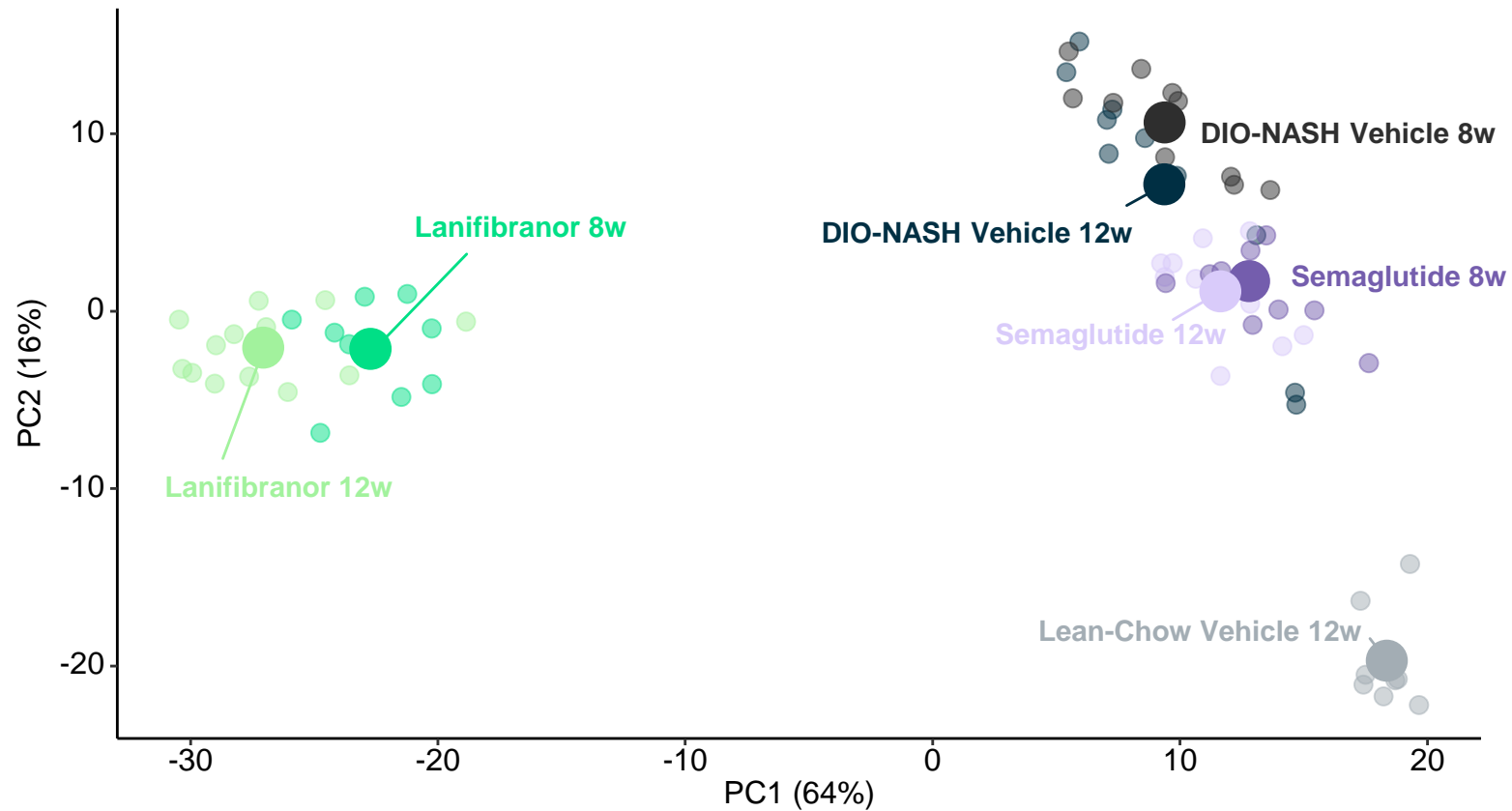**DATA REPORT**

RNA-seq data package

Based on data published in Møllerhøj et al.,

gubra

SCIENCE OF CERTAINTY

# Principal component analysis

Sample variation within and between study groups

Principal component analysis based on the top 500 most varying genes across the dataset. The axes describe the percentage of variability explained by the principal component 1 (PC1) and principal component 2 (PC2). Each small point represents a sample, and each big point represents the study group average. Samples with similar gene expression profile will locate closer together than samples with different gene expression profiles. Thus, if there is pronounced gene expression differences between groups, the samples will cluster together with members of the same group and away from samples from other groups.

Differential gene expression analysis
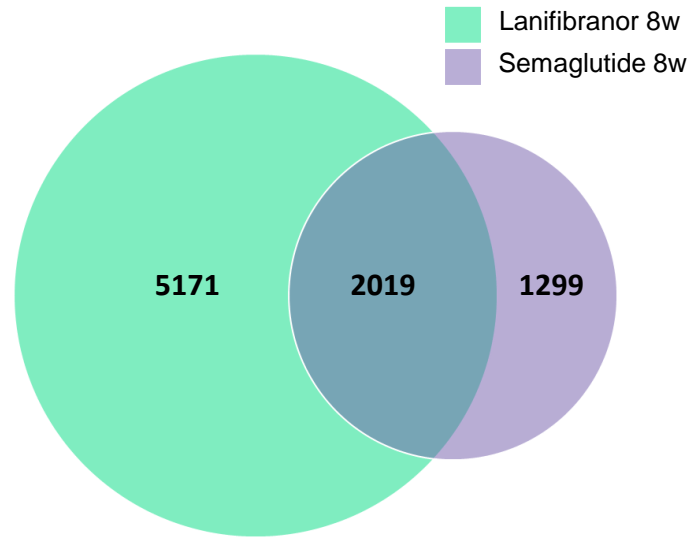
Number of differentially expressed genes
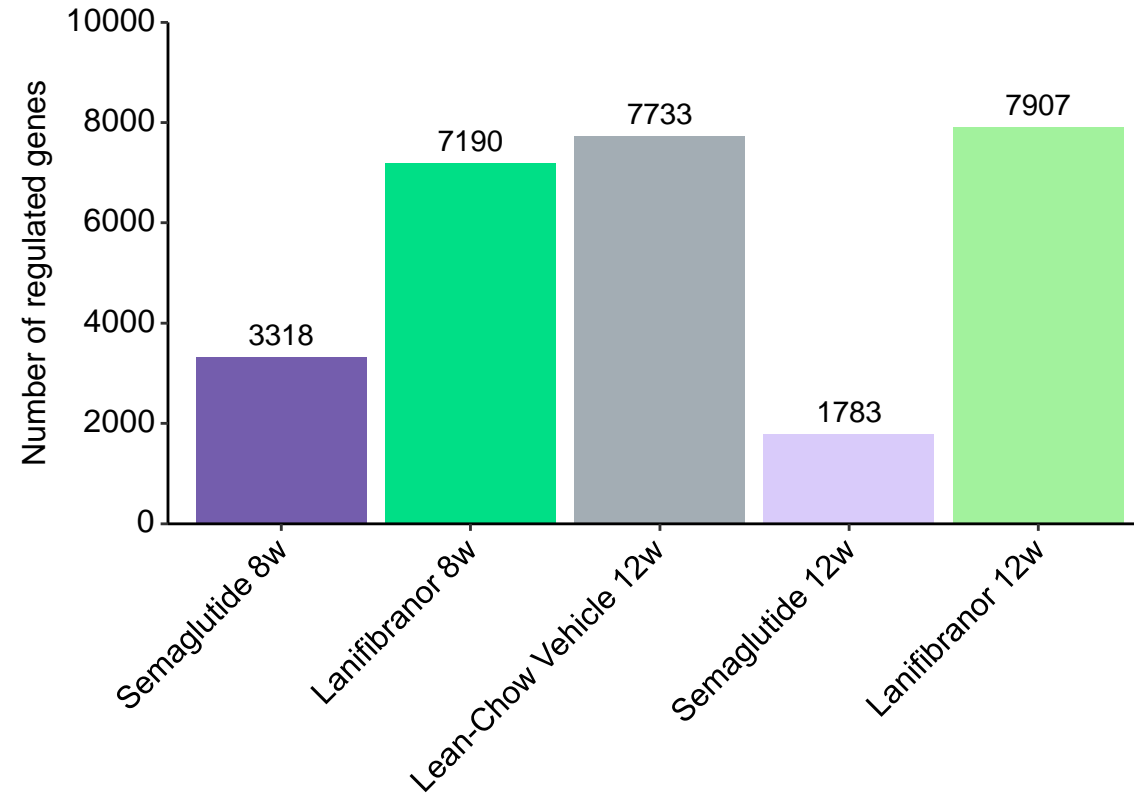
Number of differentially expressed genes for each study group compared to the corresponding DIO-NASH
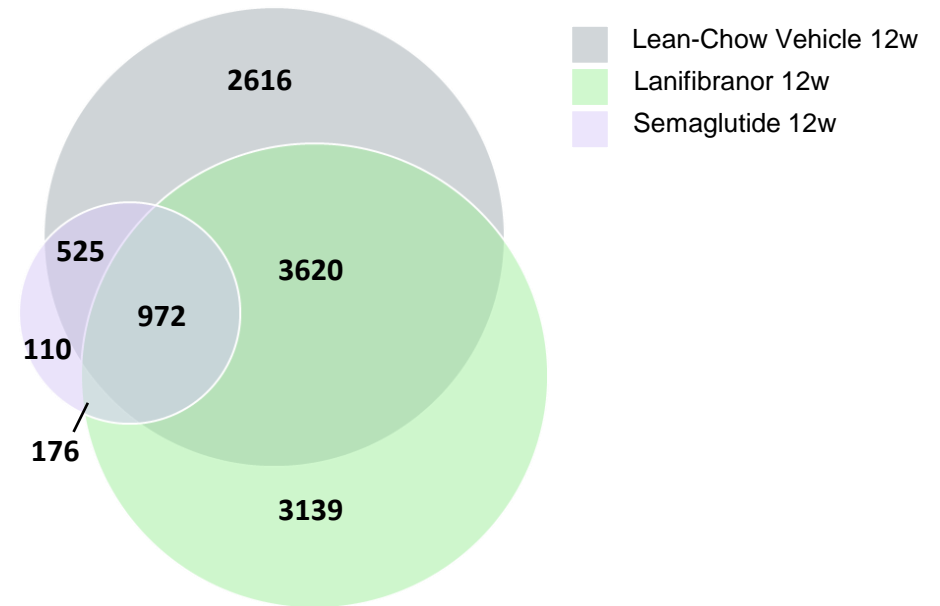Vehicle at significance level padj < 0.05 after correcting for multiple testing. n = 8 - 13.

