changes. **(A)** An overview of the applied workflow for the RNA-seq analysis. **(B)** Number of up- (red) and down-regulated (blue) differential expression (DE) genes in BxPC-3 and HCT-15 cells after treatment with M (10 nM), B (200 nM), or DMSO 0.5% (ctrl) for 6 hr and 12 hr (adjusted p threshold = 0.05, shrinkage = TRUE, and multiple testing method = independent hypothesis weighting [IHW]). **(C)** Gene ontology (GO) database functional enrichment gene set enrichment analysis (GSEA) on cell-specific and shared up- and down-regulated DE

Figure 2 continued on next page

*Figure 2 continued*

genes. For each identified biological process, enrichments in terms of count and p-value of representative terms are reported ( $p < 0.05$ ). **(D)** Expression level of cell-specific 6 hr DE genes across test conditions. GSEA was performed on modules with similar regulation identified by hierarchical clustering: for each cluster, representative GO terms and genes of the associated load are reported.

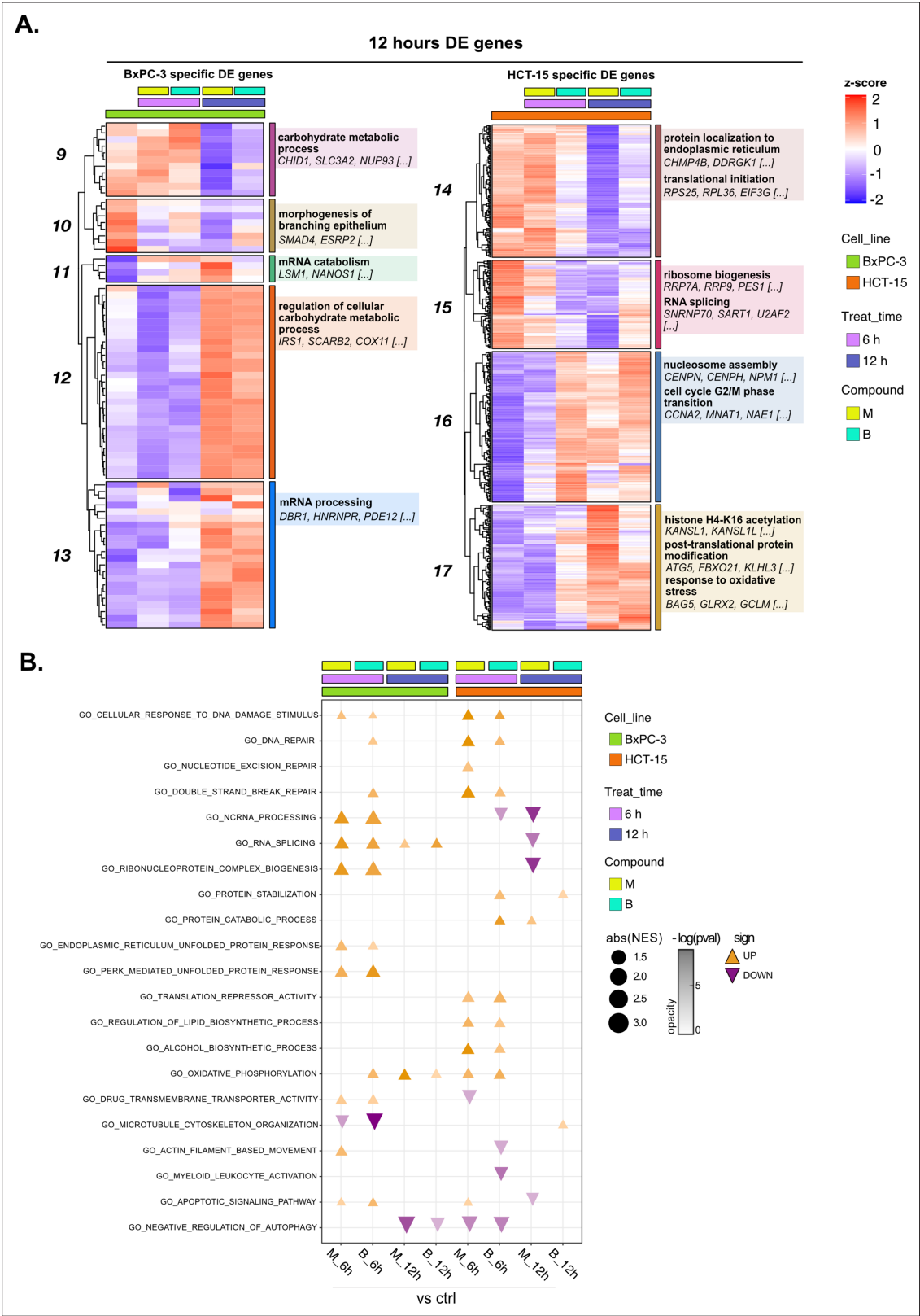


**Figure 2—figure supplement 1.** Treatments elicit cell-specific transcriptional changes—supporting data. (A) Scheme of the applied workflow for the RNA-seq analyses. (B) Principal component analysis (PCA) post surrogate variable analysis (SVA) batch correction of RNA-seq data: PC1 vs PC2 showed sample separation by cell line, PC2 vs PC3 (cell lines depicted separately) showed treatment and time point separation. (C) Volcano plots reporting up- and down-regulated differential expression (DE) genes in all treated vs control comparisons (adjusted  $p < 0.05$ ). (D) Venn plot reporting the number

Figure 2—figure supplement 1 continued on next page

*Figure 2—figure supplement 1 continued*

of specific and shared up- and down-regulated DE genes between BxPC-3 and HCT-15 cells (union of DE genes in all treated vs control comparisons). (E) Enriched gene ontology (GO) terms ( $p < 0.05$ ) derived from gene set enrichment analysis (GSEA) on BxPC-3 and HCT-15 DE genes (union of DE genes in all treated vs control comparisons) and on shared DE genes (up- and down-regulated separately). (F) Schematic representation of enrichment map-based selection of representative GO terms to be reported in **Figure 2C**.

