

GENE EXPRESSION PREDICTS HISTOLOGICAL SEVERITY AND REVEALS DISTINCT MOLECULAR PROFILES OF NONALCOHOLIC FATTY LIVER DISEASE

Stephen A. Hoang PhD¹, Abdul Oseini MD², Ryan E. Feaver PhD¹, Banumathi Cole PhD¹, Amon Asgharpour MD³, Robert Vincent MD PhD², Mohammad Siddiqui MD², Mark J. Lawson PhD¹, Nathan C. Day¹, Justin M. Taylor¹, Brian R. Wamhoff PhD¹, Faridoddin Mirshahi MS², Melissa J Contos MD², Michael Idowu MD² and Arun J. Sanyal MBBS MD²

Corresponding authors:

Stephen A. Hoang PhD
Title: Head of Computational Biology, HemoShear Therapeutics
Address: 501 Locust Avenue
Charlottesville, VA 22902
Phone: (434) 872 0196
Fax:
Email: hoang@hemoshear.com

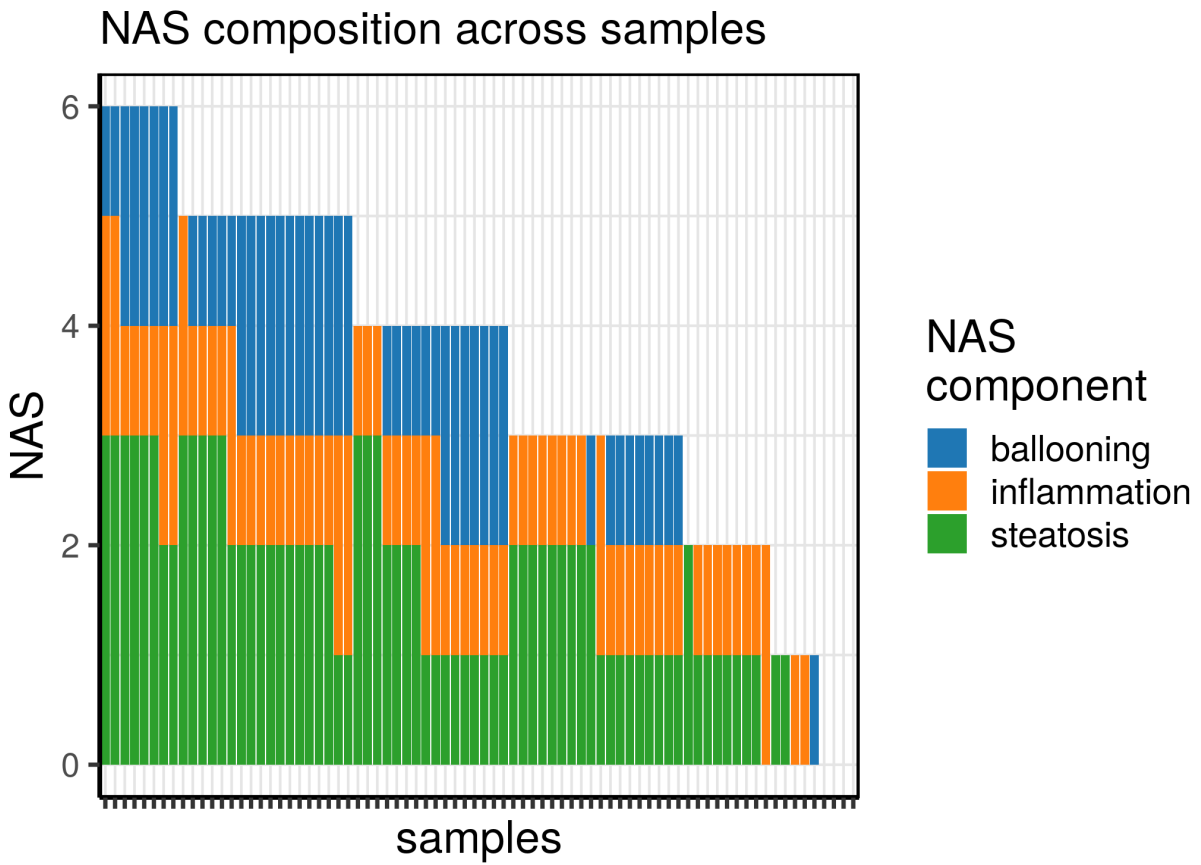
Senior author:

