Supplemental Information

Supplementary Methods and Materials

Ten male cynomolgus monkeys (*macaca fascicularis*) were part of a 21-month experimental time line (Figure 1A). For the first year (Naïve samples), monkeys, 5-6 years of age (average weight 4.5 kg), were acclimatized to the study environment and operant instrumentation. During this acclimatization period, the monkeys were also trained to present their leg for venipuncture to collect blood from the saphenous and/or femoral veins without the use of an anesthetic agent. Plasma samples were collected for endocrine tests, monitoring of blood alcohol levels, and for biomarker discovery and validation. Monkeys were induced to consume liquids under a schedule of food pellet deliveries (i.e., schedule-induced polydipsia (1)) as described previously for these animals (2). Induction conditions did not require food deprivation and were not associated with weight loss. The ethanol was presented in the form of 4% w/v ethanol. Following one month of 0.5 g/kg/day ethanol (two drink equivalents, Induction samples), the animals were escalated to drink 1.0 g/kg/day for 30 consecutive days, and finally, 1.5 g/kg/day for 30 consecutive days. Following the 90-day induction period of alcohol consumption, animals were given unlimited access (22 hours per day) to either ethanol or water for the next six months (Drinking samples) (2,3). Two independent samples (A and B) were collected from each state in the experimental time line (Naïve, Induction, Drinking; Figure 1A). These two independent samples were collected on different days to determine reproducibility of the plasma protein concentrations.

Plasma protein profiling was performed at Rules-Based Medicine, Inc. (Austin, TX) using standard Luminex technology (4). Plasma samples (in triplicate) were subjected to Human

Antigen MAP antigen analysis for 90 different circulating proteins. Sample and capture microspheres were thoroughly mixed and incubated at room temperature for 1 hour. Multiplexed cocktails of biotinylated reporter antibodies were then added and incubated for 1 hour. Multiplexes were developed using an excess of streptavidin-phycoerythrin solution. Analysis was performed in a Luminex 100 instrument. Unknown values for each of the analytes localized in a specific multiplex were determined using 4 and 5 parameter, weighted and non-weighted curve fitting algorithms included in the data analysis package.

Differential abundance of individual plasma proteins was determined using a conservative approach with a one-way repeated measures ANOVA and Bonferroni multiple testing correction ($p < 0.05$). To identify the most consistent plasma protein changes, only those differences significant by a Student Newman-Keuls pair-wise post-hoc test ($p < 0.05$) between two time points and for *both* the A and B samples were considered. Support Vector Machine (SVM) classification analysis was performed using GeneSpring 7.3 (Agilent, Santa Clara, CA) using a polynomial kernel function with no scaling factor as described previously (5). A statistical tool only recently applied to biomedical research, SVM is a supervised machine learning method that allows for classification of samples based on a hyperplane constructed in *n*-dimensional space where *n* is the number of quantitative measures. In this study, the plasma concentrations of the proteins in the biomarker panel were used to classify drinking status. Accuracy, sensitivity, specificity, PPV (positive predictive value), and NPV (negative predictive value) were determined according to standard definitions from classification analysis results (6) (Figure S1). Database searching for tissue origin of proteins was performed using Ingenuity Pathway Analysis software and Swissprot identifiers (Ingenuity, Redwood City, CA).