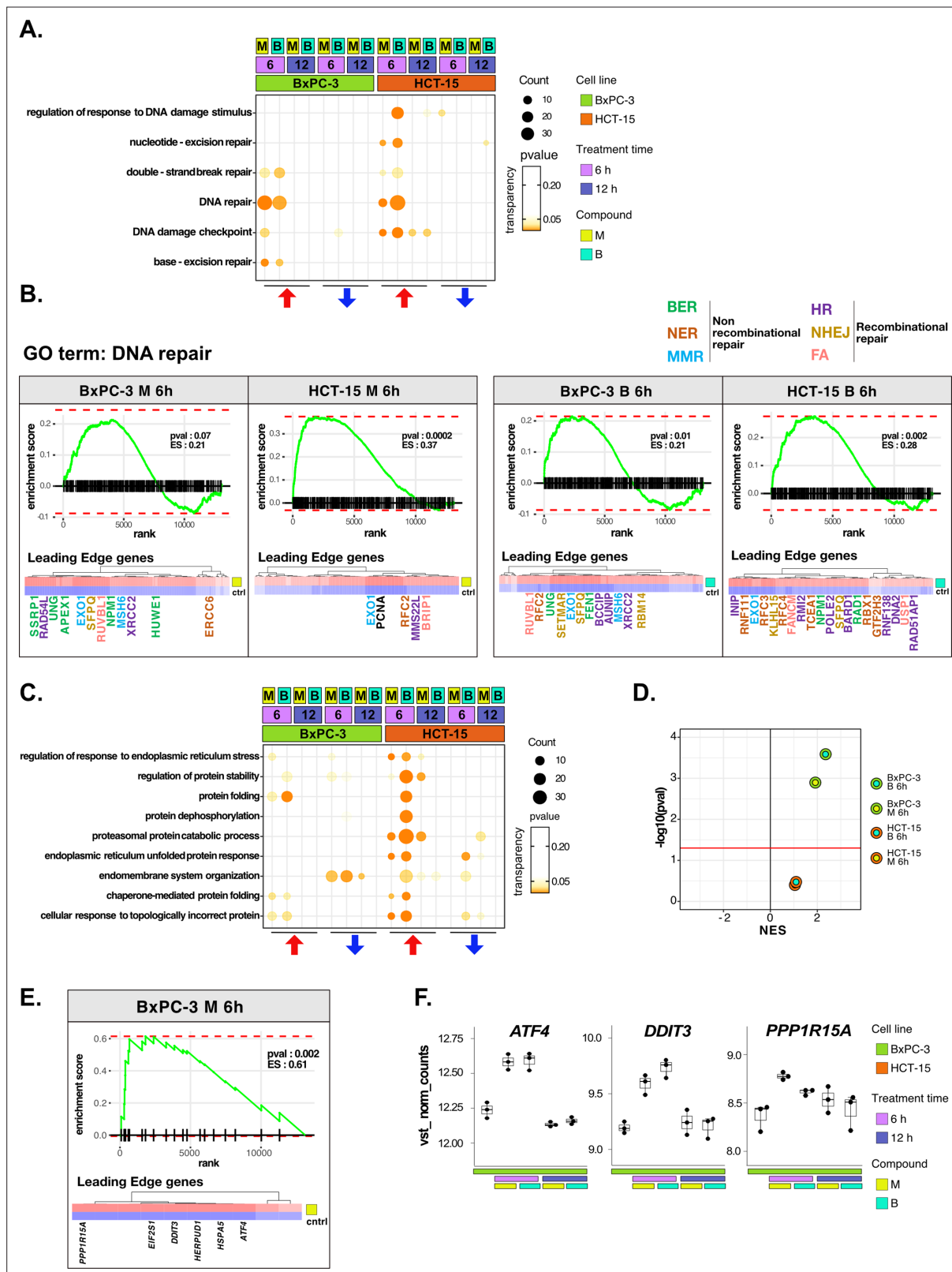


**Figure 2—figure supplement 2.** Treatments elicit cell-specific transcriptional changes—supporting data 2. **(A)** Expression level of cell-specific 12 hr differential expression (DE) genes across test conditions. Gene set enrichment analysis (GSEA) was performed on modules with similar regulation identified by hierarchical clustering: for each cluster, representative gene ontology (GO) terms and genes of the associated load are reported. **(B)** GO database functional enrichment (GSEA) obtained from log<sub>2</sub>FC ranks in all treated vs control comparison both in BxPC-3 and HCT-15 cells. For each

Figure 2—figure supplement 2 continued on next page

Figure 2—figure supplement 2 continued

identified biological process, enrichments in terms of absolute normalized enrichment score (abs[NES]) and  $-\log(p)$  of representative terms are reported ( $p < 0.05$ ).



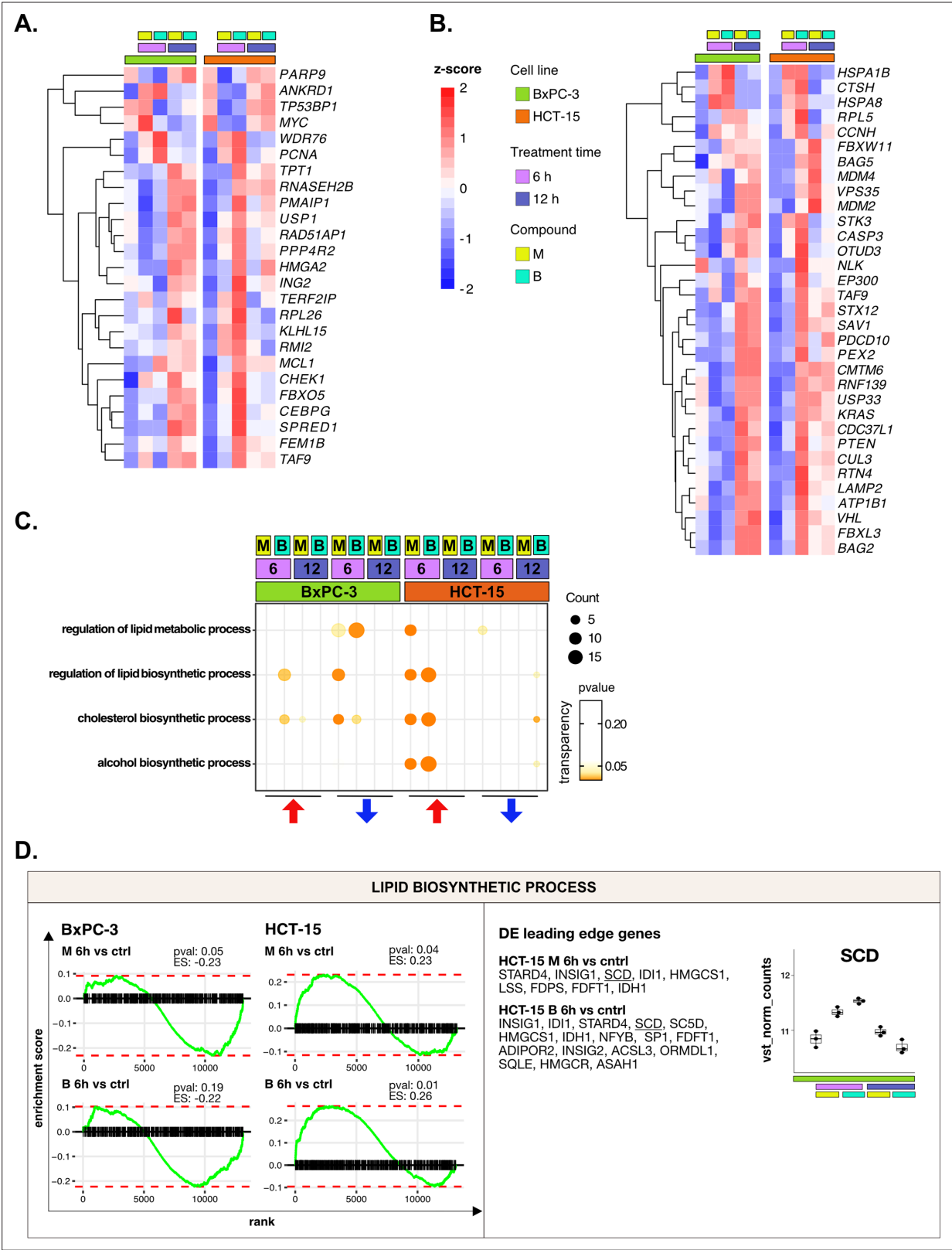
**Figure 3.** DNA repair and proteostasis are key modulators of the response to 3-chloropiperidines (3-CePs). **(A)** Gene set enrichment analysis (GSEA) for terms related to DNA damage and repair performed on differential expression (DE) genes detected in each of the considered treated vs control comparisons. For each gene ontology (GO) term ( $p < 0.05$ ), enrichments in terms of count and p-value are reported. **(B)** GSEA enrichment plots for the DNA repair pathway obtained from  $\log_2FC$  ranks for each of the considered treated vs control comparisons. The expression of leading edge genes