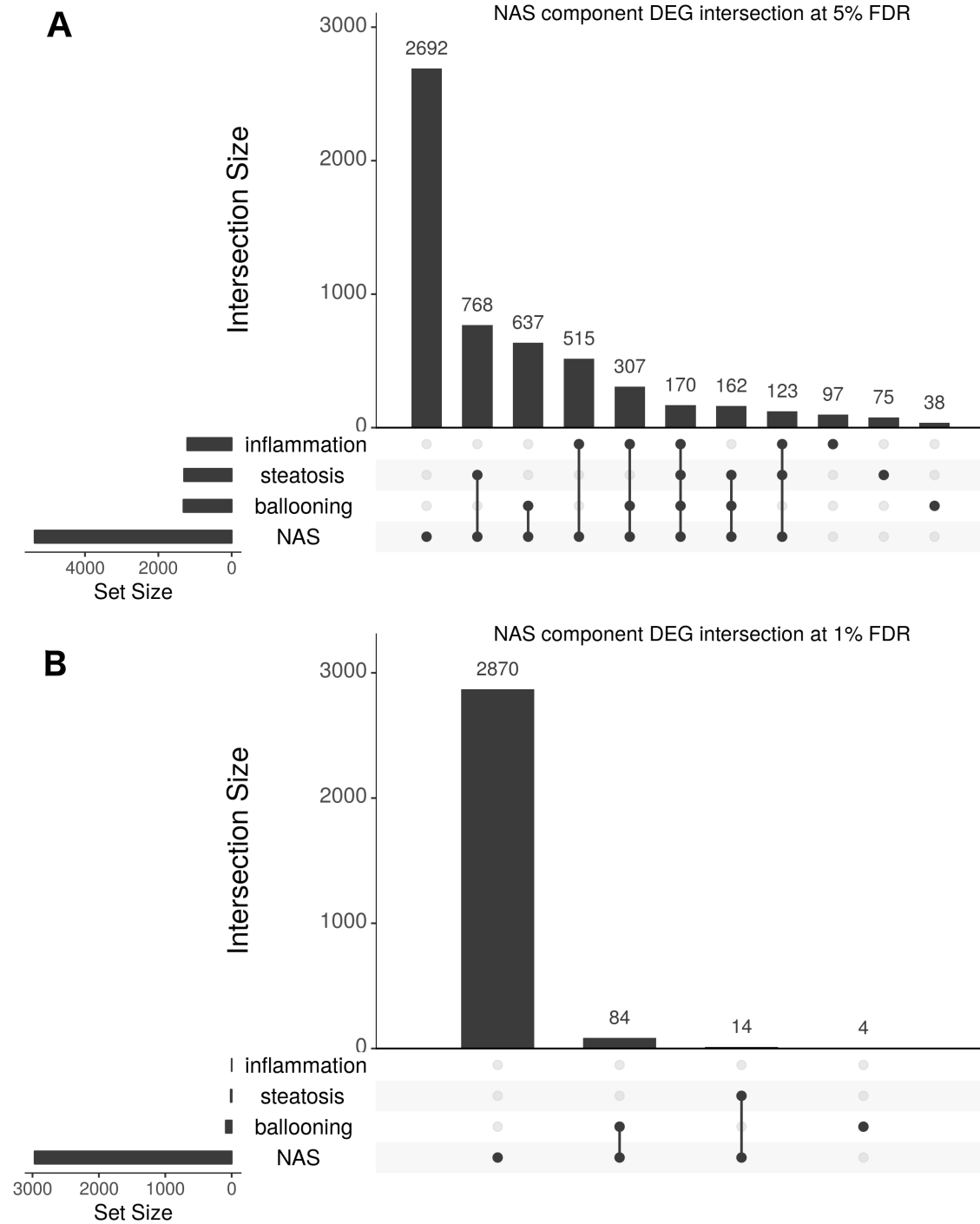
Arun J. Sanyal
Title: Z. Reno Vlahcevic Professor of Medicine, Physiology and Molecular Pathology
Address: MCV Box 980341
Richmond, VA 23298-0341
Phone: (804) 828 6314
Fax: (804) 828 2992
Email: arun.sanyal@vcuhealth.org

Affiliations:

1. HemoShear Therapeutics, LLC, Charlottesville, Virginia
2. Div. of Gastroenterology, Hepatology and Nutrition, Dept. of Internal Medicine, Virginia Commonwealth University School of Medicine, Richmond, VA
3. Div. of Surgical Pathology, Dept. of Pathology, Virginia Commonwealth University School of Medicine, Richmond, VA