Gene regulation overlaps
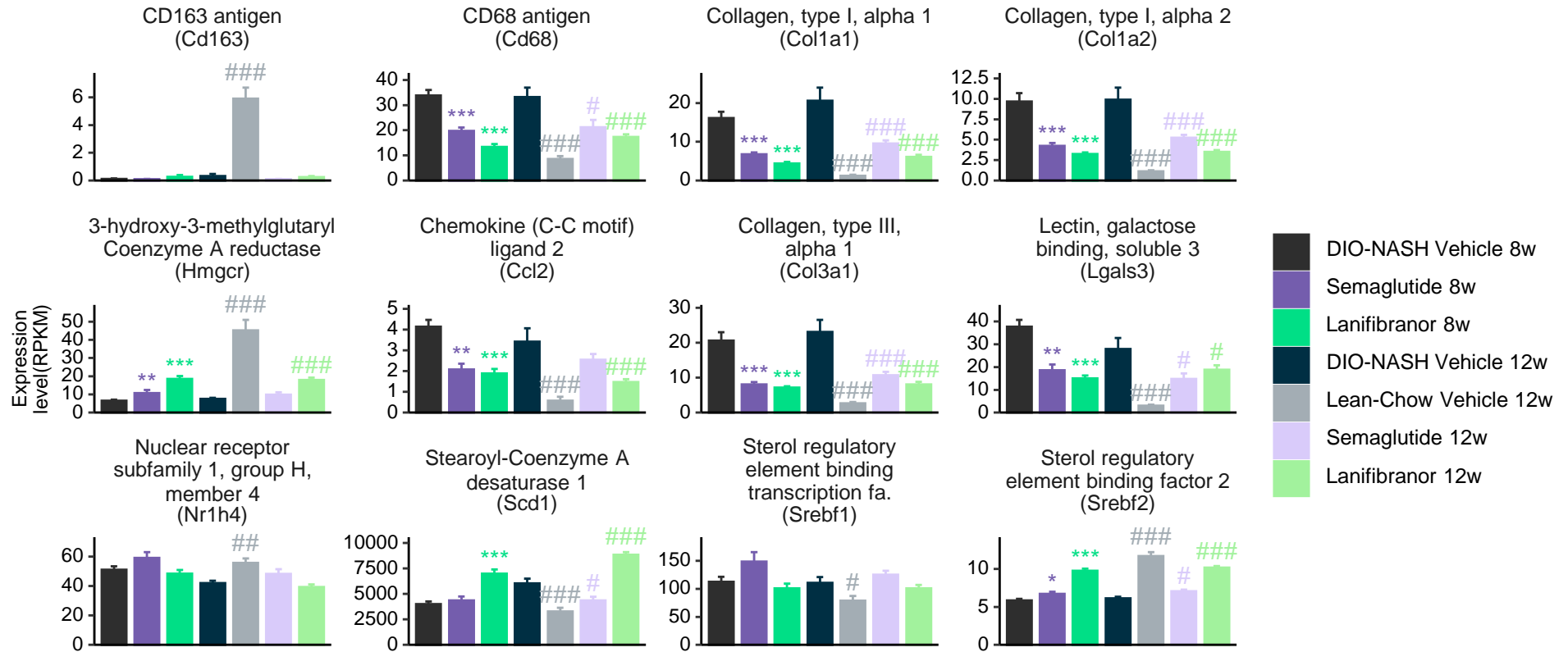
Overview of shared regulations

Week 8

Week 12

Number of differentially expressed genes regulated by either one or more study groups when compared to the corresponding DIO-NASH Vehicle. n = 8 - 13.

MASH: Genes of interest

Selected genes associated with NASH biology

Gene expression of selected genes associated with MASH biology. Values are expressed as study group mean + SEM. */#:padj < 0.05, **/##:padj < 0.01, ***/###:padj < 0.001 compared to the corresponding DIO-NASH Vehicle after correcting for multiple testing. RPKM: Reads Per Kilobase Million. n = 8 - 13.

Example: RNA-seq data report. All data belong to Gubra.

Gubra curated gene panels

Regulation of disease-related genes
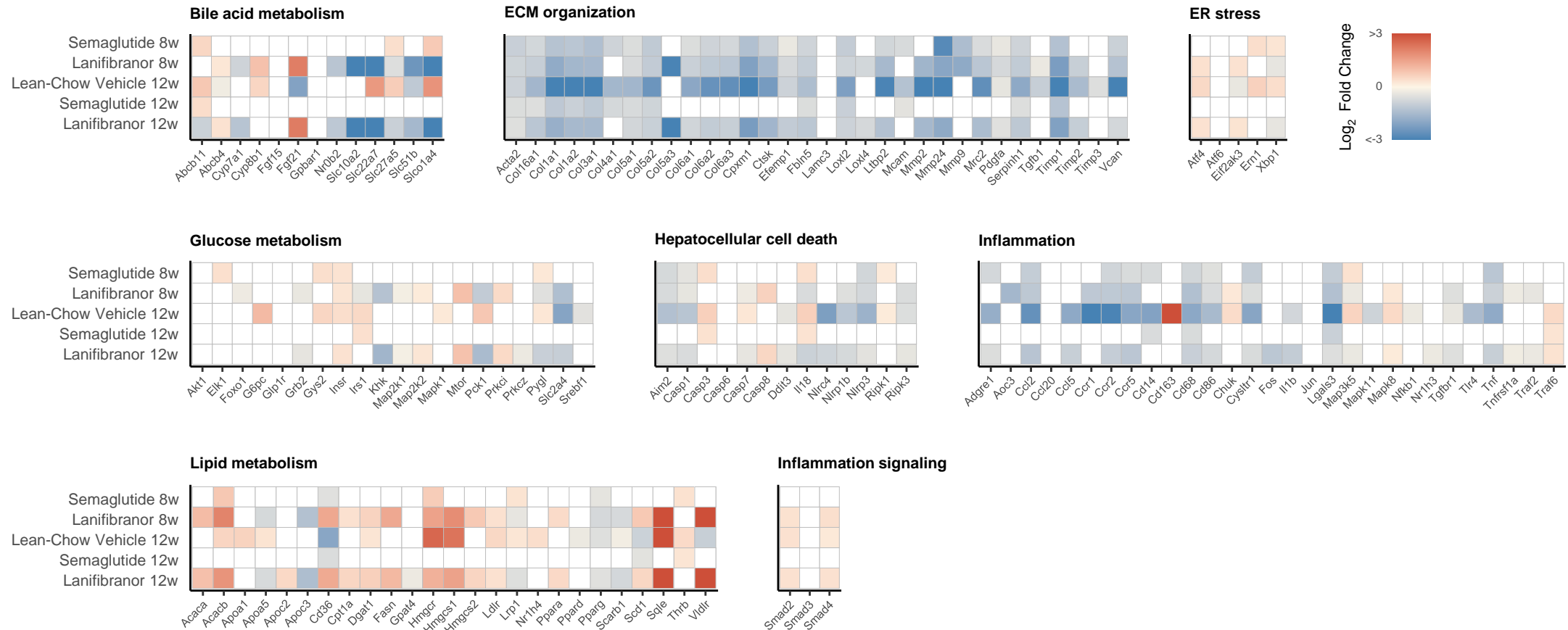
Heatmap of gene regulation in Gubra curated gene panels. Log2 fold change is shown for significantly regulated genes (padj < 0.05 after correcting for multiple testing). Red and blue colours indicate up- and down-regulation, respectively, when compared to the corresponding DIO-NASH Vehicle. Non-significantly regulated genes for a given comparison are shown in white. n = 8 - 13.

Example: RNA-seq data report. All data belong to Gubra.
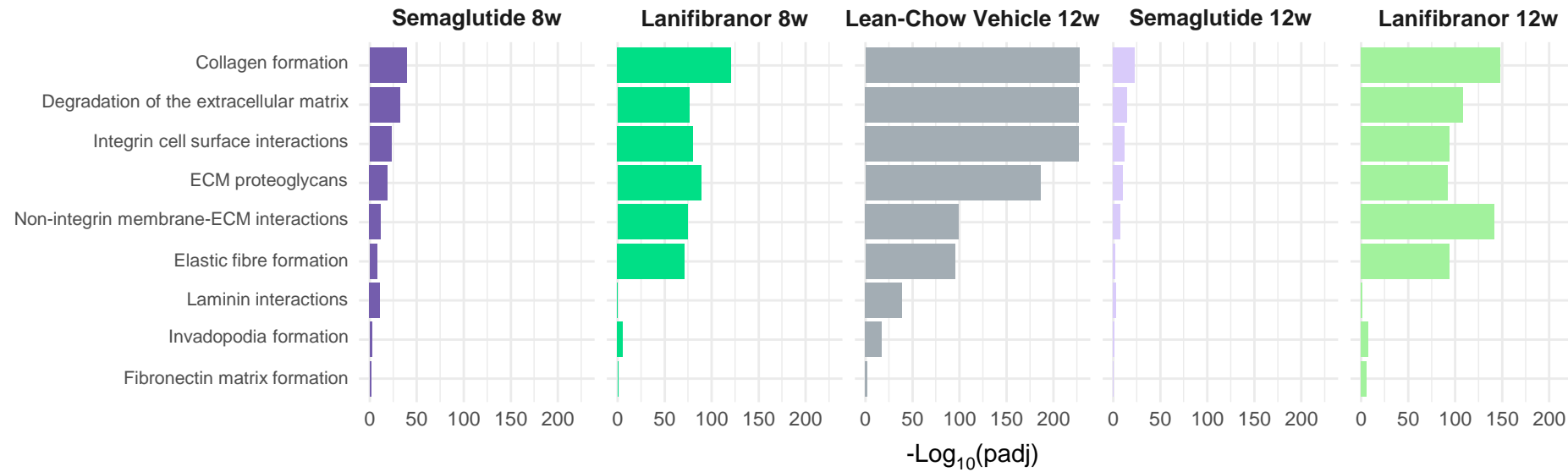
Unbiased pathway analysis

Top-level pathways
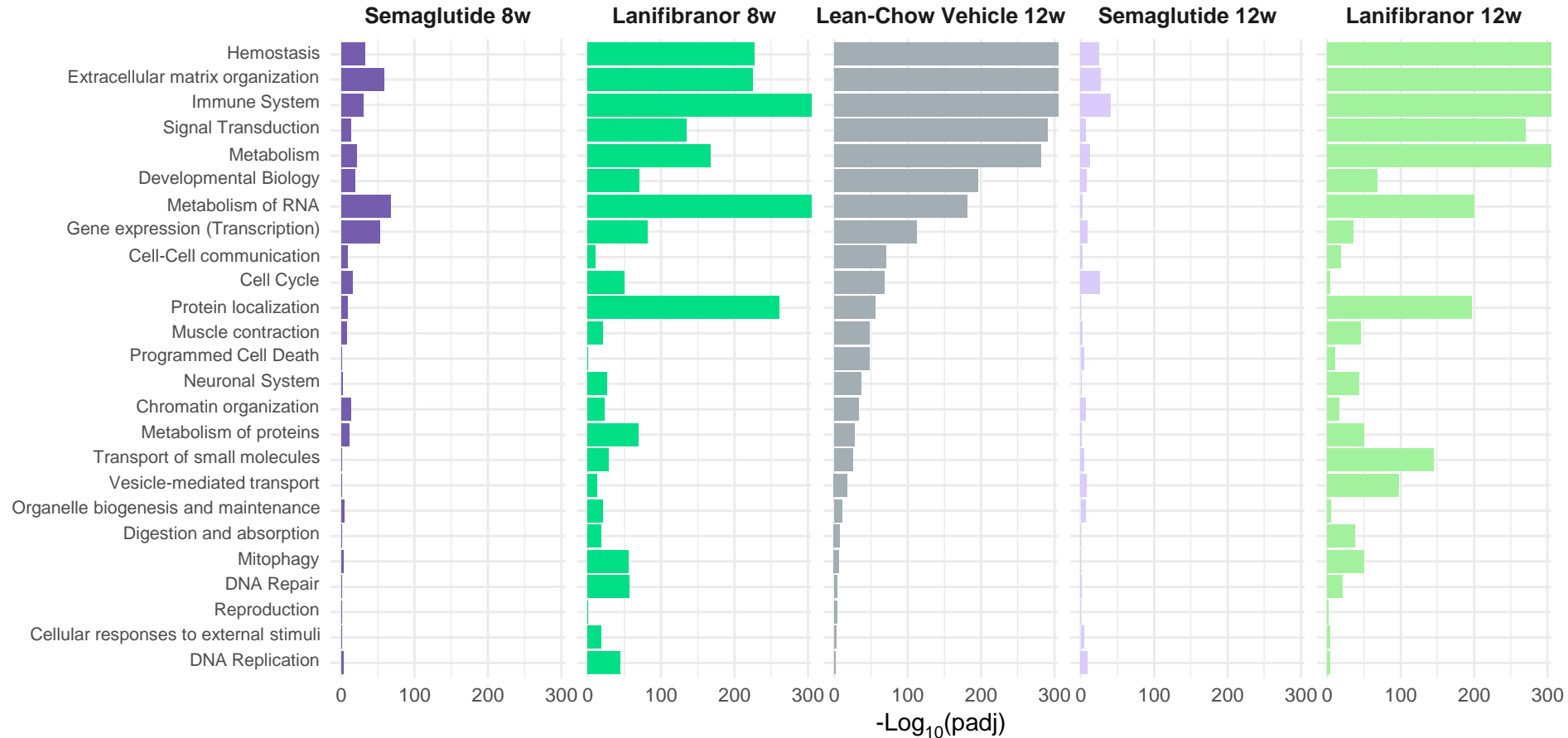
Enrichment of top-level pathways (Reactome database) for the shown study groups compared to the corresponding DIO-NASH Vehicle. The x-axis illustrates the level of enrichment depicted as the statistical significance for each gene-set analysis after correcting for multiple testing. n = 8 - 13.