Figure 3 continued on next page

## Figure 3 continued

is also shown, where key DE genes are reported with the same color of their associated DNA repair pathways (BER = base excision repair, NER = nucleotide-excision repair, MMR = mismatch repair, HR = homologous recombination, NHEJ = non-homologous end joining, and FA = Fanconi anemia pathway) (Cantor et al., 2001; Charles Richard et al., 2016; Ferretti et al., 2016; Fortini et al., 2003; Fousteri and Mullenders, 2008; Jaafar et al., 2017; Johnson et al., 1999; Kunkel and Erie, 2005; Liu and Kong, 2021; Li and Jin, 2012; Li et al., 2009a; Lou et al., 2017; Lu et al., 2007; Mazin et al., 2010; McVey et al., 2016; Nijman et al., 2005; Niraj et al., 2019; Overmeer et al., 2010; Parsons et al., 2009; Pascucci et al., 2018; Piwko et al., 2016; Poletto et al., 2014; Poulsen et al., 2013; Prasad et al., 2000; Rajendra et al., 2014; Satoh and Hanawalt, 1996; Smirnova et al., 2005; Tellier and Chalmers, 2019; Westermarck et al., 2003; Xu et al., 2008; Yard et al., 2016). (C) GSEA for terms related to protein stability and endoplasmic reticulum (ER) load performed on DE genes detected in each comparison. For each GO term, enrichments in terms of count and p-value are reported. (D) NES (normalized enrichment score) and  $-\log_{10}$  pval for the  $\log_2$ FC rank-based GSEA enrichment of the GO term PERK-mediated unfolded protein response (UPR) in treated vs control comparisons. (E) GSEA enrichment plot for the PERK-mediated UPR pathway obtained from  $\log_2$ FC rank in the M 6 hr vs control comparison in BxPC-3 cells. The expression of leading edge genes is also shown, where key DE genes of the mentioned pathway are reported. (F) Boxplots showing the expression level of *ATF4*, *DDIT3*, and *PPP1R15A* (vst-transformed normalized counts) in BxPC-3 cells (n=3).



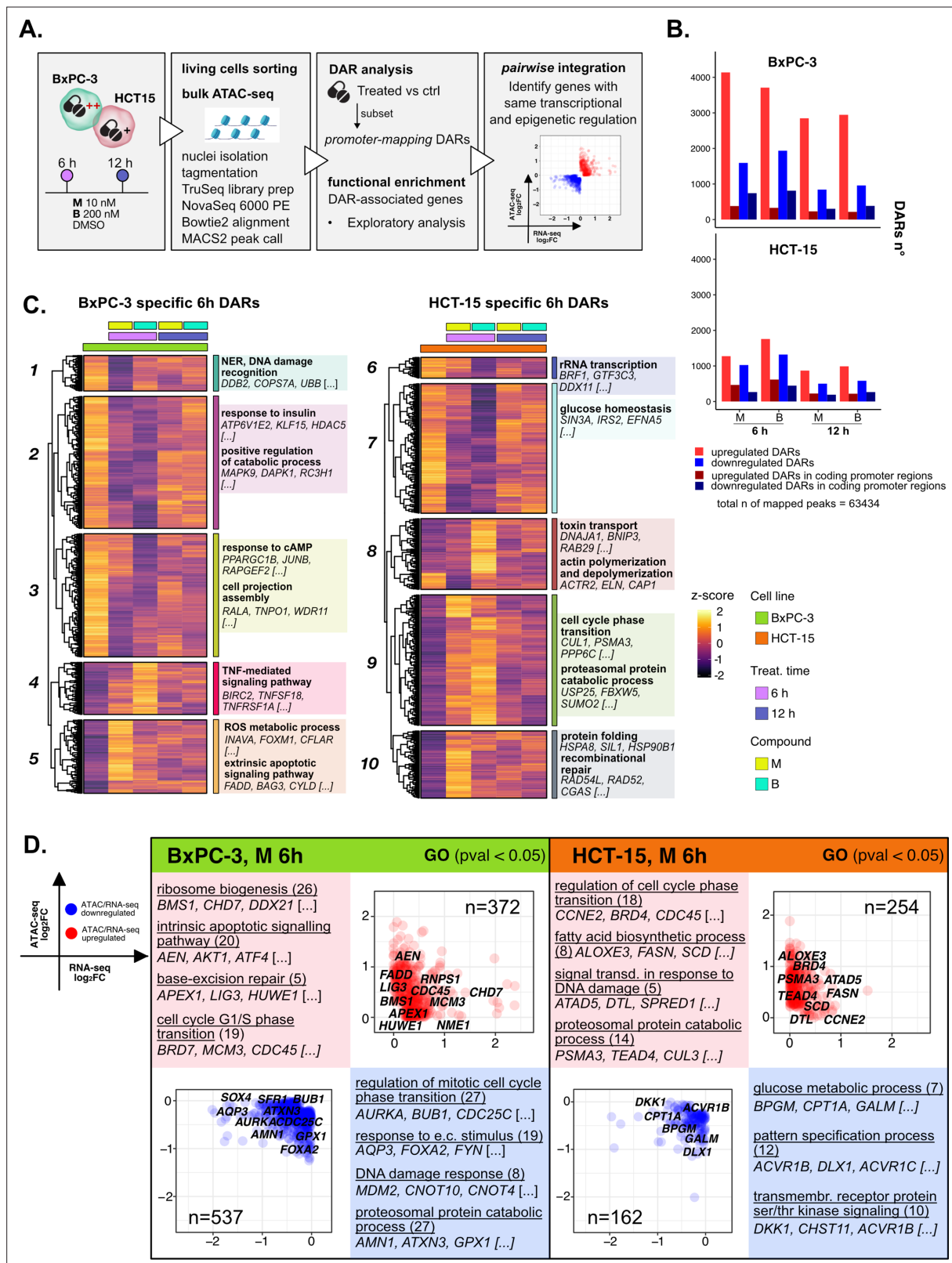


**Figure 3—figure supplement 1.** DNA repair and proteostasis are key modulators of the response to 3-chloropiperidines (3-CePs)—supporting data. (A) Expression level of differential expression (DE) genes included in the load of the gene ontology (GO) term regulation of response to DNA damage stimulus (HCT-15 cells,  $p < 0.05$ ) in BxPC-3 and HCT-15 cells. (B) Expression level of DE genes included in the load of the GO term proteasomal protein catabolic process (HCT-15 cells,  $p < 0.05$ ) in BxPC-3 and HCT-15 cells. (C) Gene set enrichment analysis (GSEA) for terms related to lipid metabolism

Figure 3—figure supplement 1 continued on next page

*Figure 3—figure supplement 1 continued*

performed on DE genes detected in each of the considered treated vs control comparisons. For each GO term ( $p < 0.05$ ), enrichments in terms of count and p-value are reported. **(D)** GSEA enrichment plots for the DNA repair pathway obtained from  $\log_2$ FC ranks for each of the considered treated vs control comparisons. Key DE genes from the leading edge are reported, together with the expression of SCD in HCT-15 cells (normalized vst-transformed batch corrected counts,  $n=3$ ).