Table S1. Plasma protein profiling results

Protein Name	Protein Symbol	Bonferroni Corrected ANOVA	Post-hoc Testing			Final Biomarker Panel
			Naive vs Induction	Naive vs Drinking	Induction vs Drinking	
Acid Phosphatase, Prostate	ACCP	ND				
Adiponectin	ADIPOQ	p<0.001	NC	Decreased	Decreased	X
Alpha 1 Antitrypsin	Serpina1a	NC				
Alpha 2 Macroglobin	A2M	p<0.001	NC	Decreased	Decreased	X
Alpha-Fetoprotein	AFP	NC				
Amyloid P Component, Serum	APCS	p<0.001	Variable	NC	Variable	
ApoAI	APOA1	p<0.001	NC	Variable	Variable	
ApoCIII	APOC3	p<0.001	NC	Variable	Variable	
ApoH	APOH	NC				
Beta 2 Microglobin	B2M	p<0.001	Variable	NC	Variable	
Brain-Derived Neurotrophic Factor	BDNF	NC				
C Reactive Protein	CRP	NC				
Calcitonin	CALCA	p<0.001	Variable	Decreased	Variable	
Carcinoembryonic Antigen-Related Cell Adhesion Molecule 5	CEACAM5	ND				
CD 40	CD40	p<0.001	Decreased	Decreased	NC	X
CD 40 Ligand	CD40LG	p<0.001	Variable	Variable	NC	
Chemokine (C motif) Ligand 1	XCL1	ND				
Chemokine (C-C motif) Ligand 11	Ccl11	p<0.001	Variable	Variable	NC	
Chemokine (C-C motif) Ligand 2	CCL2	NC				
Chemokine (C-C motif) Ligand 22	CCL22	NC				
Chemokine (C-C motif) Ligand 3	CCL3	p<0.001	Decreased	Decreased	NC	X
Chemokine (C-C motif) Ligand 4	CCL4	p<0.001	Decreased	Decreased	NC	X
Chemokine (C-C motif) Ligand 5	CCL5	NC				
Chemokine (C-X-C motif) Ligand 5	CXCL5	p<0.001	Decreased	Decreased	NC	X
Colony Stimulating Factor 2	CSF2	p<0.001	Variable	NC	Variable	
Colony Stimulating Factor 3	CSF3	ND				
Complement 3	C3	p<0.001	NC	Decreased	Decreased	X
Creatine Kinase, Muscle	CKM	ND				
Endothelin-1	EDN1	ND				
Epidermal Growth Factor	EGF	p<0.001	Variable	Variable	NC	
Erythropoietin	EPO	ND				
Factor III	F3	p<0.001	Variable	Decrease	NC	
Factor VII	F7	p<0.001	Decreased	Decreased	NC	X
Fatty Acid Binding Protein 3	FABP3	p<0.001	Variable	Decreased	NC	
Ferritin	FT	NC				
Fibrinogen	FG	p<0.001	Variable	Variable	NC	
Fibroblast Growth Factor 2	FGF2	ND				
Glutamic-Oxaloacetic Transaminase 1	GOT1	p<0.001	Increased	Increased	NC	X
Glutathione S-Transferase	GST	ND				
Growth Hormone	GH1	NC				
Haptoglobin	HP	NC				
IFN-gamma	IFNG	ND				
IgA	IgA	NC				
IgE	IgE	p<0.001	Decreased	Decreased	NC	X
IgM	IgM	NC				
Insulin	INS	p<0.001	Variable	Variable	NC	
Insulin-Like Growth Factor 1	IGF1	p<0.001	Decreased	Decreased	NC	X
Intercellular Adhesion Molecule 1	ICAM1	ND				
Interleukin 1 Receptor Antagonist	IL1RN	NC				
Interleukin 10	IL10	p<0.001	Variable	NC	Variable	
Interleukin 12 Subunit p40	IL12-p40	ND				
Interleukin 12 Subunit p70	IL12-p70	p<0.001	Variable	Decreased	Variable	
Interleukin 13	IL-13	p<0.001	Variable	NC	Variable	
Interleukin 15	IL-15	NC				
Interleukin 16	IL-16	ND				
Interleukin 18	IL-18	p<0.001	Variable	Decreased	NC	
Interleukin 1alpha	IL-1alpha	ND				
Interleukin 1beta	IL-1beta	ND				
Interleukin 2	IL2	p<0.001	Decreased	Decreased	NC	X
Interleukin 3	IL-3	ND				