Supplementary figure 1

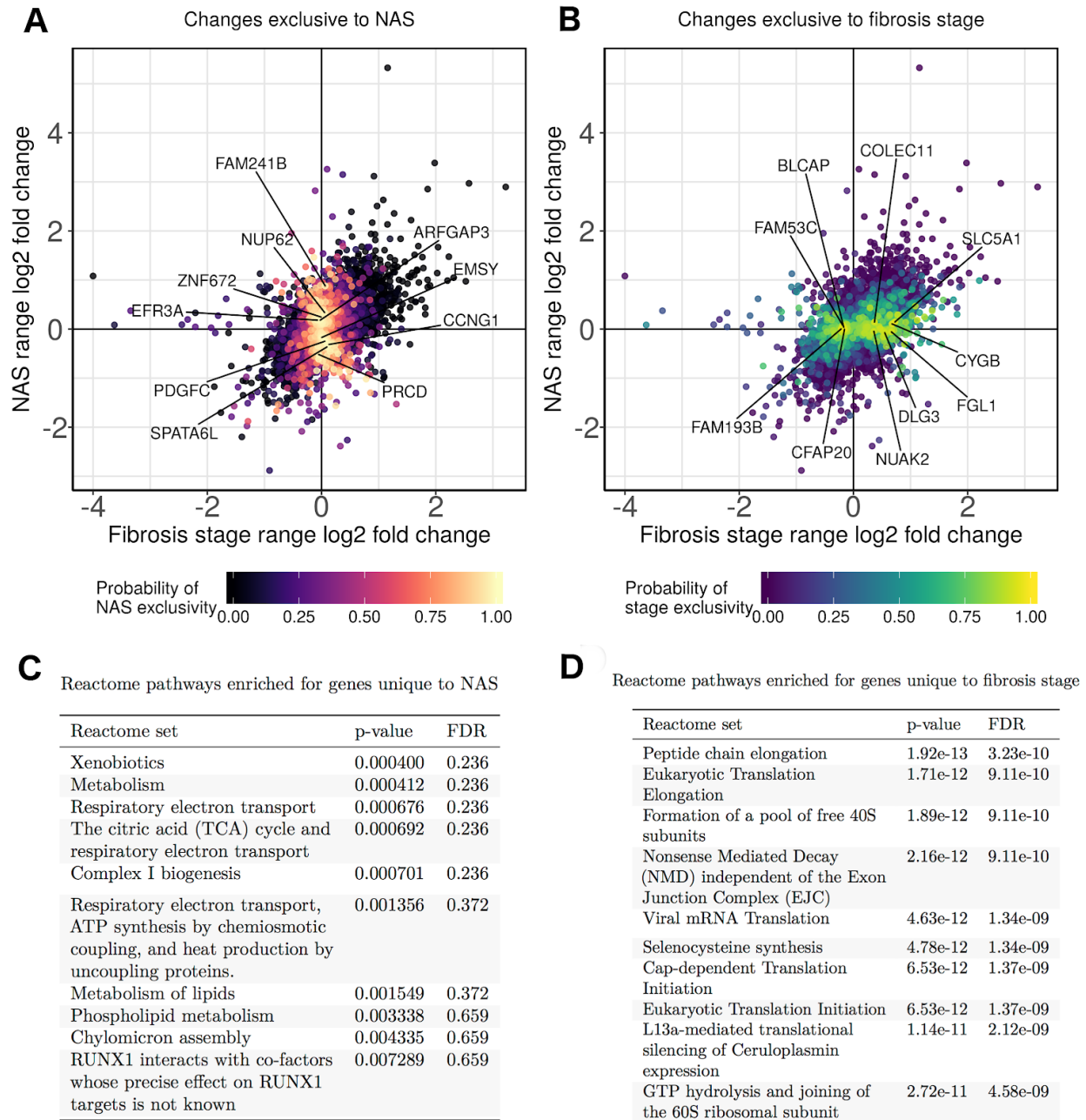


Supplementary figure 2

Supplementary Figure 2: Differential expression analysis of NAS components. Plot shows the set intersection cardinality for differentially expressed genes at 5% (A) and 1% (B) FDR thresholds. Differentially expressed genes were assessed with respect to lobular inflammation, steatosis, cytological ballooning, and the composite NAS. Individual set sizes are shown on the

lower left corner of the plots, intersection sizes are on the y-axis, and intersection labels are shown along the x-axis.