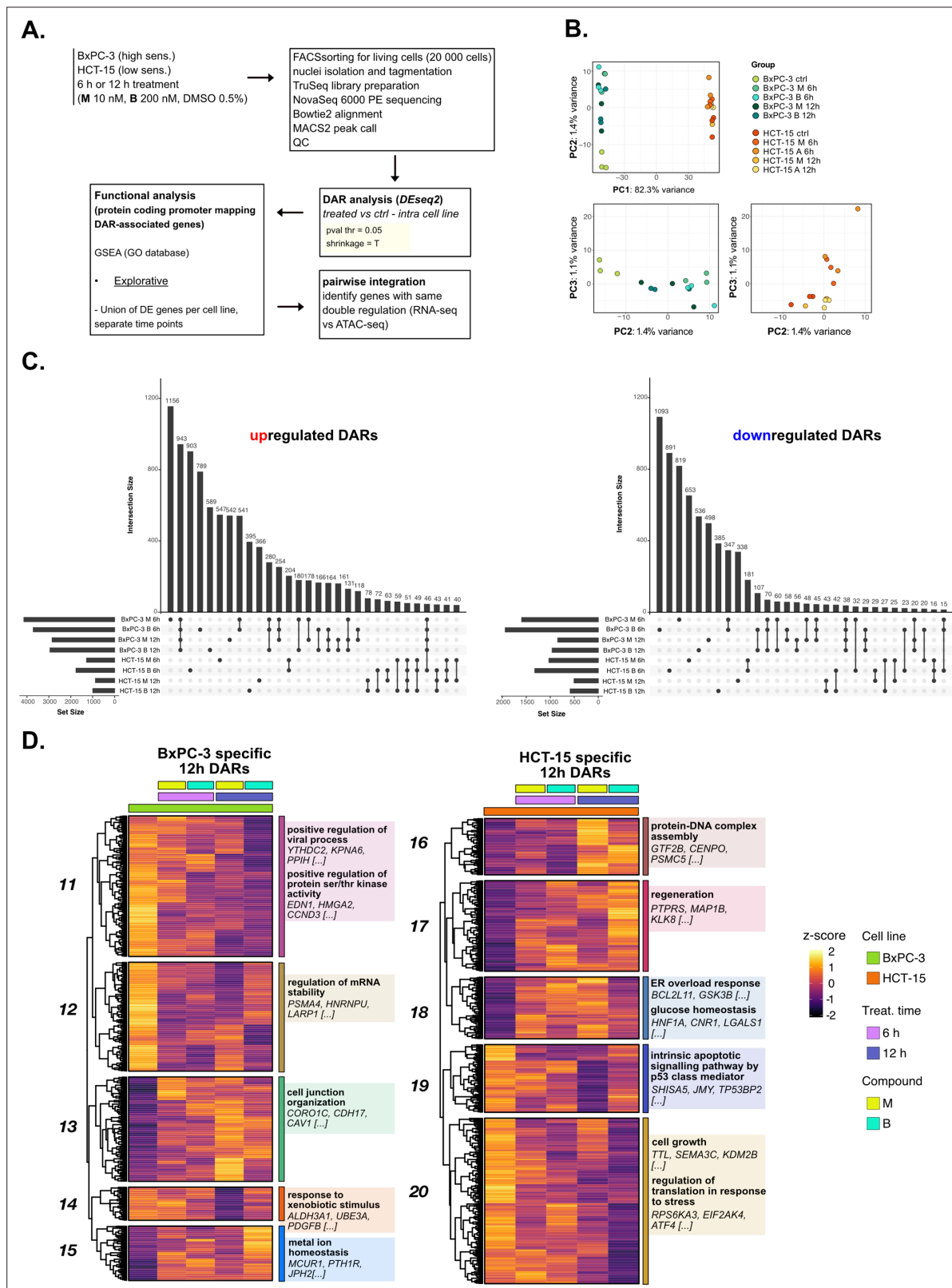


**Figure 4.** The response to 3-chloropiperidines (3-CePs) is further regulated at the chromatin level. **(A)** An overview of the applied workflow for the ATAC-seq analysis. **(B)** Number of up- (red) and down-regulated (blue) differentially accessible regions (DARs) in BxPC-3 and HCT-15 cells after treatment with M (10 nM), B (200 nM), or DMSO 0.5% (ctrl) for 6 hr and 12 hr (p-value threshold = 0.05, shrinkage = TRUE). Light blue/red = all detected DARs, dark blue/red = protein coding DARs mapping in promoter regions. **(C)** Accessibility level of cell-specific 6 hr DARs across test conditions. Gene

Figure 4 continued on next page

*Figure 4 continued*

set enrichment analysis (GSEA) was performed on genes associated with DARs with similar regulation, grouped in modules identified by hierarchical clustering; for each cluster, representative gene ontology (GO) terms and genes of the associated load are reported. **(D)** Pairwise integration: ratio-ratio plots report the RNA- and ATAC-seq  $\log_2$ FCs of genes showing the same direction of transcriptional and chromatin accessibility regulation and their GSEA. Integration was performed not only at the same time point of 6 hr in both omic layers. For each GO term ( $p < 0.05$ ), enrichments in terms of count and p-value are reported.