Example: RNA-seq data report. All data belong to Gubra.

Unbiased pathway analysis

Extracellular matrix organization sub-pathways

Enrichment of *Extracellular matrix organization* sub-pathways (Reactome database, level 2) for the shown study groups compared to the corresponding DIO-NASH Vehicle. The x-axis illustrates the level of enrichment depicted as the statistical significance for each gene-set analysis after correcting for multiple testing. n = 8 - 13.

Example: RNA-seq data report. All data belong to Gubra.

Gene expression data in GubraView

We believe that the best way to fully exploit the potential of an RNA-seq dataset is to to enable our customers to explore and raise questions to the data. For this purpose, we have developed a panel of interactive functionalities in GubraView.

An instructional video showcasing how to use the interactive RNA-seq module in GubraView can be found here:
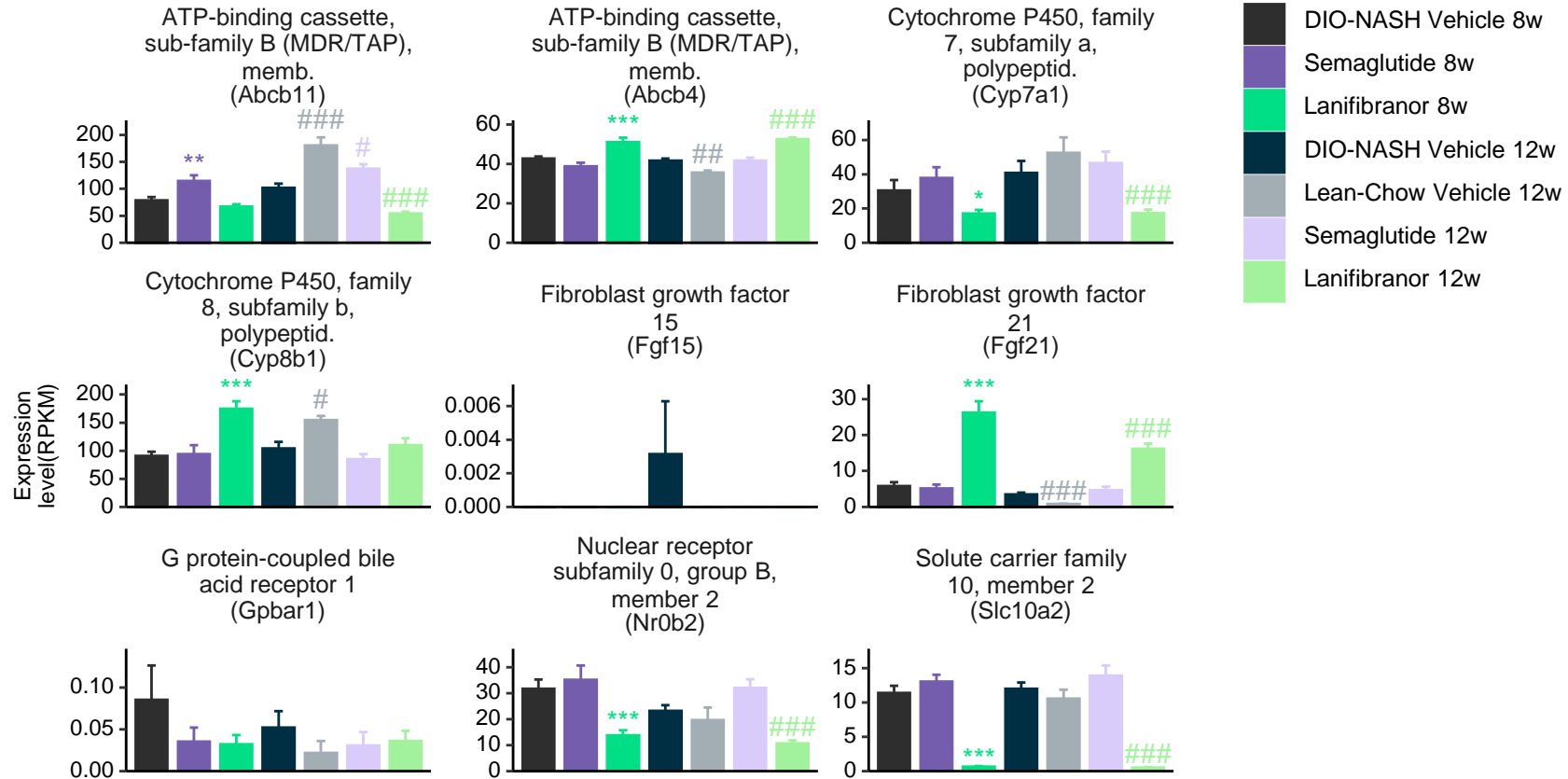
Video: RNA-seq in GubraView

9

Example: RNA-seq data report. All data belong to Gubra.

Appendix

Example: RNA-seq data report. All data belong to Gubra.
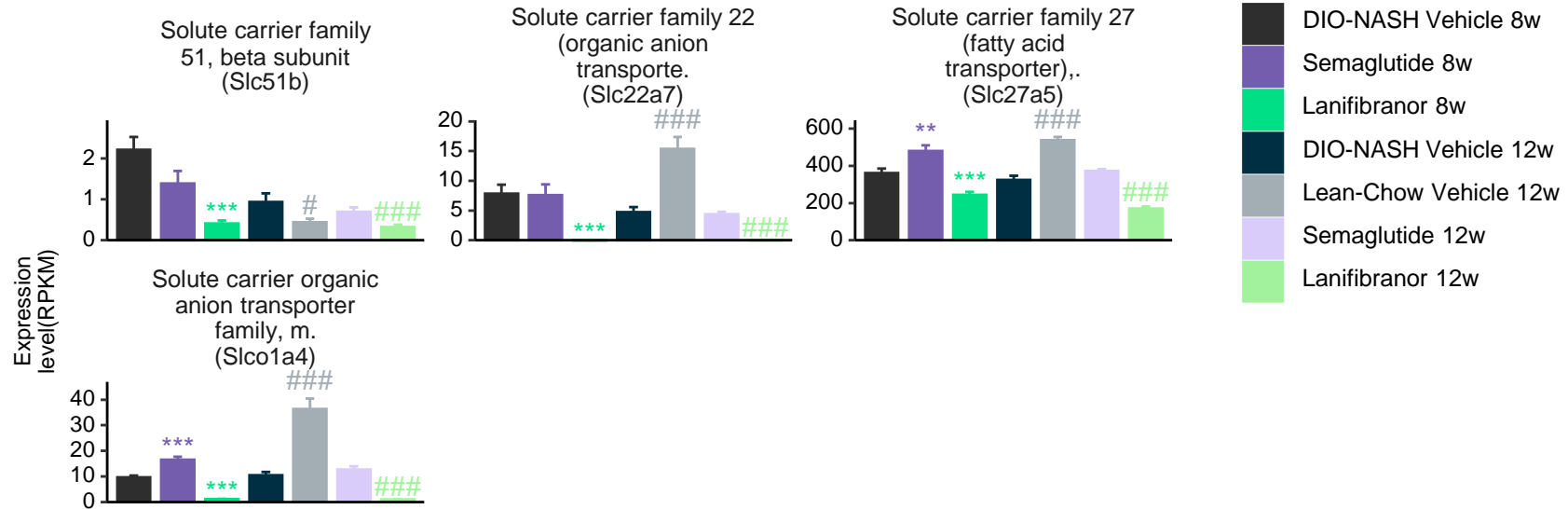
Bile acid metabolism

Bile acid metabolism - 1/2

Gene expression of genes found in Gubra curated gene panels. Values are expressed as group mean + SEM.
*/#:padj < 0.05, **/##:padj < 0.01, ***/###:padj < 0.001 compared to the corresponding DIO-NASH Vehicle
after correcting for multiple testing. RPKM: Reads Per Kilobase Million. n = 8 - 13.
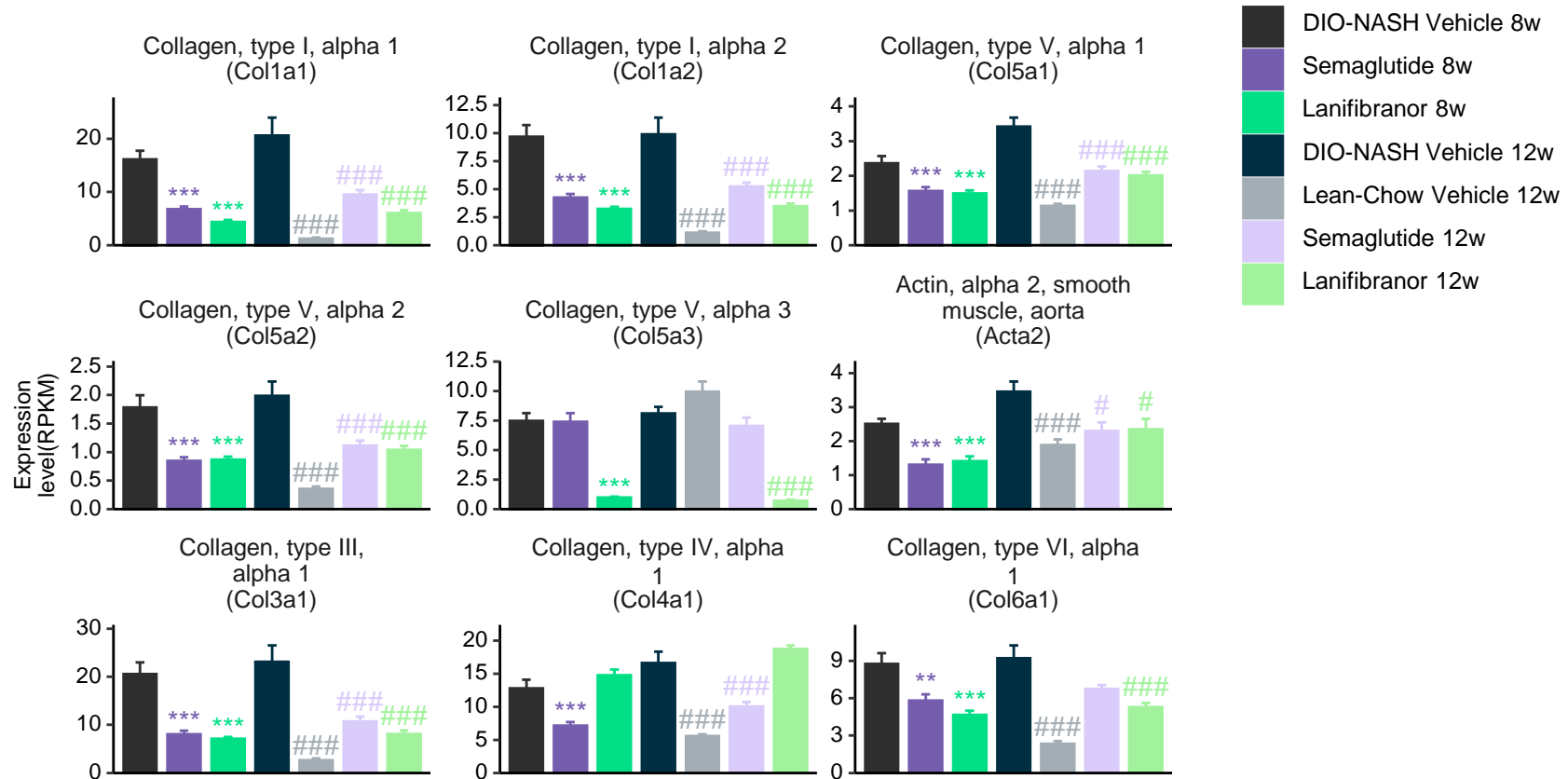
Bile acid metabolism

Bile acid metabolism - 2/2

Gene expression of genes found in Gubra curated gene panels. Values are expressed as group mean + SEM.
#:padj < 0.05, **:padj < 0.01, ***/###:padj < 0.001 compared to the corresponding DIO-NASH Vehicle after
correcting for multiple testing. RPKM: Reads Per Kilobase Million. n = 8 - 13.