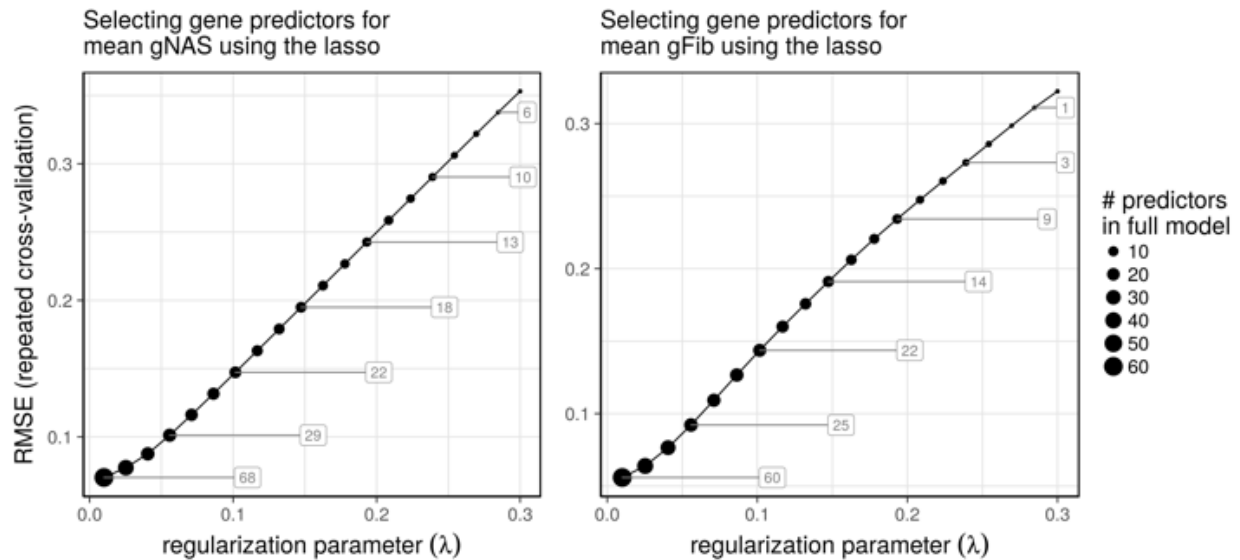
Supplementary figure 3



Supplementary Figure 3: An analysis of gene expression responses that are unique to NAS (A) or fibrosis stage (B). Figures A & B are fold change-fold change plots (NAS vs fibrosis stage), where each point represents a gene, and color represents the posterior probability that a given gene is uniquely differentially regulated with respect to NAS (A) or fibrosis stage (B) (i.e. the

probability approaches one when a gene shows strong evidence of differential regulation in one condition, but not the other). The top 10 genes are labeled in plots A & B. The tables (C & D) show the Reactome gene sets that are enriched for genes with a high degree of uniqueness.

Supplementary Figure 4



Supplementary Figure 4: The cross-validation root-mean-squared error of the lasso models as a function of the regularization parameter. The number of predictors in each model is represented as point size.

Supplemental Table 1: Characteristics of the populations studied

	Histologically normal control N=6	NAFLD N=72
Age (yrs) mean \pm S.D.	49.8 \pm 21	52.8 \pm 12
Gender (males) n	2	28
Caucasians (n)	5	64
Body Mass Index (kg/m ²) mean \pm S.D.	25.9 \pm 5.8	33.3 \pm 5.5
Diabetes (n)	0	29
Hypertension (n)	1	46
Hyperlipidemia (n)	0	42
AST (IU/L) mean \pm S.D.	35.7 \pm 11	60 \pm 40
ALT (IU/L) mean \pm S.D.	37.9 \pm 12	82 \pm 56
Alk Phos (IU/L) mean \pm S.D.	96 \pm 73	93 \pm 33
Bilirubin (mg/dl) mean \pm S.D.	0.6 \pm 0.2	0.6 \pm 0.6
Albumin (gm/dl) mean \pm S.D.	4.3 \pm 0.2	4.4 \pm 0.5
Steatohepatitis (n)	0	58
Steatosis grade: mean	-	2
Lobular inflammation grade: mean	-	1.4
Hepatocellular Ballooning: mean	-	1.8
Fibrosis stage: mean	-	1.3
Distribution of Steatosis grade		
0	6	2
1	-	29
2	-	27
3	-	14
Distribution of Lobular inflammation grade		
0	6	5
1	-	55
2	-	12
Distribution of Hepatocellular ballooning grade		
0	6	14
1	-	28
2	-	30
Distribution of NAS		
0	6	0
1	-	2
2	-	9
3	-	18
4	-	17
5	-	18

6	-	8
Distribution of fibrosis stage		
0	6	19
1	-	28
2	-	9
3	-	14
4	-	2