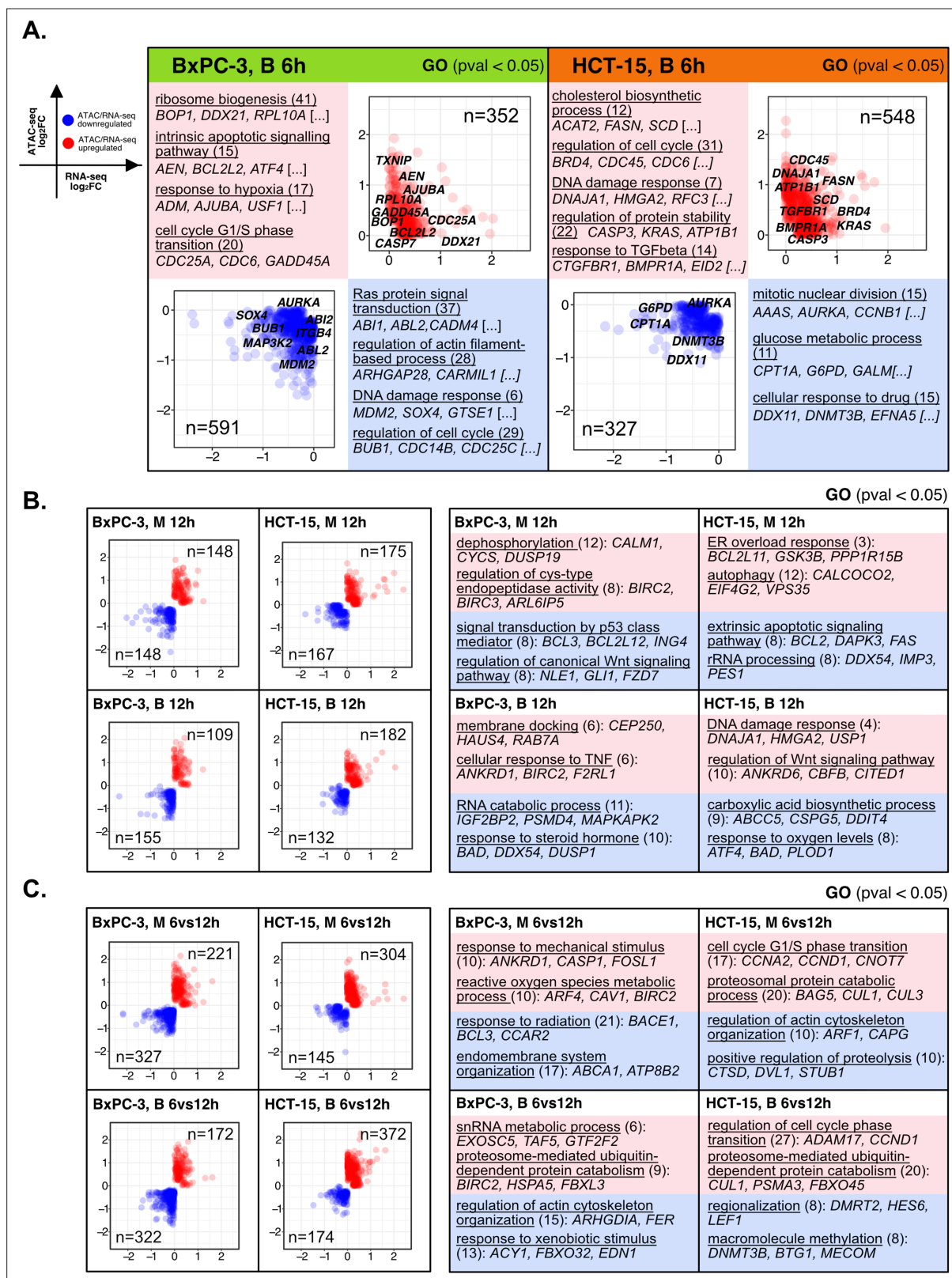


**Figure 4—figure supplement 1.** The response to 3-chloropiperidines (3-CePs) is further regulated at the chromatin level—supporting data. **(A)** Scheme of the applied workflow for the ATAC-seq analyses. **(B)** Principal component analysis (PCA) of ATAC-seq data: PC1 vs PC2 showed samples separation by cell line and PC2 vs PC3 (cell lines depicted separately) showed treatment and time point separation. **(C)** Upset plots reporting up- and down-regulated differentially accessible regions (DARs) ( $p < 0.05$ ) and their overlap between all treated and control comparisons in both cell lines. **(D)** Accessibility level

Figure 4—figure supplement 1 continued on next page

*Figure 4—figure supplement 1 continued*

of cell-specific 12 hr DARs across test conditions. Gene set enrichment analysis (GSEA) was performed on genes associated with DARs with similar regulation, grouped in modules identified by hierarchical clustering: for each cluster, representative gene ontology (GO) terms and genes of the associated load are reported.



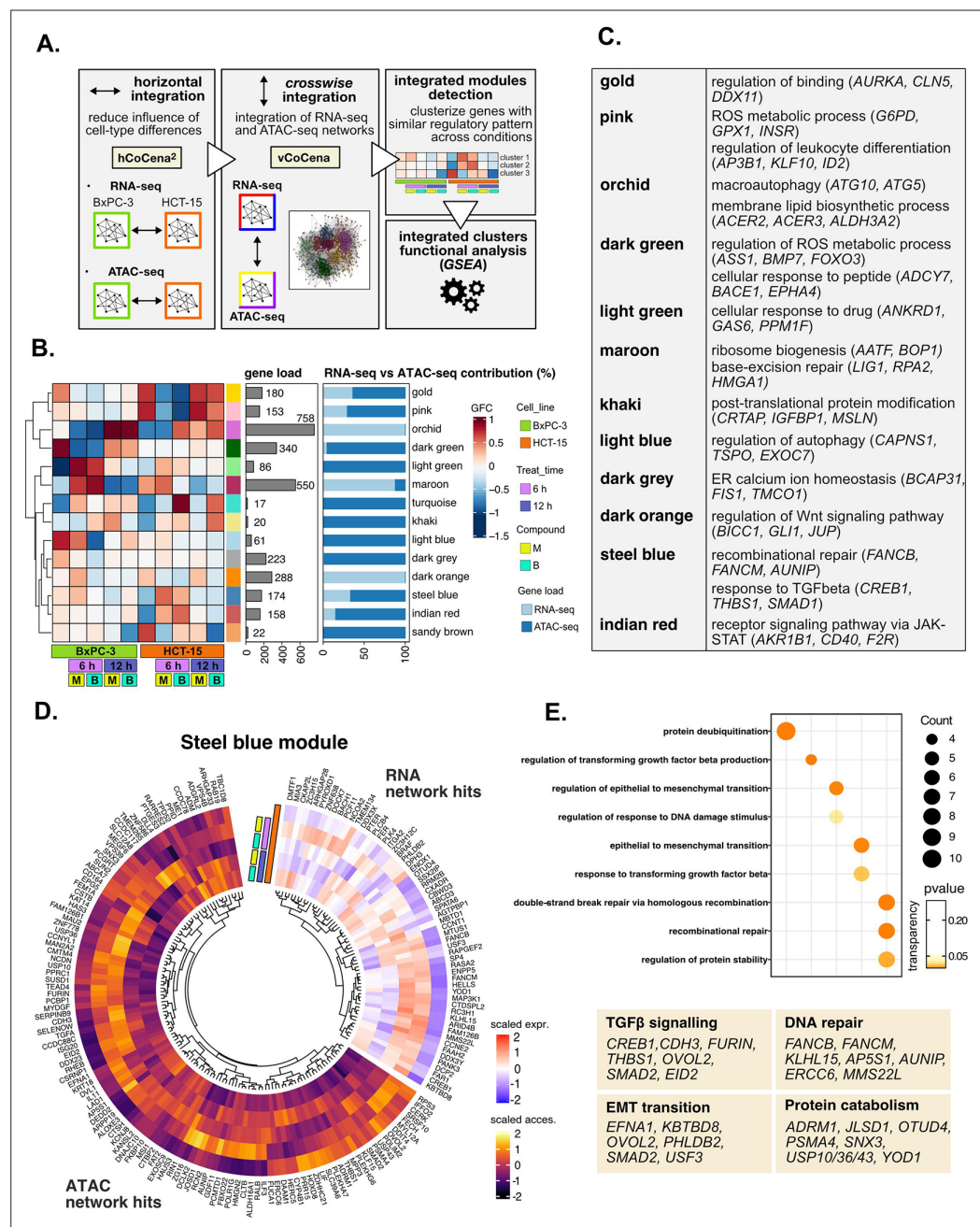
**Figure 4—figure supplement 2.** The response to 3-chloropiperidines (3-CePs) is further regulated at the chromatin level—supporting data 2. **(A)** Pairwise integration: ratio-ratio plots report the RNA- and ATAC-seq log<sub>2</sub>FCs of genes showing the same direction of transcriptional and chromatin accessibility regulation. Integration was performed not at the same time point of 6 hr in both omic layers for compound B. Associated key gene set enrichment analysis (GSEA) terms with p<0.05 are also reported. **(B)** Pairwise integration performed at the same time point of 12 hr in both omic layers