Protein Name	**Protein Symbol**	**Bonferroni Corrected ANOVA**	**Post-hoc Testing**			**Final Biomarker Panel**
			Naive vs Induction	**Naive vs Drinking**	**Induction vs Drinking**	
Interleukin 4	IL-4	ND				
Interleukin 5	IL5	NC				
Interleukin 6	IL-6	ND				
Interleukin 7	IL7	p<0.001	Increased	Increased	NC	X
Interleukin 8	IL8	NC				
Kallikrein-Related Peptidase 3	KLK3	p<0.001	Decreased	Decreased	NC	X
KIT Ligand	KITLG	p<0.001	Variable	Decreased	NC	
Leptin	LSL	ND				
Lipoprotein (a)	LPA	ND				
Lymphotoxin Alpha	LTA	NC				
Matrix Metalloproteinase 2	MMP2	p<0.001	Decreased	Decreased	NC	X
Matrix Metalloproteinase 3	MMP3	ND				
Matrix Metalloproteinase 9	MMP9	p<0.001	NC	Variable	NC	
Myeloperoxidase	MPO	NC				
Myoglobin	MB	NC				
Neighbor of BRCA1 Gene 1	NBR1	ND				
NMDA receptor regulated 1	NARG1	ND				
Pregnancy-Associated Plasma protein A, Pappalysin 1	PAPPA	ND				
S100 calcium binding protein A12	S100A12	ND				
Serine (or cysteine) Peptidase Inhibitor, Clade E, Member 1	Serpine1	p<0.001	Variable	NC	Increase	
Serpin Peptidase Inhibitor, Clade A (alpha-1 antiproteinase, antitrypsin), Member 7	Serpina7	ND				
Sex Hormone-Binding Globulin	SHBG	p<0.001	Variable	Variable	NC	
Thrombopoietin	THPO	p<0.001	Decreased	Decreased	NC	X
Thyroid Stimulating Hormone	TSH	ND				
Tissue Inhibitor of Matrix Metalloproteinase-1	TIMP1	p<0.001	Variable	Decrease	NC	
Tumor Necrosis Factor Alpha	TNFA	ND				
Tumor Necrosis Factor Receptor Superfamily, Member 1B	TNFRSF1B	NC				
Vascular Cell Adhesion Molecule 1	VCAM1	NC				
Vascular Endothelial Growth Factor A	VEGFA	p<0.001	Decreased	Decreased	NC	X
von Willebrand Factor	VWF	ND				

ND, not detected; NC, no change
Note: Proteins in the final biomarker panel are indicated in bold and in the rightmost column.

Table S2. Tissues-of-origin for proteins in biomarker panel. This graph illustrates the anatomical origin for the various components of the biomarker panel. Clearly, proteins originate from a number of different organ systems and some proteins may come from multiple organ systems. This table was generated with information in the Ingenuity Pathway Analysis system (Ingenuity Systems, Redwood City, CA).

Protein Symbol	Alternate Symbol	Entrez Gene Name	Blood	Peripheral Immune Cells	Brain	Adipose	Bladder	Epidermis	Heart	Kidney	Large Intestine	Liver	Lung	Mammary Gland	Ovary	Pancreas	Placenta	Prostate Gland	Retina	Salivary Gland	Skeletal Muscle	Small Intestine	Spleen	Stomach	Testis	Thymus	Uterus
A2M		alpha-2-macroglobulin	X	X					X	X		X	X		X		X		X								X
ADIPOQ		adiponectin	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X		X	X	X	X		X	X	X
C3		complement component 3	X	X	X	X	X	X	X		X	X	X	X	X		X	X	X	X	X		X	X		X	X
CCL3	MIP-1alpha	chemokine (C-C motif) ligand 3	X	X	X																						X
CCL4	MIP-1beta	chemokine (C-C motif) ligand 4	X	X						X		X											X	X			X
CD40		CD40 molecule, TNF receptor family member 5	X	X	X	X	X			X	X	X	X	X	X	X	X		X			X	X	X	X	X	X
CXCL5	ENA-78	chemokine (C-X-C motif) ligand 5	X	X																							
F7	Factor 7	coagulation factor VII	X	X								X															
GOT1	SGOT	glutamic-oxaloacetic transaminase 1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X		X	X
IgE		immunoglobulin E	X	X																							
IGF1		insulin-like growth factor 1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
IL2		interleukin 2	X	X												X							X				
IL7		interleukin 7	X	X		X																X	X			X	
KLK3	PSA	kallikren-related peptidase 3	X											X				X		X							
MMP2		matrix metallo-peptidase 2	X	X		X	X	X	X	X	X	X	X		X	X	X		X									X
THPO		thrombopoietin	X	X	X							X															
VEGFA		vascular endothelial growth factor A	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