Example: RNA-seq data report. All data belong to Gubra.
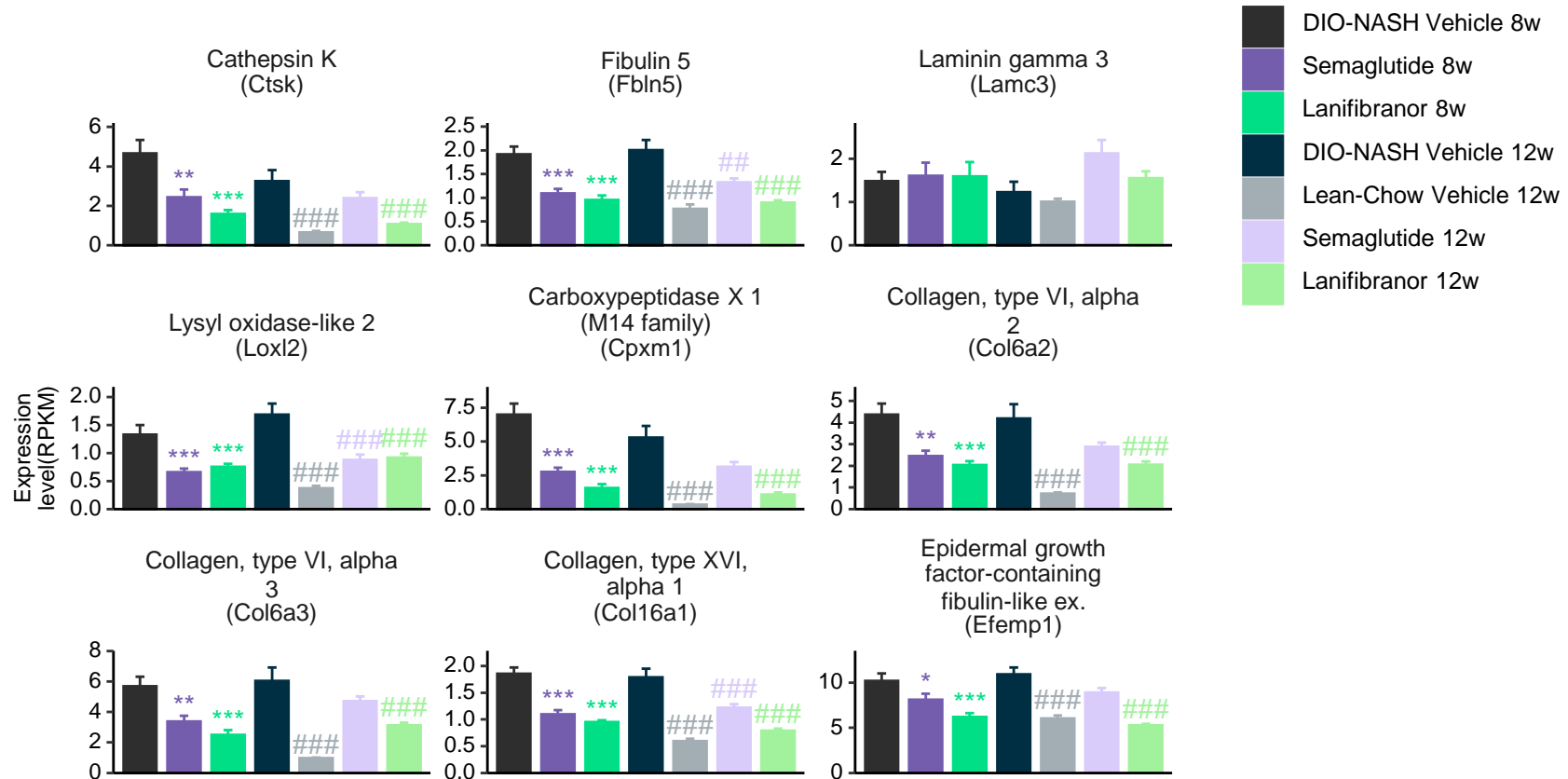
ECM organization

ECM organization - 1/4

Gene expression of genes found in Gubra curated gene panels. Values are expressed as group mean + SEM. #:padj < 0.05, **:padj < 0.01, ***/###:padj < 0.001 compared to the corresponding DIO-NASH Vehicle after correcting for multiple testing. RPKM: Reads Per Kilobase Million. n = 8 - 13.
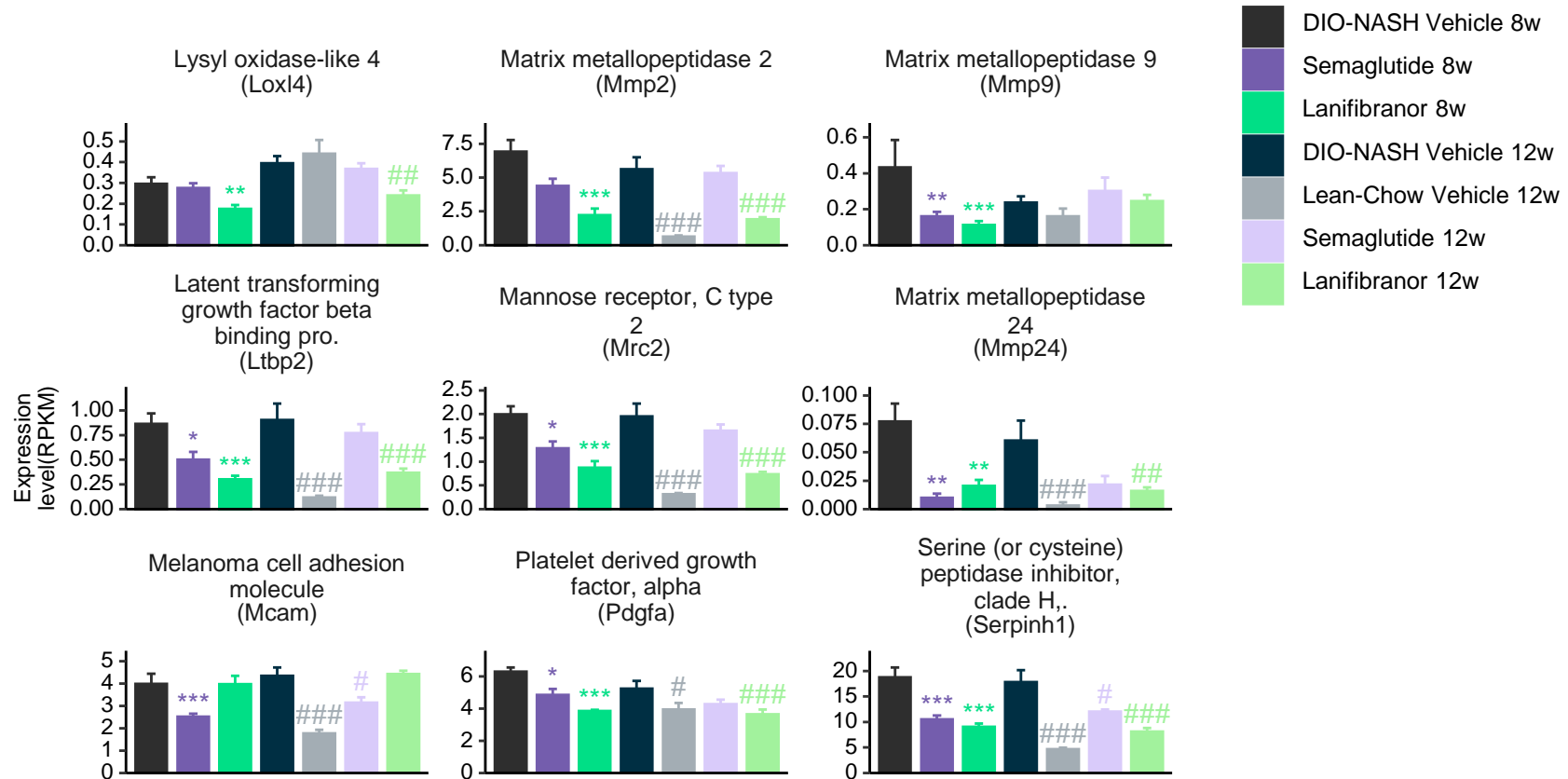
ECM organization

ECM organization - 2/4

Gene expression of genes found in Gubra curated gene panels. Values are expressed as group mean + SEM.
*:padj < 0.05, **/##:padj < 0.01, ***/###:padj < 0.001 compared to the corresponding DIO-NASH Vehicle
after correcting for multiple testing. RPKM: Reads Per Kilobase Million. n = 8 - 13.
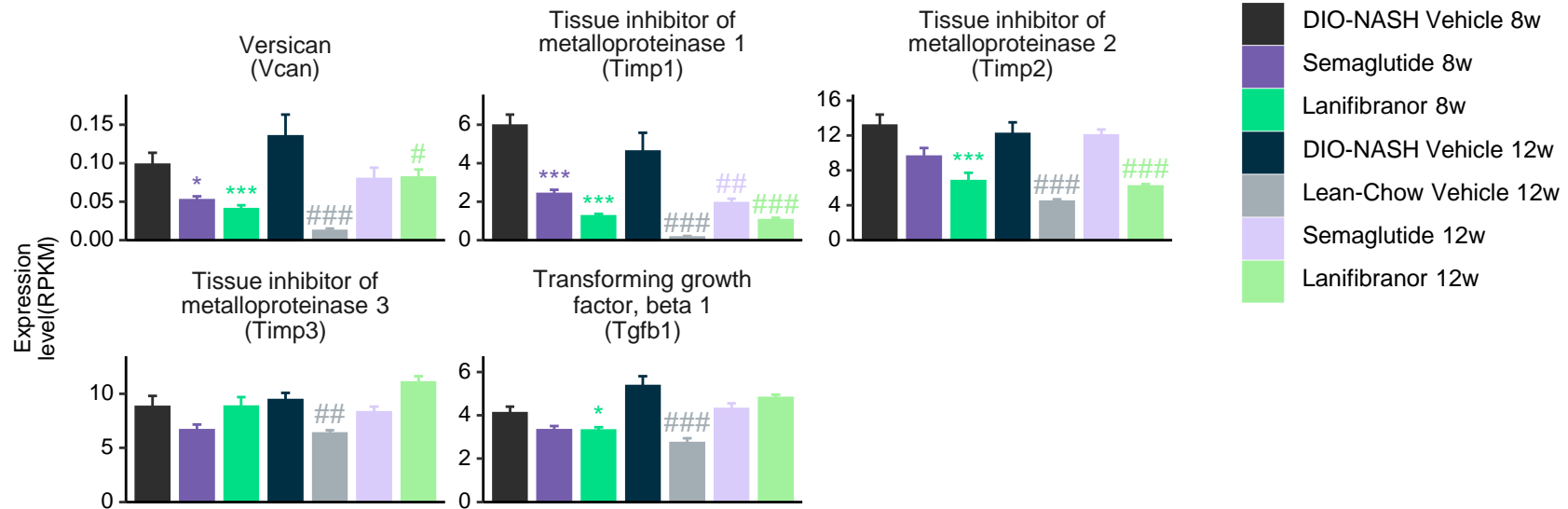
ECM organization