Figure 4—figure supplement 2 continued on next page

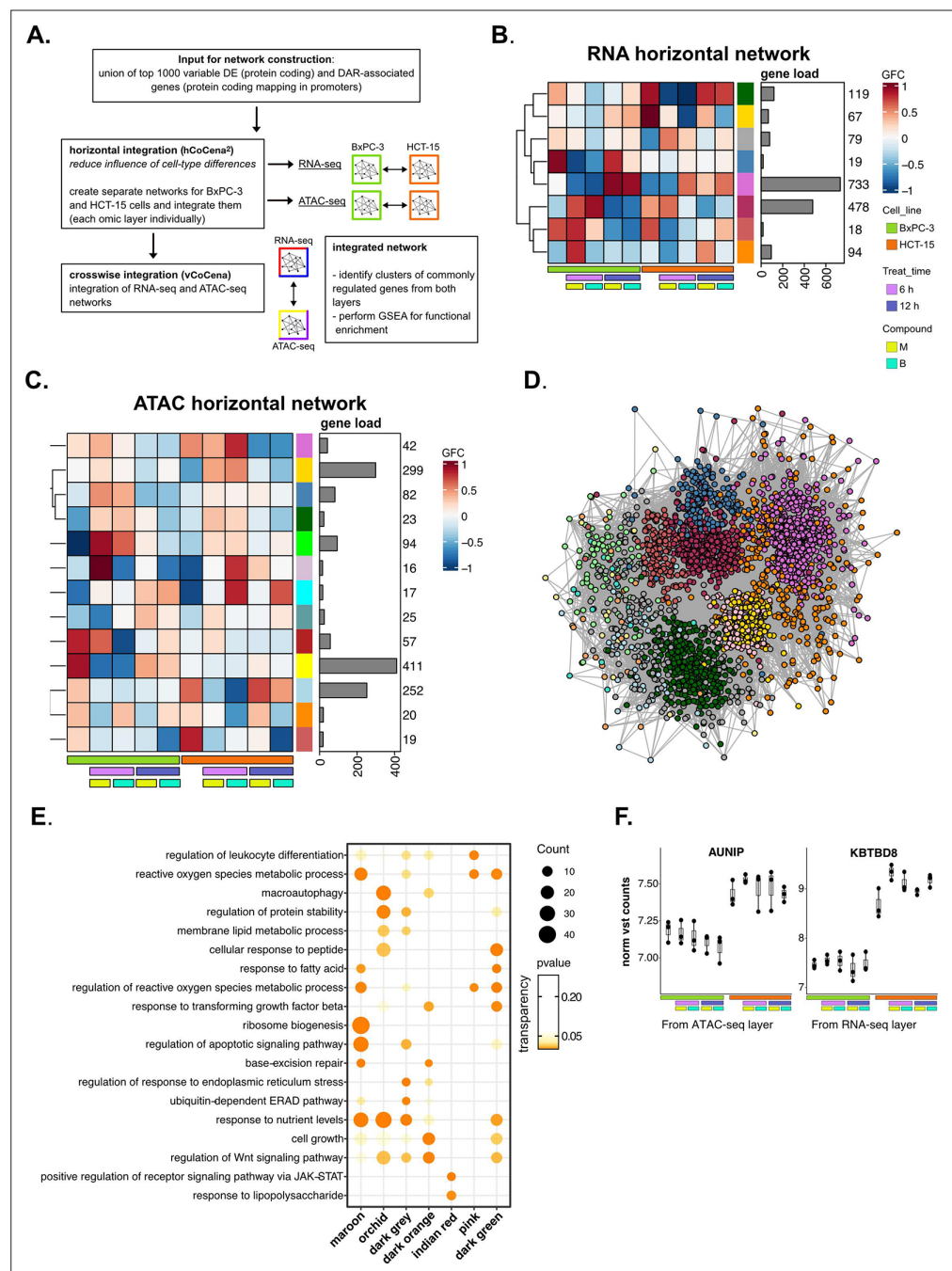
Figure 4—figure supplement 2 continued

in response to M and B. Associated key GSEA terms with  $p < 0.05$  are also reported. **(C)** The integration was also performed between chromatin changes at 6 hr and transcriptional responses at 12 hr. Associated key GSEA terms with  $p < 0.05$  are also reported.