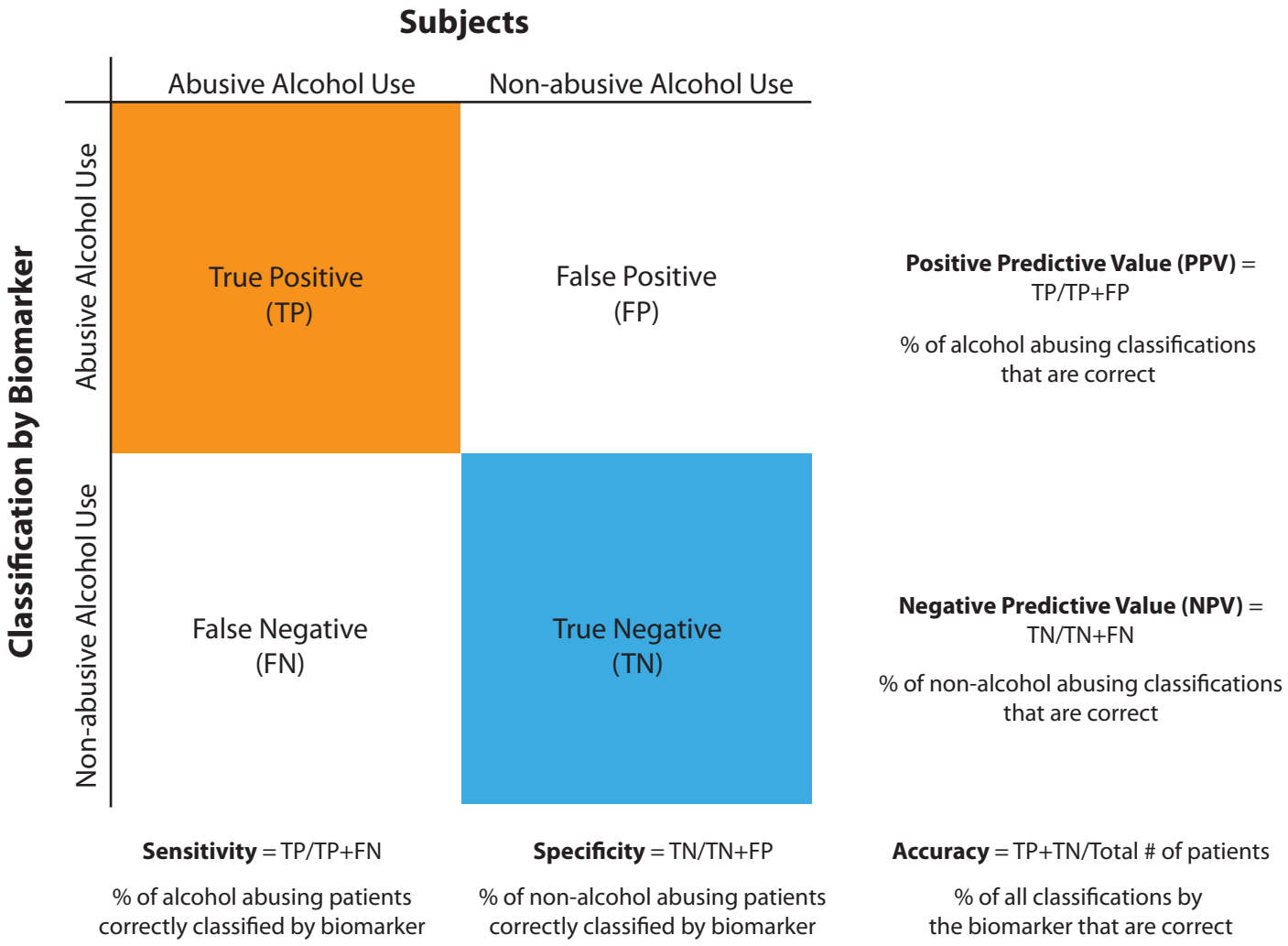

Figure S1. Statistical evaluation of biomarker performance.
Unlike traditional statistics that examine differences between populations, diagnostics are intended to be informative of the individual patient. A biomarker's ability to correctly classify subjects can be summarized by sensitivity, specificity, positive predictive value, negative predictive value, and accuracy measures. In this example, the ability to classify abusive and non-abusive alcohol use is portrayed. Sensitivity and specificity are the most commonly used measures, and, in this case, they provide the percentages of subjects correctly identified as abusing and not abusing alcohol respectively.