ECM organization - 3/4

Gene expression of genes found in Gubra curated gene panels. Values are expressed as group mean + SEM.
*/#:padj < 0.05, **/##:padj < 0.01, ***/###:padj < 0.001 compared to the corresponding DIO-NASH Vehicle
after correcting for multiple testing. RPKM: Reads Per Kilobase Million. n = 8 - 13.
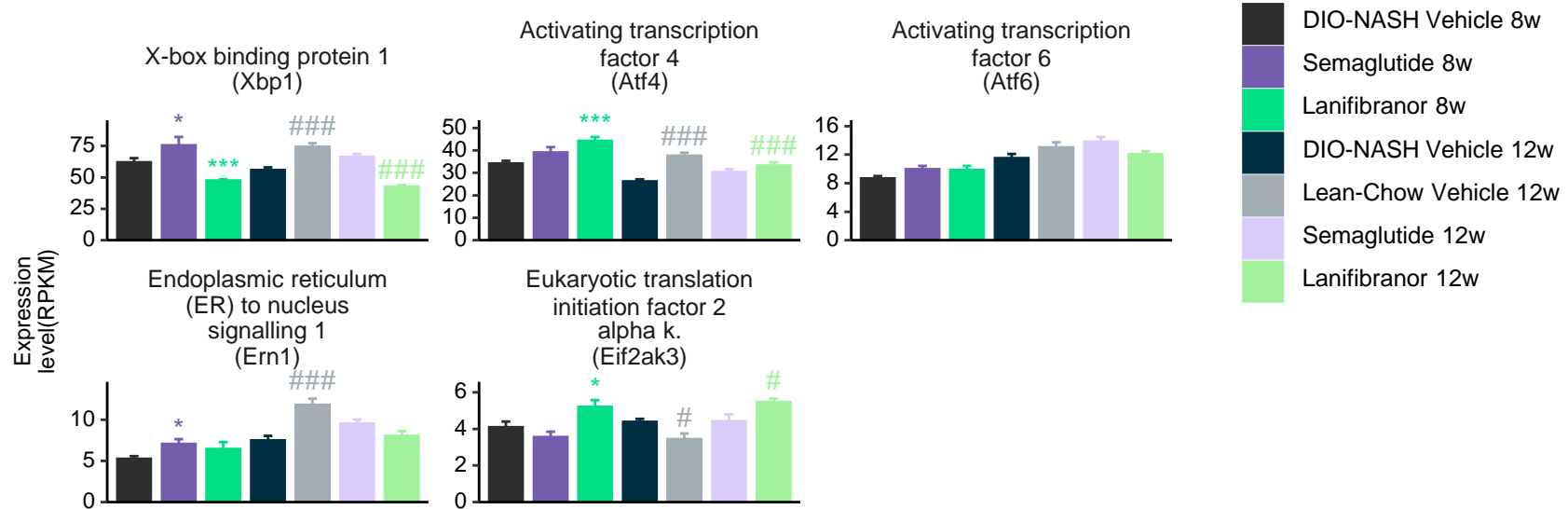
ECM organization

ECM organization - 4/4

Gene expression of genes found in Gubra curated gene panels. Values are expressed as group mean + SEM.
*/#:padj < 0.05, ##:padj < 0.01, ***/###:padj < 0.001 compared to the corresponding DIO-NASH Vehicle
after correcting for multiple testing. RPKM: Reads Per Kilobase Million. n = 8 - 13.
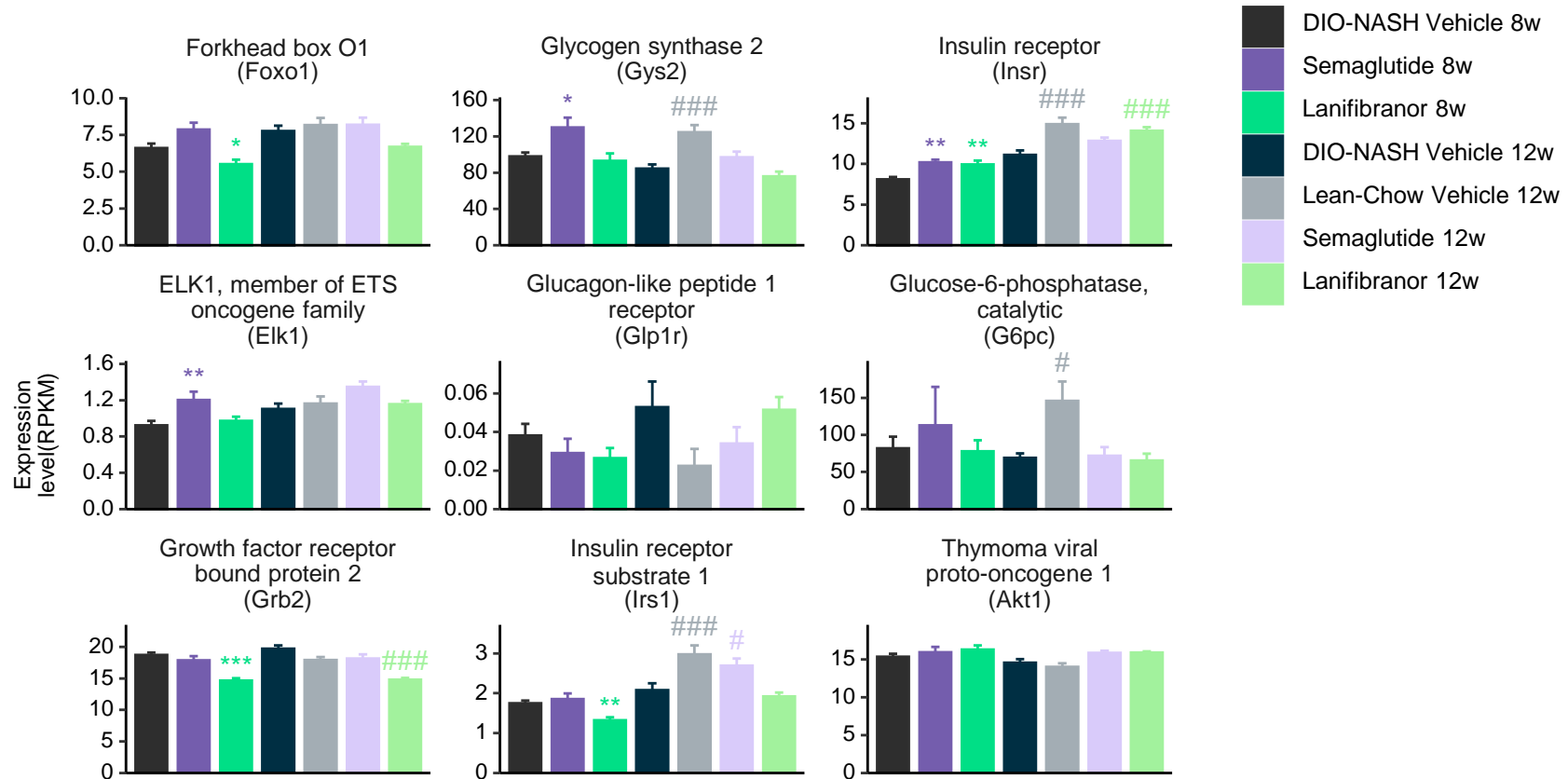
ER stress

ER stress

Gene expression of genes found in Gubra curated gene panels. Values are expressed as group mean + SEM. */#:padj < 0.05, ***/###:padj < 0.001 compared to the corresponding DIO-NASH Vehicle after correcting for multiple testing. RPKM: Reads Per Kilobase Million. n = 8 - 13.

Example: RNA-seq data report. All data belong to Gubra.
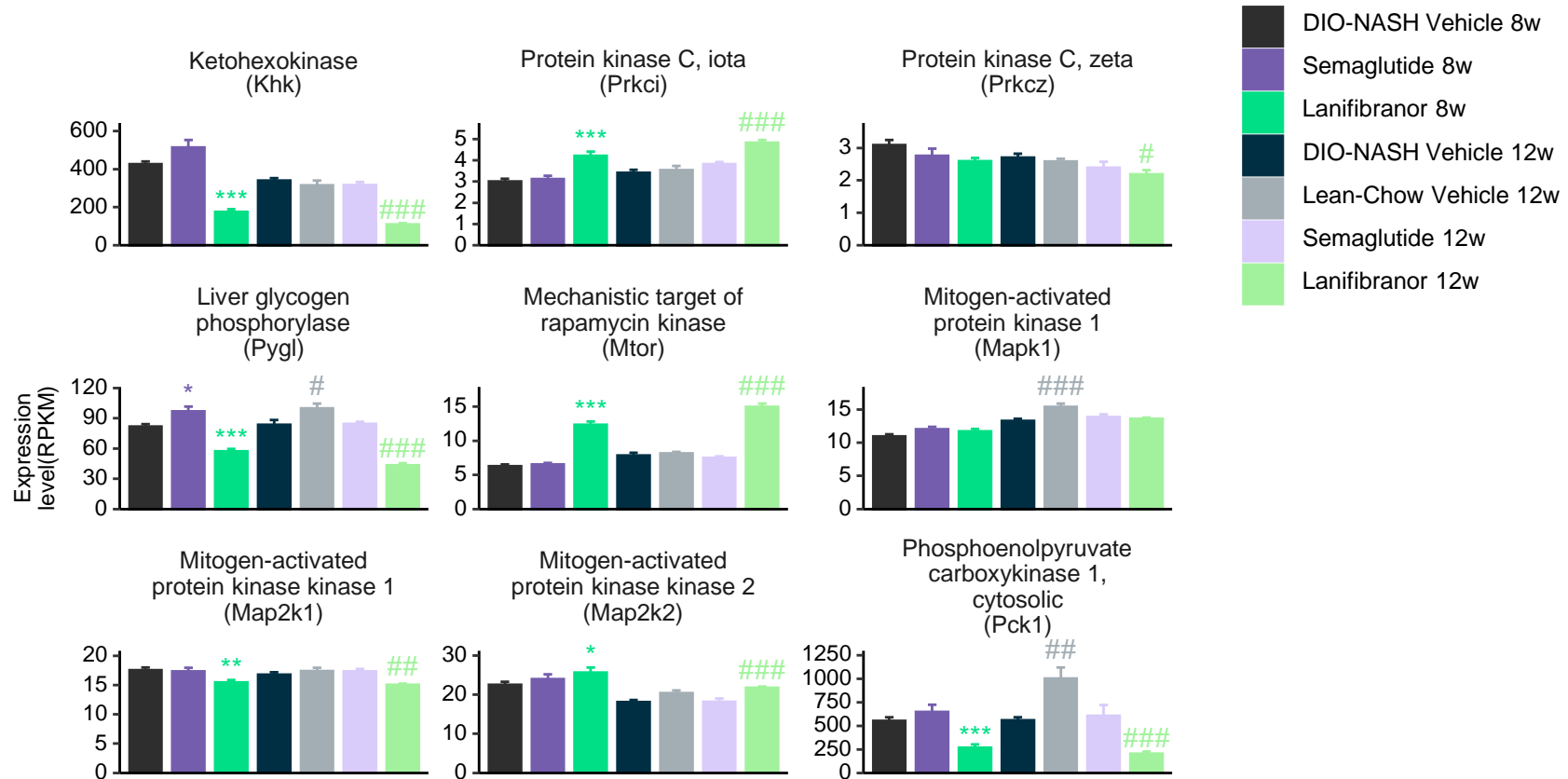
Glucose metabolism

Glucose metabolism - 1/3

Gene expression of genes found in Gubra curated gene panels. Values are expressed as group mean + SEM. */#:padj < 0.05, **:padj < 0.01, ***/###:padj < 0.001 compared to the corresponding DIO-NASH Vehicle after correcting for multiple testing. RPKM: Reads Per Kilobase Million. n = 8 - 13.

Example: RNA-seq data report. All data belong to Gubra.
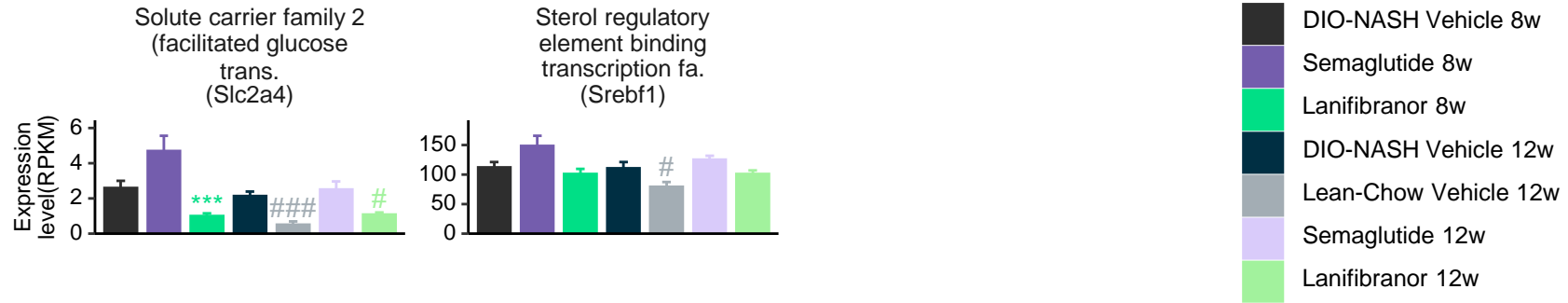
Glucose metabolism

Glucose metabolism - 2/3

Gene expression of genes found in Gubra curated gene panels. Values are expressed as group mean + SEM.
*/#:padj < 0.05, **/##:padj < 0.01, ***/###:padj < 0.001 compared to the corresponding DIO-NASH Vehicle
after correcting for multiple testing. RPKM: Reads Per Kilobase Million. n = 8 - 13.