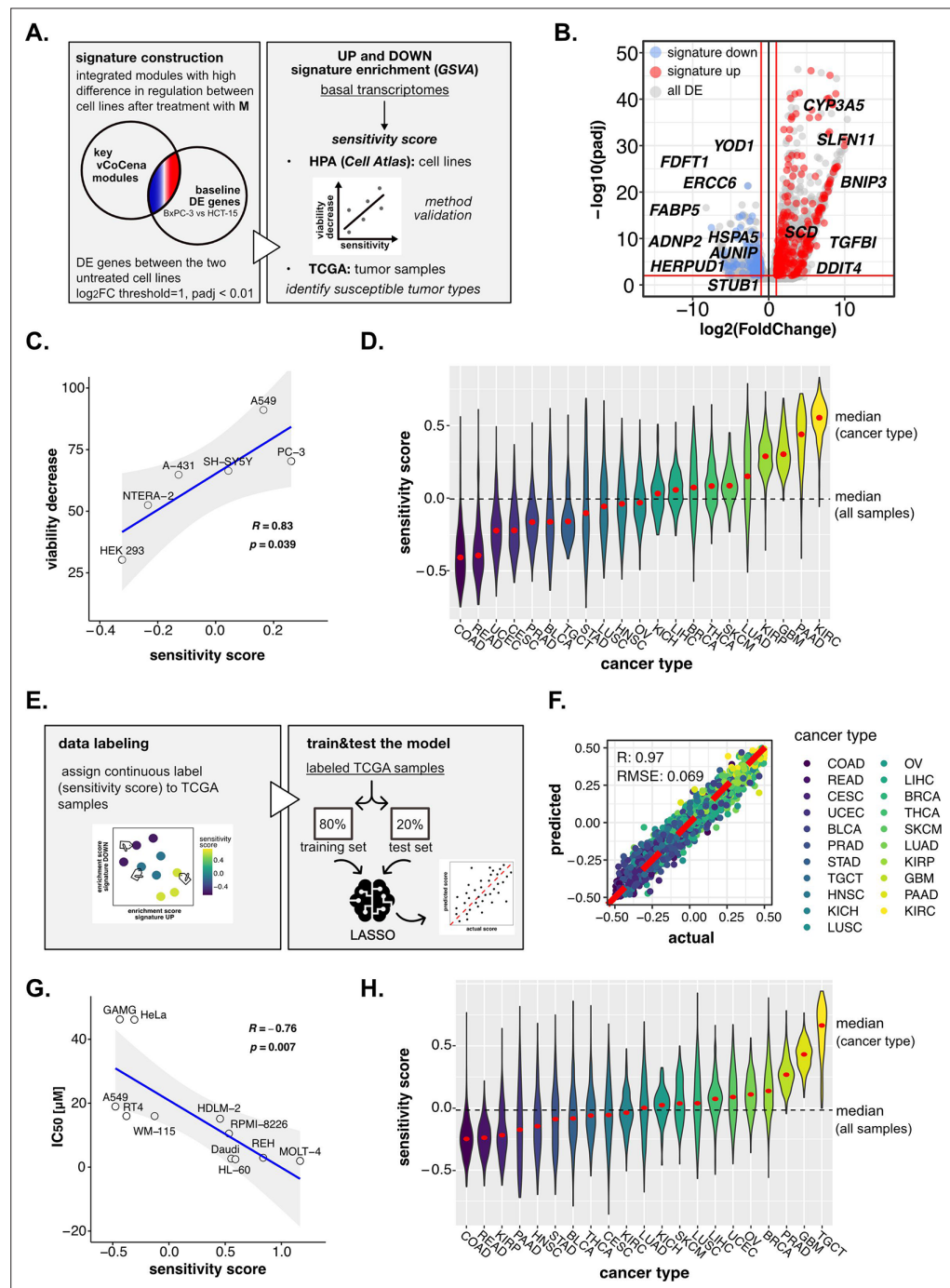




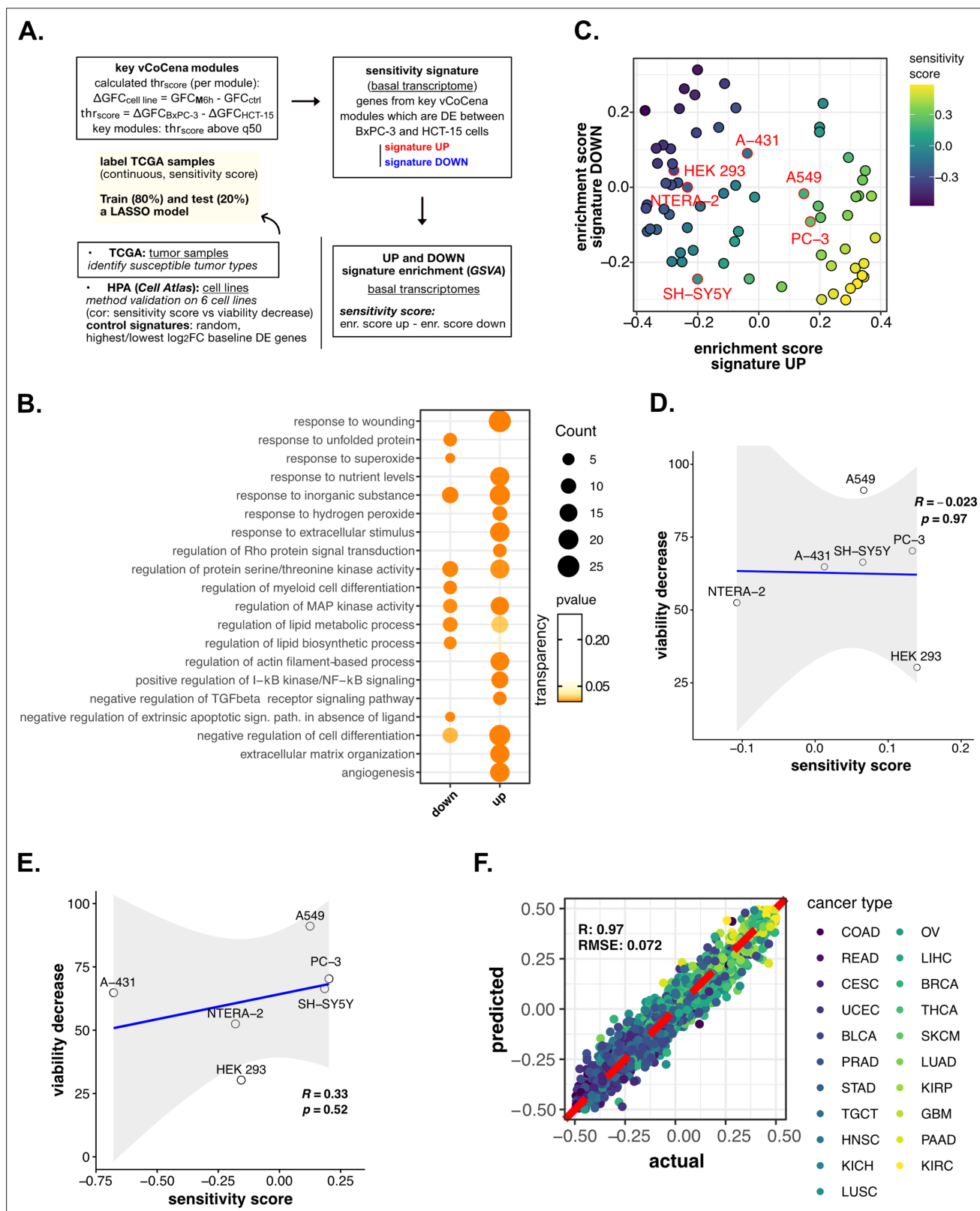
**Figure 5.** Crosswise integration expedites the comprehension of multi-omic data. **(A)** Overview of the applied workflow for the crosswise integration analysis. **(B)** Integrated modules of genes from the RNA- and ATAC-seq layers obtained with vertical construction of co-expression network analysis (vCoCena) and associated group fold change (GFC) pattern of regulation across conditions. The relative contribution of hits from the RNA- or ATAC-seq layers is also reported for each module. **(C)** Representative gene ontology (GO) terms (p < 0.05) for the most relevant modules of genes, identified by gene set enrichment analysis (GSEA). Enrichments in terms of count and p-value are reported. **(D)** Expression and chromatin accessibility levels in HCT-15 cells of genes included in the *steelblue* module (nodes can come from the RNA- or ATAC-seq layer). **(E)** Most representative GO terms from GSEA on genes of the *steelblue* module (key areas: EMT (Epithelial-Mesenchymal Transition), protein catabolism, TGF β signaling, and DNA repair). For each GO term (p < 0.05), enrichments in terms of count and p-value are reported.



**Figure 5—figure supplement 1.** Crosswise integration expedites the comprehension of multi-omic data—supporting data. (A) Scheme of the applied workflow for the crosswise integration analysis. (B) Horizontally integrated modules of genes from the RNA-seq layer and associated group fold change (GFC) pattern of regulation across conditions. (C) Horizontally integrated modules of genes from the ATAC-seq layer and associated GFC pattern of regulation across conditions. (D) Crosswise integrated vertical construction of co-expression network analysis (vCoCena) network. (E) Most representative gene ontology (GO) terms from gene set enrichment analysis (GSEA) on genes of the maroon, orchid, dark orange, dark grey, indian red, pink, and dark green modules. For each GO term ( $p < 0.05$ ), enrichments in terms of count and p-value are reported. (F) Boxplots showing the expression level of AUNIP and KBTBD8 (vst-transformed normalized counts, batch corrected in RNA-seq,  $n=3$ ) in BxPC-3 and HCT-15 cells.