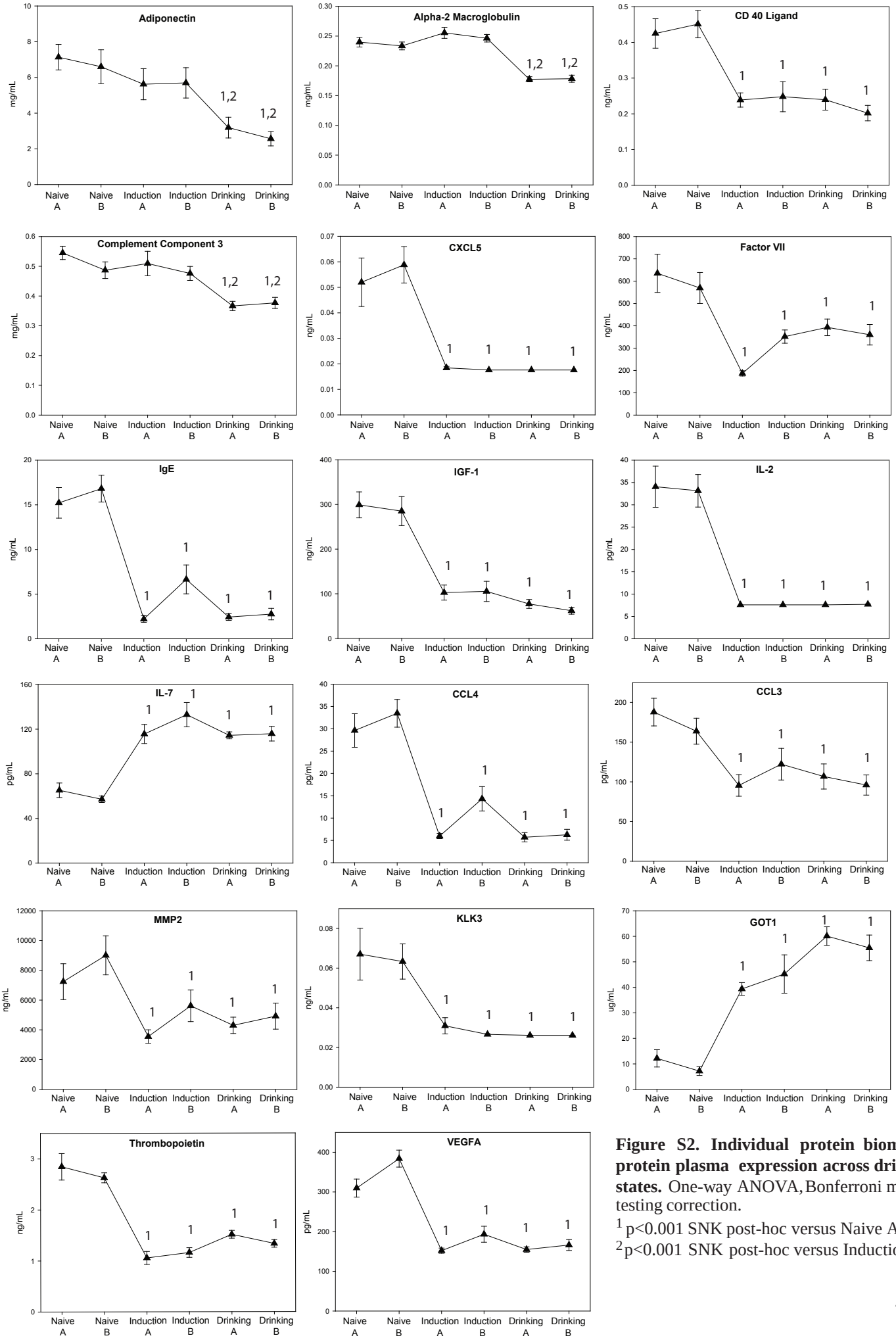

Figure S2. Individual protein biomarker protein plasma expression across drinking states. One-way ANOVA, Bonferroni multiple testing correction.