Example: RNA-seq data report. All data belong to Gubra.
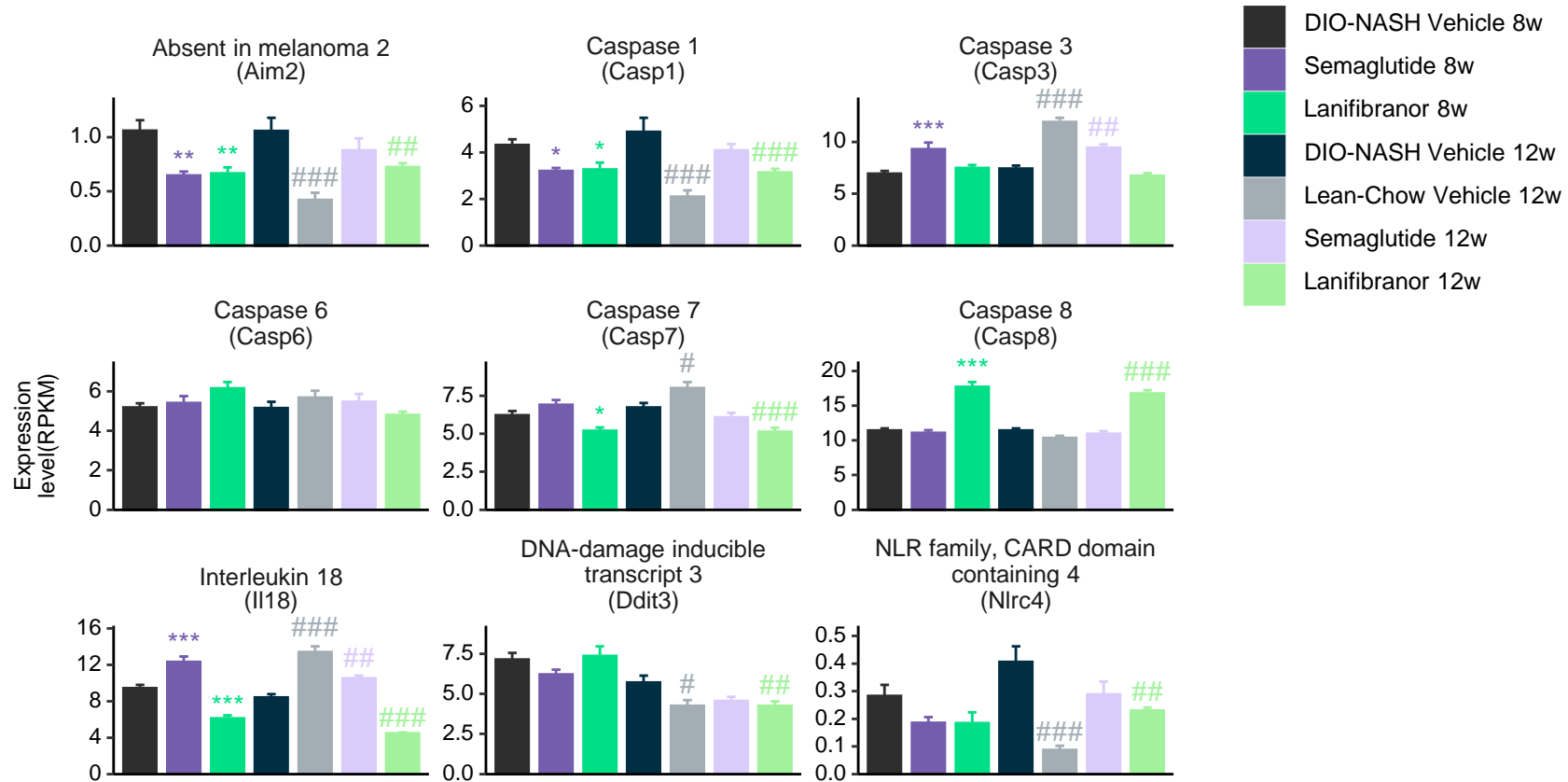
Glucose metabolism

Glucose metabolism - 3/3

Gene expression of genes found in Gubra curated gene panels. Values are expressed as group mean + SEM. #:padj < 0.05, ***/###:padj < 0.001 compared to the corresponding DIO-NASH Vehicle after correcting for multiple testing. RPKM: Reads Per Kilobase Million. n = 8 - 13.
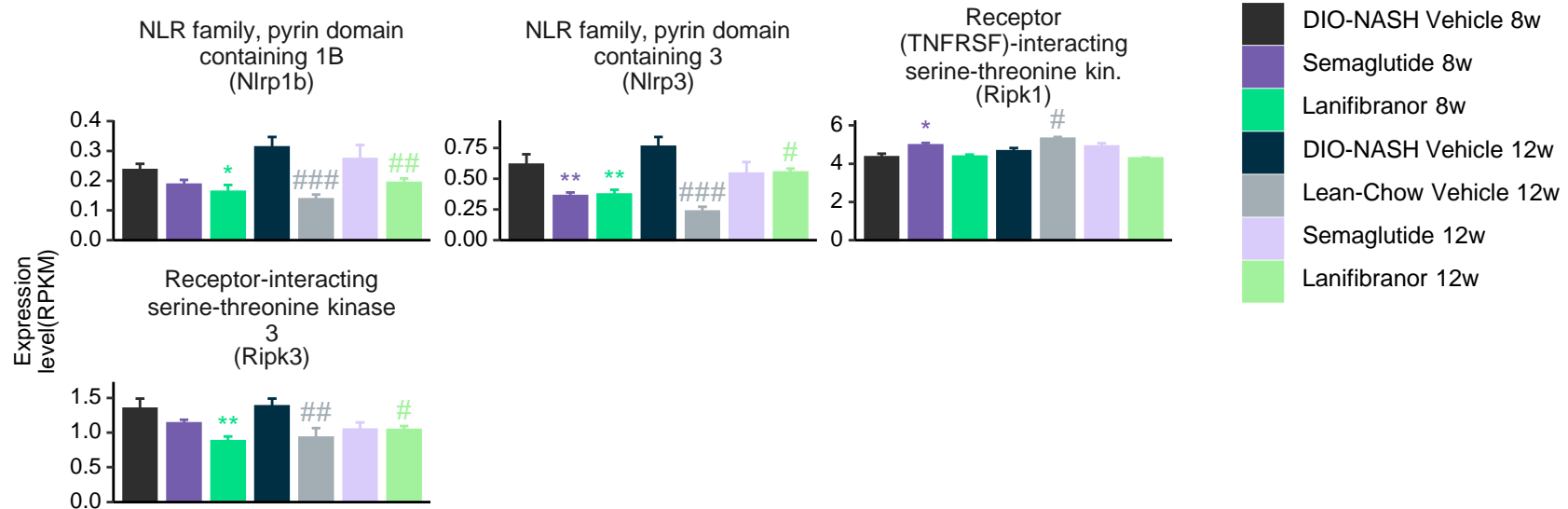
Hepatocellular cell death

Hepatocellular cell death - 1/2

Gene expression of genes found in Gubra curated gene panels. Values are expressed as group mean + SEM.
*/#:padj < 0.05, **/##:padj < 0.01, ***/###:padj < 0.001 compared to the corresponding DIO-NASH Vehicle
after correcting for multiple testing. RPKM: Reads Per Kilobase Million. n = 8 - 13.
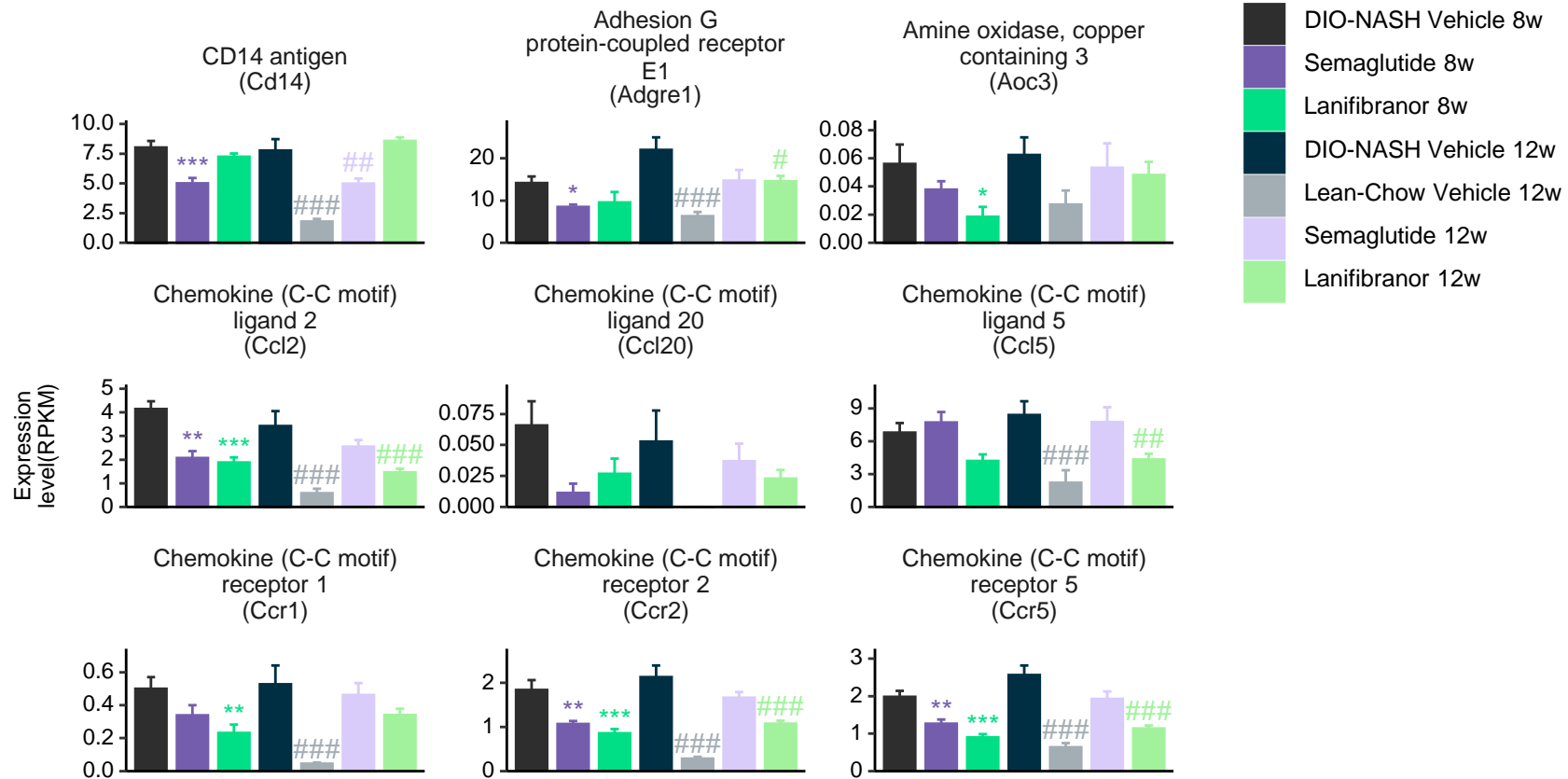
Hepatocellular cell death