**Figure 6.** Perturbation-informed basal signatures efficiently predict sensitivity to our candidate drugs. **(A)** Overview of the applied workflow for the sensitivity signature construction and associated drug susceptibility prediction. **(B)** M sensitivity signature genes (red = signature up and blue = signature down) pinpointed from all differential expression (DE) genes in the BxPC-3 vs HCT-15 baseline comparison. **(C)** Pearson correlation between predicted sensitivity score and viability decrease in a subset of human protein atlas (HPA) (cell atlas) cell lines (validation set). **(D)** Sensitivity scores predicted from gene set variation analysis (GSVA) enrichment of our up and down signatures in RNA-seq profiles of TCGA tumor samples. Median values for all sample scores and within each tumor type are reported. **(E)** Overview of the applied workflow for the LASSO-based ML setup. **(F)** Predictive outcome of the trained model (Pearson correlation R and RMSE (root-mean-square error) are reported). **(G)** Pearson correlation between predicted cisplatin sensitivity score and  $\text{IC}_{50}$  in a subset of HPA (cell atlas) cell lines (validation set). **(H)** Cisplatin sensitivity scores predicted from GSVA enrichment of our up and down signatures in RNA-seq profiles of TCGA tumor samples. Median values for all sample scores and within each tumor type are reported.



**Figure 6—figure supplement 1.** Perturbation-informed basal signatures efficiently predict sensitivity to our candidate drugs—supporting data.

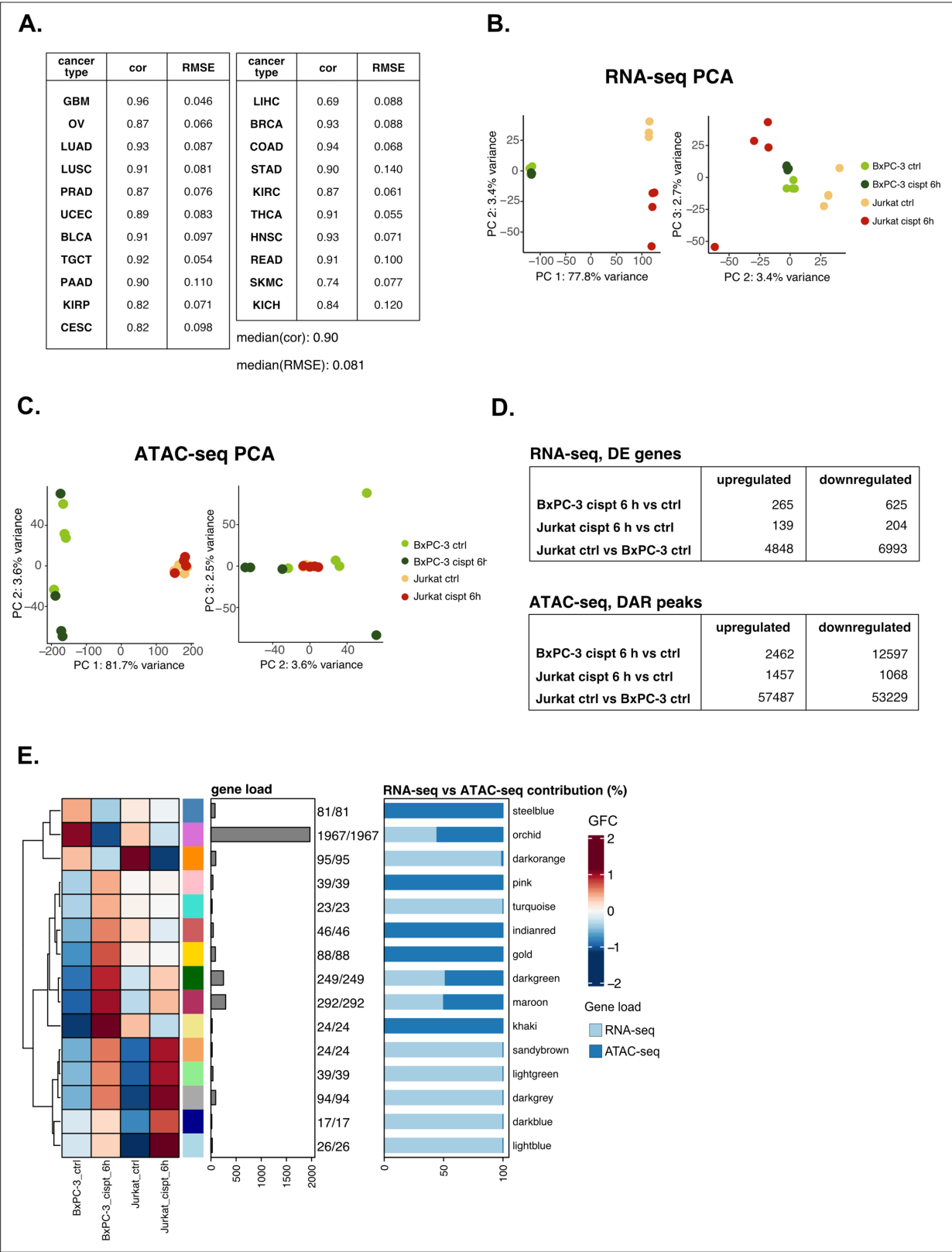
(A) Scheme of the applied workflow for the sensitivity signature construction and associated drug susceptibility prediction. (B) Representative gene ontology (GO) terms ( $p < 0.05$ ) for genes of the M sensitivity signature (up and down), identified by gene set enrichment analysis (GSEA). Enrichments in terms of count and p-value are reported. (C) Human protein atlas (HPA) (cell atlas) cell lines separation based on gene set variation analysis (GSVA)

Figure 6—figure supplement 1 continued on next page

*Figure 6—figure supplement 1 continued*

enrichment of our newly constructed up vs down signatures of sensitivity to M. Color scale reflects samples predicted sensitivity score (up signature enrichment - down signature enrichment). **(D)** Pearson correlation between predicted sensitivity score and viability decrease in a subset of HPA (cell atlas) cell lines (validation set) using a random signature. **(E)** Pearson correlation between predicted sensitivity score and viability decrease in a subset of HPA (cell atlas) cell lines (validation set) using a control signature composed by differential expression (DE) genes with top up and down  $\log_2$ FC. **(F)** Predictive performance after the exclusion of genes belonging to our signature from training and test set transcriptomes (Pearson correlation R and RMSE are reported).





**Figure 6—figure supplement 2.** Validation analysis on cisplatin sensitivity confirms the versatility of the pipeline. **(A)** Predicting performance of our model on tumor-specific subtypes (Pearson's correlation and RMSE are reported). **(B)** RNA-seq principal component analysis (PCA) (PCs 1 vs 2 and 2 vs 3) for the cisplatin dataset. **(C)** ATAC-seq PCA (PCs 1 vs 2 and 2 vs 3) for the cisplatin dataset. **(D)** Identified number of differential expression (DE) and differentially accessible regions (DARs) in all comparisons for the cisplatin datasets using parameters described in Methods section. **(E)** Crosswise integrated cisplatin vertical construction of co-expression network analysis (vCoCena) network.