[1] $p<0.001$ SNK post-hoc versus Naive A & B

[2] $p<0.001$ SNK post-hoc versus Induction A & B

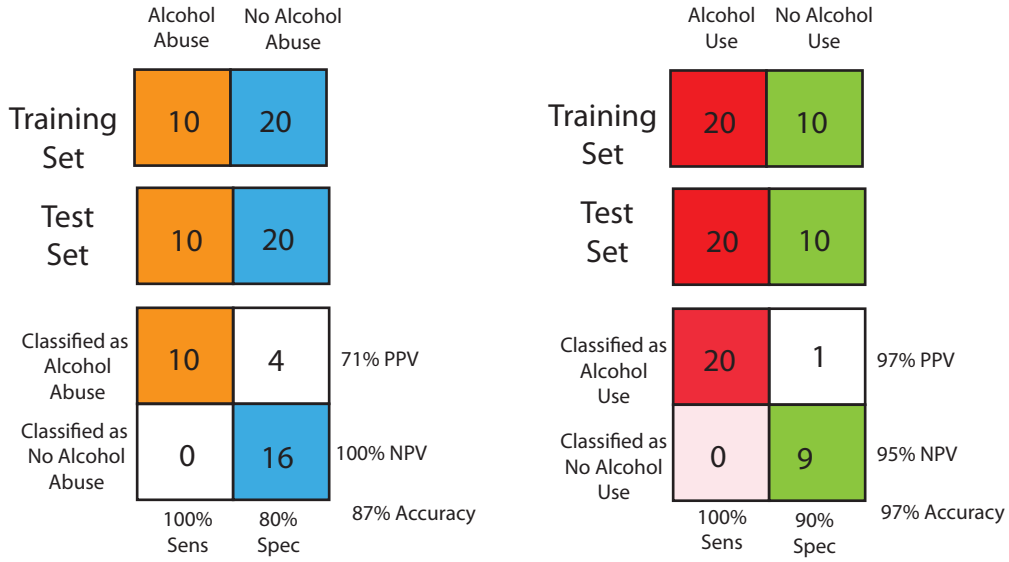

Figure S3. Classification analysis using a training set/test set approach. Using A samples from each drinking state as the training set the B samples were classified using a support vector machine approach. Similar results were obtained from flipping of training and test sets. PPV, positive predictive value; NPV, negative predictive value; Sens, sensitivity; Spec, specificity.

1. Falk D (1961): Production of polydipsia in normal rats by an intermittent food schedule. *Science* 133: 195-196.

2. Grant KA, Leng X, Green HL, Szeliga KT, Rogers LS, Gonzales SW (2008): Drinking typography established by scheduled induction predicts chronic heavy drinking in a monkey model of ethanol self-administration. *Alcohol Clin Exp Res* 32: 1824-1838.

3. Vivian JA, Green HL, Young JE, Majerksy LS, Thomas BW, Shively CA, *et al.* (2001): Induction and maintenance of ethanol self-administration in cynomolgus monkeys (Macaca fascicularis): long-term characterization of sex and individual differences. *Alcohol Clin Exp Res* 25: 1087-1097.

4. Vignali DA (2000): Multiplexed particle-based flow cytometric assays. *J Immunol Methods* 243: 243-255.

5. Freeman WM, Bixler GV, Brucklacher RM, Lin CM, Patel KM, Vanguilder HD, *et al.* (2009): A multistep validation process of biomarkers for preclinical drug development. *Pharmacogenomics J* Dec 8 [Epub ahead of print].

6. Freeman WM, Vrana KE (2010): Future prospects for biomarkers of alcohol consumption and alcohol-induced disorders. *Alcohol Clin Exp Res* In press.