Hepatocellular cell death - 2/2

Gene expression of genes found in Gubra curated gene panels. Values are expressed as group mean + SEM. */#:padj < 0.05, **/##:padj < 0.01, ###:padj < 0.001 compared to the corresponding DIO-NASH Vehicle after correcting for multiple testing. RPKM: Reads Per Kilobase Million. n = 8 - 13.
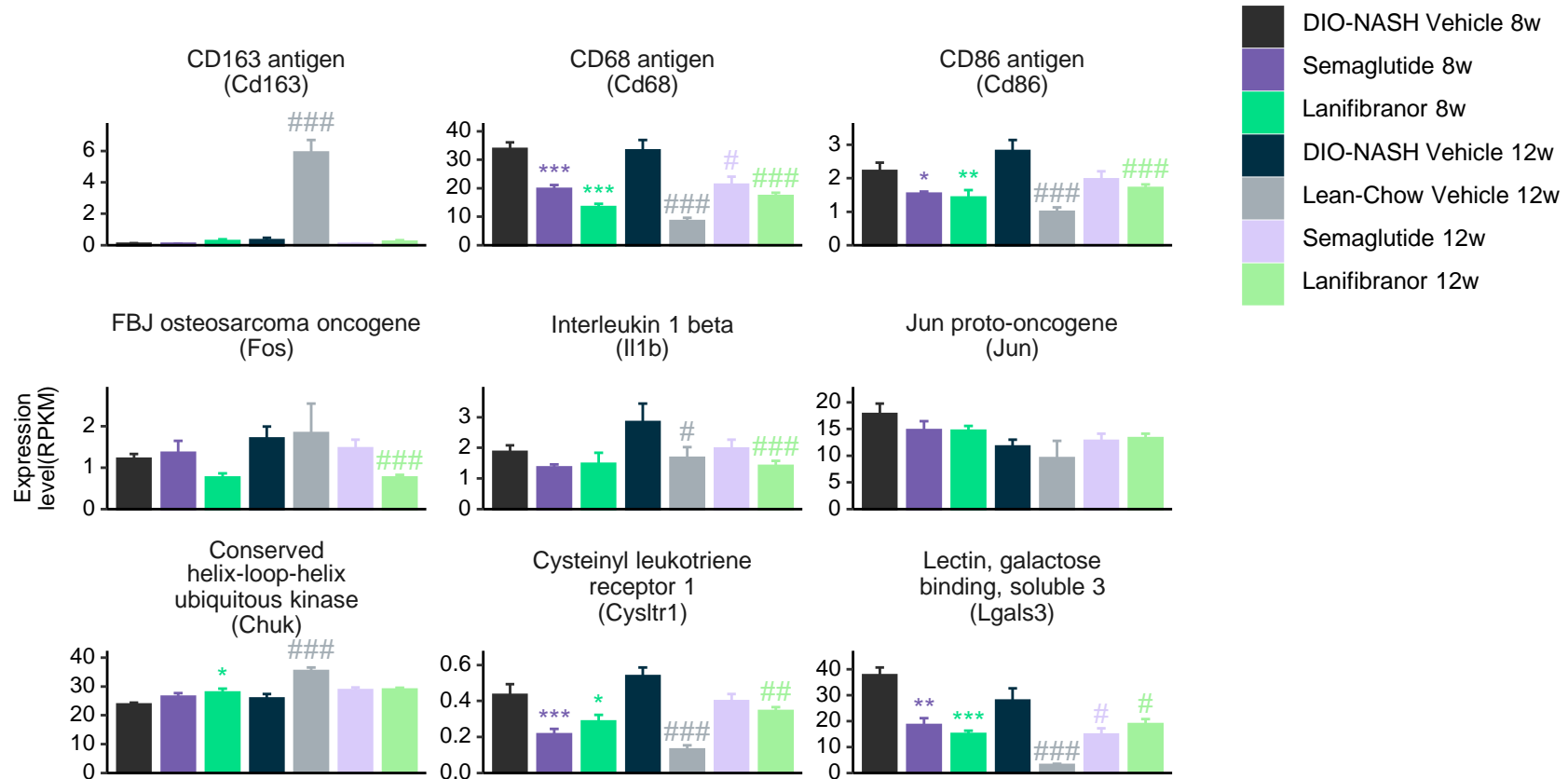
Inflammation

Inflammation - 1/4

Gene expression of genes found in Gubra curated gene panels. Values are expressed as group mean + SEM.
*/#:padj < 0.05, **/##:padj < 0.01, ***/###:padj < 0.001 compared to the corresponding DIO-NASH Vehicle
after correcting for multiple testing. RPKM: Reads Per Kilobase Million. n = 8 - 13.
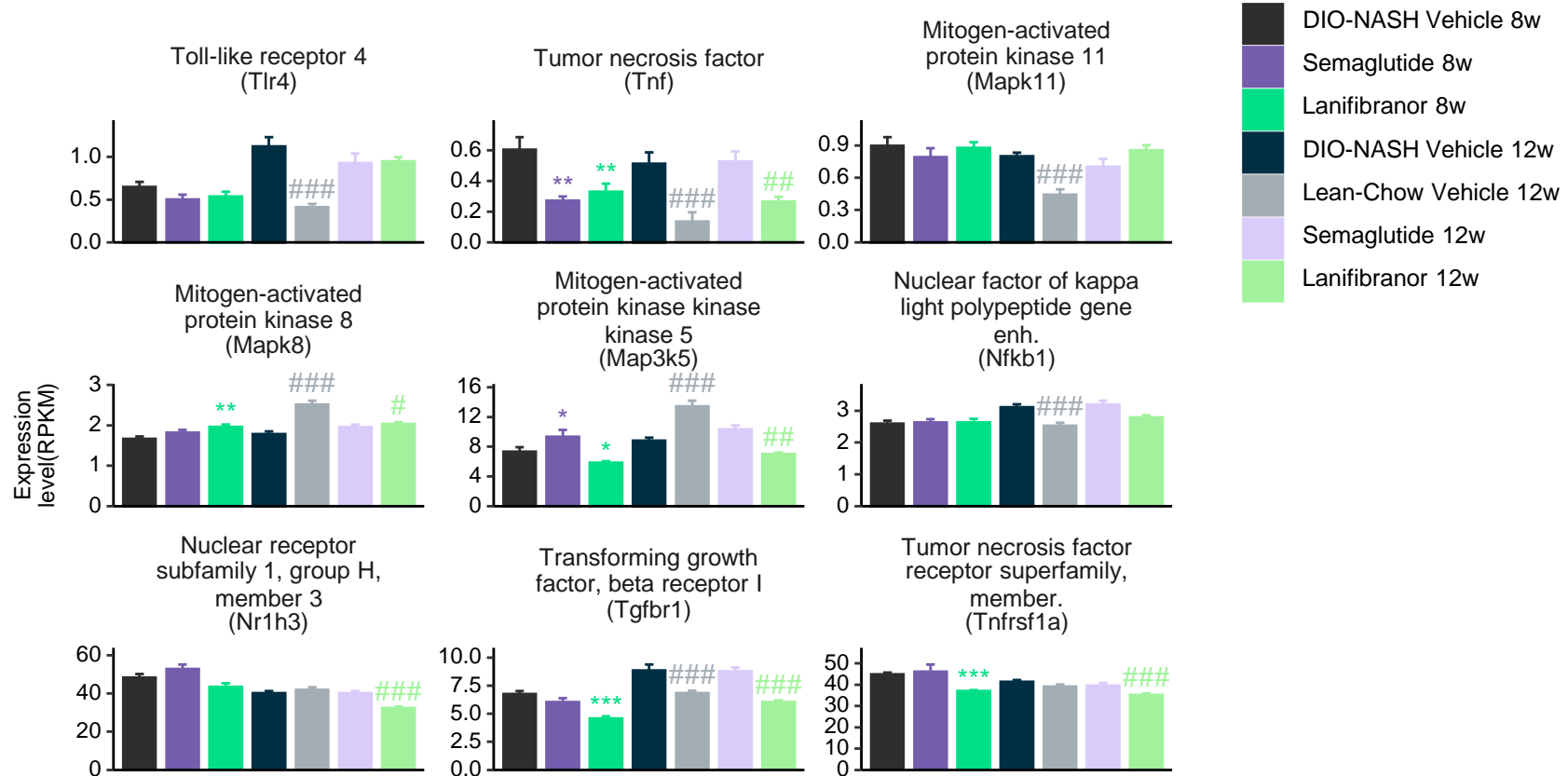
Inflammation

Inflammation - 2/4

Gene expression of genes found in Gubra curated gene panels. Values are expressed as group mean + SEM.
*/#:padj < 0.05, **/##:padj < 0.01, ***/###:padj < 0.001 compared to the corresponding DIO-NASH Vehicle
after correcting for multiple testing. RPKM: Reads Per Kilobase Million. n = 8 - 13.
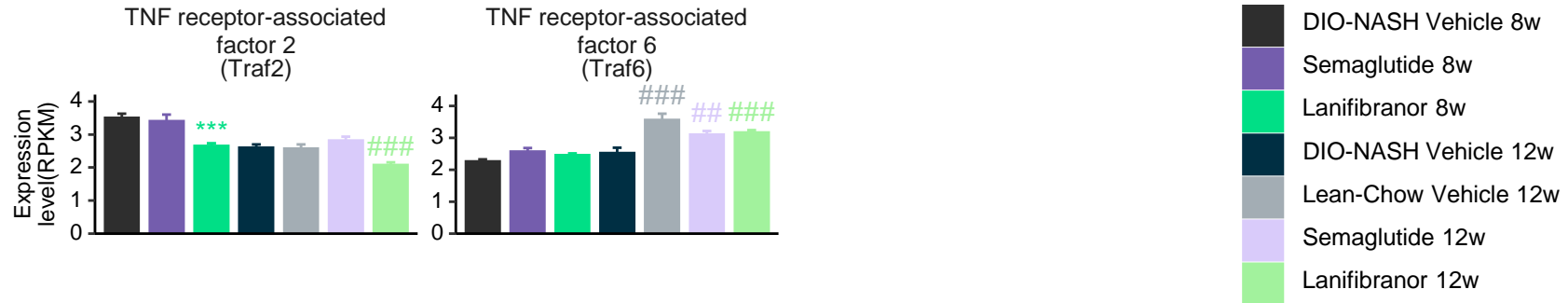
Inflammation

Inflammation - 3/4

Gene expression of genes found in Gubra curated gene panels. Values are expressed as group mean + SEM.
*/#:padj < 0.05, **/##:padj < 0.01, ***/###:padj < 0.001 compared to the corresponding DIO-NASH Vehicle after correcting for multiple testing. RPKM: Reads Per Kilobase Million. n = 8 - 13.
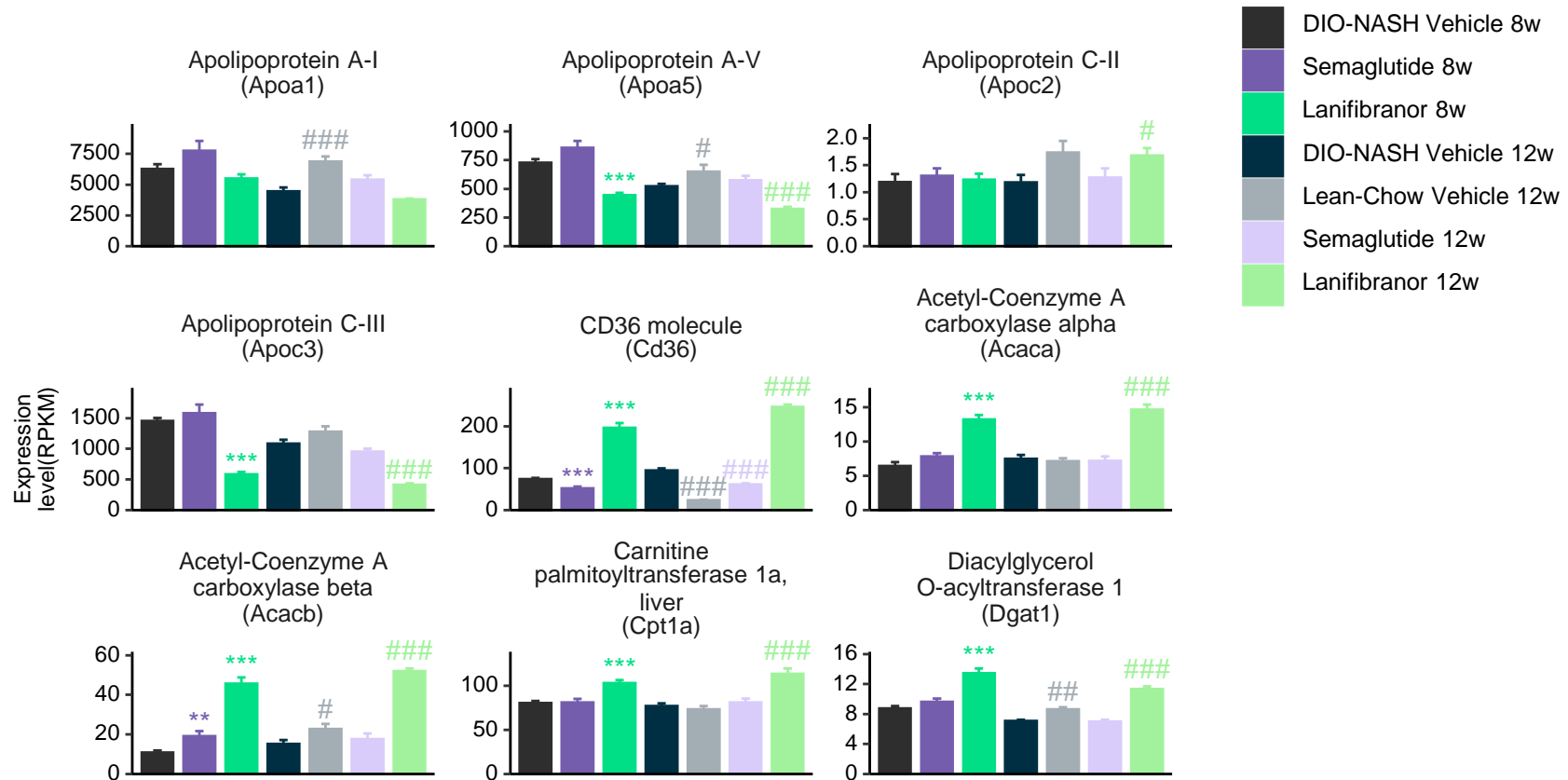
Inflammation

Inflammation - 4/4

Gene expression of genes found in Gubra curated gene panels. Values are expressed as group mean + SEM. ##:padj < 0.01, ***/###:padj < 0.001 compared to the corresponding DIO-NASH Vehicle after correcting for multiple testing. RPKM: Reads Per Kilobase Million. n = 8 - 13.
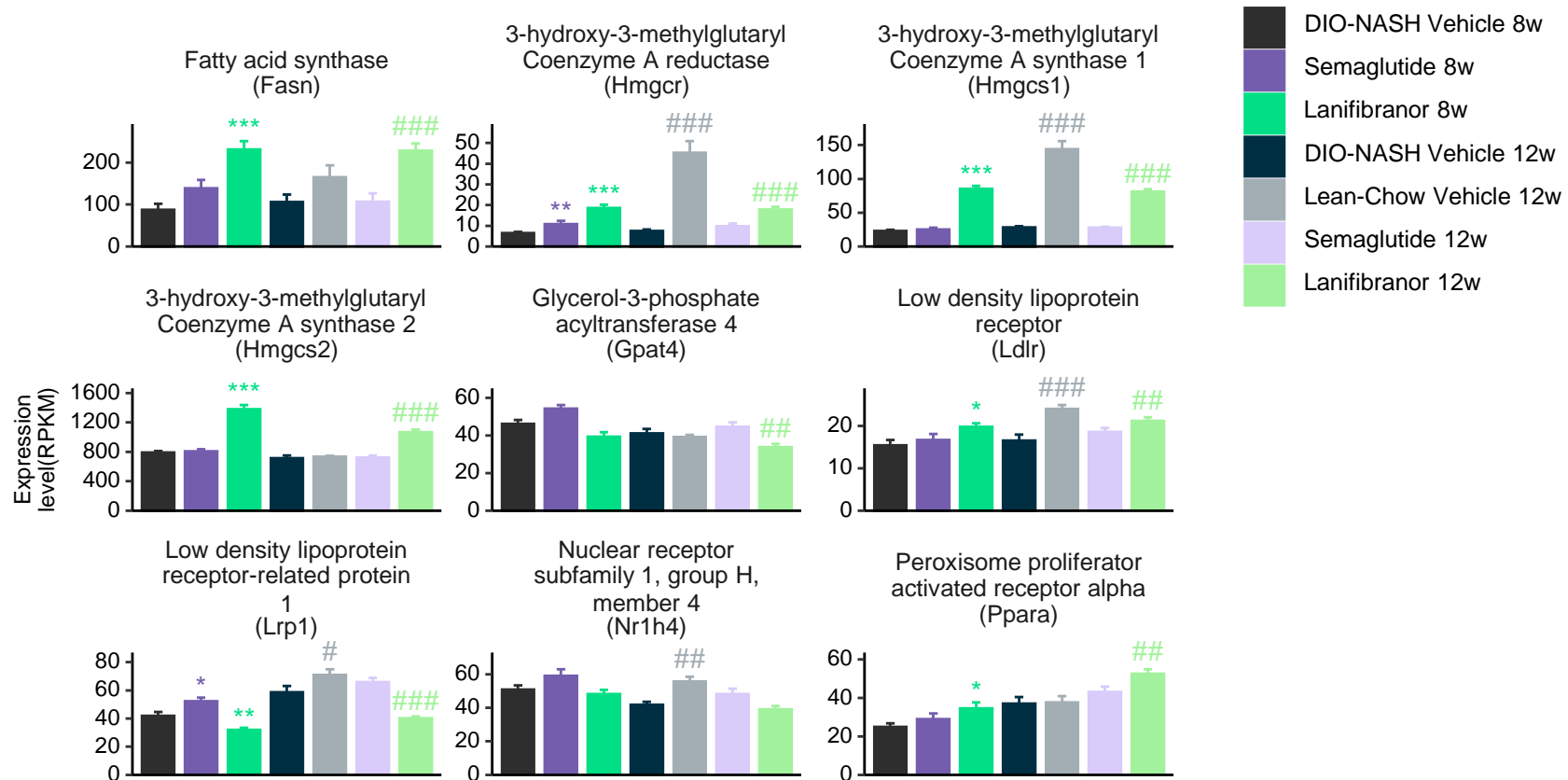
Lipid metabolism

Lipid metabolism - 1/3

Gene expression of genes found in Gubra curated gene panels. Values are expressed as group mean + SEM.
#:padj < 0.05, **/##:padj < 0.01, ***/###:padj < 0.001 compared to the corresponding DIO-NASH Vehicle
after correcting for multiple testing. RPKM: Reads Per Kilobase Million. n = 8 - 13.
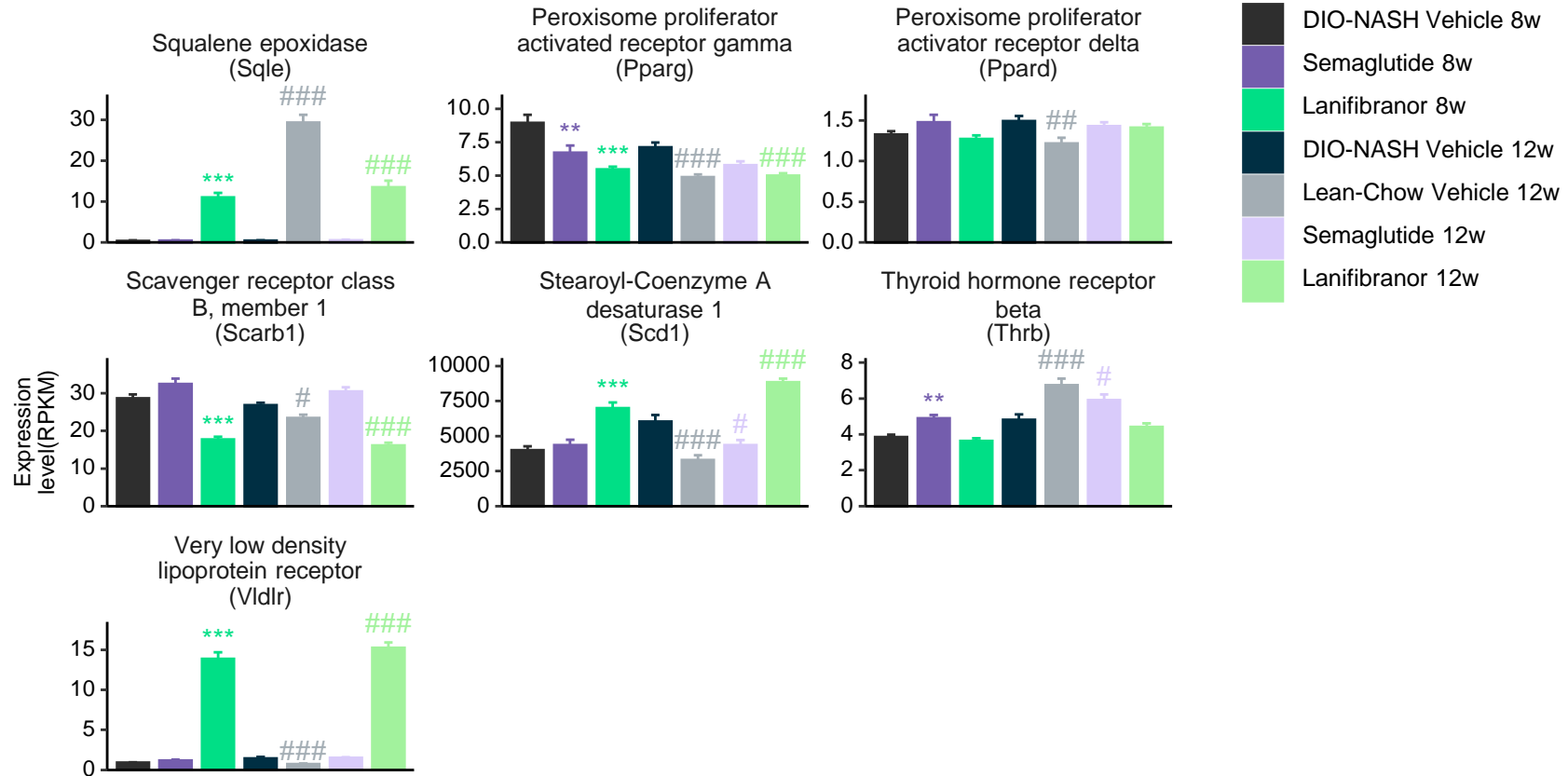
Lipid metabolism

Lipid metabolism - 2/3

Gene expression of genes found in Gubra curated gene panels. Values are expressed as group mean + SEM.
*/#:padj < 0.05, **/##:padj < 0.01, ***/###:padj < 0.001 compared to the corresponding DIO-NASH Vehicle
after correcting for multiple testing. RPKM: Reads Per Kilobase Million. n = 8 - 13.
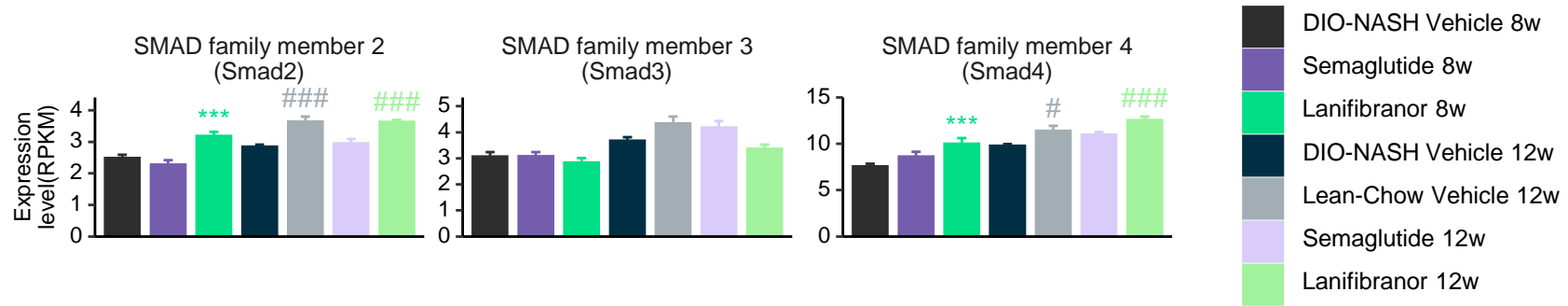
Lipid metabolism - 3/3

Gene expression of genes found in Gubra curated gene panels. Values are expressed as group mean + SEM.
#:padj < 0.05, **/##:padj < 0.01, ***/###:padj < 0.001 compared to the corresponding DIO-NASH Vehicle
after correcting for multiple testing. RPKM: Reads Per Kilobase Million. n = 8 - 13.

Inflammation signaling

Inflammation signaling

Gene expression of genes found in Gubra curated gene panels. Values are expressed as group mean + SEM.
#:padj < 0.05, ***/###:padj < 0.001 compared to the corresponding DIO-NASH Vehicle after correcting for
multiple testing. RPKM: Reads Per Kilobase Million. n = 8 - 13.

Example: RNA-seq data report. All data belong to Gubra.