

Dione T. Kobayashi, PhD 1; Brett Chung 1; Laurie Stephen, PhD 2; Karri L. Ballard, PhD 2; Ruth Dewey 2; Jing Shi, MD/PhD 2; Michael Walker, PhD 2; Kathleen McCarthy 1; Sergey Paushkin, MD/PhD 1; Cynthia Joyce 1; James Mapes, PhD 2; Karen S. Chen, PhD 1

1 SMA Foundation, New York, NY; 2 Rules-Based Medicine, Austin, TX

## Introduction

### Background:

Biomarkers are measures of biological state that are useful in assessing disease severity, designing proof of concept trials, tracking therapeutic responses, and generating hypotheses about disease pathophysiology. Biomarkers in Spinal Muscular Atrophy (SMA) related to muscle and nerve physiology, imaging, SMN protein and transcript are being explored in other initiatives, and there is also a need to investigate potential new plasma protein biomarkers that could aid in prognosis or monitoring treatment efficacy in SMA patients.

### Objective:

Our objective was to use a multiplexed immunoassay panel (MAP) technology to expand on the plasma protein biomarker discovery aspect of the prospective Biomarkers for SMA (BforSMA) clinical study conducted in 2009 and create a custom biomarker panel for SMA for use in the research and clinical communities.

### Strategy:

Using two existing Rules-Based Medicine Multi-analyte Profile (MAP) panels comprising 267 markers (DiscoveryMAP and OncologyMAP), SMA and control plasma samples collected from the BforSMA study were analyzed to identify novel plasma protein biomarker candidate hits that correlate to the Modified Hammersmith Functional Motor Scale (MHFMS) measure of disease severity in SMA patients. New immunoassays for the best hits identified by LC/MS that were not in existing MAPs were also developed.

### Results:

Testing in the existing MAPs produced 72 new candidate marker hits. Another 8 markers hits identified separately in the BforSMA study were reproduced in the MAPs. Working assays for the top 9 BforSMA hits have been produced in singleplex and pilot multiplex formats, and a final panel of the top 35 combined hits will be tested with new SMA plasma samples to verify the candidate markers.

## Methods

### Data and Samples

The data herein was generated in an effort to identify a marker or panel of markers in plasma from a wide range of SMA patients that segregates with measures of clinical severity. Markers were identified as hits for their association with disease severity as determined by motor function or other functional outcome measures. Data was generated across three different discovery studies or campaigns. The first platform was a LC/MS ITRAQ analysis done by BG Medicine. The second and third were the DiscoveryMAP and OncologyMAP multiplexed immunoassay panels developed by Rules-Based Medicine. Samples used were plasma samples collected in the Biomarkers for SMA Study. This study was a multi-center, pilot study enrolling 130 subjects, age 2 to 12 years from 18 academic pediatric neuromuscular clinics. Each subject was seen for a single visit, during which an assessment of functional ability, pulmonary status and nutritional status was performed. No therapeutic intervention occurred.

Three groups of SMA patients and one cohort of control children were enrolled according to the classifications below:

- Children with **type I** SMA (n=17).
- Children with **type II** SMA (n=49).
- Children with **type III** SMA (n=42).
- Control children (n=22)

**129 plasma samples were collected from the SMA patients and matched control subjects. These samples were used in ensuing LC/MS and MAP platform biomarker analyses.**

Other samples from the Pediatric Neuromuscular Clinical Research Network (PNCN) SMA Natural History Study (NHS) from Columbia University, Boston Children's Hospital and Children's Hospital of Philadelphia were used for assay development. This study was a multisite, longitudinal prospective study enrolling 101 patients age 4 months to 45 years from three academic pediatric neuromuscular clinics. Subjects were seen at 0, 6, 12, 18, 24, 30 and 36 month visits, during which assessments using multiple motor scales and tests were made; pulmonary status, strength, nerve/muscle physiology and nutritional status assessments were also performed. No therapeutic intervention occurred.

Three SMA groups from the PNCN NHS were enrolled according to the classifications below:

- Subjects with **type I** SMA (n=27).
- Subjects with **type II** SMA (n=40).
- Subjects with **type III** SMA (n=34).

### Statistical Analysis

The LC/MS analyte data was analyzed using univariate and multivariate regression methods. Age and gender were controlled for in the analysis. Hits were identified by significant p-values whereby the lowest p-values were the strongest hits. The MAPs were analyzed using ANOVA, t-test, Fisher's Exact test, and Pearson's correlation and by multivariate regression analysis (linear, lasso, random forest). Analytes that had a high number of missing values (greater than 40% below limit of detection) were excluded from the analyses. Analytes whose values were driven by a few outliers generally fell out of the hitlist using these analyses. The false discovery rate (FDR) was controlled by the Benjamin Hochberg method and the FDR q-value cutoff was set at 0.10.

Further multivariate analyses were performed on both the LC/MS and MAP data to identify the best hits from all three discovery campaigns (Figure 1). The ability of top hits to predict SMA function scores is represented in Figure 2.

### LC/MS ITRAQ

Abundant proteins were depleted in order to facilitate a good dynamic range of protein measurements. In this project a dual affinity depletion strategy was implemented: in the first stage, 14 highly abundant proteins (serum albumin, IgG, fibrinogen, transferrin, IgA, IgM, haptoglobin, α-2-macroglobulin, α-1 acid glycoprotein, α-1 antitrypsin, Apo A-I, Apo A-II, complement C3, and Apo B-100) were depleted by IgY antibody column. The remaining proteins were extracted from non proteinaceous components by reverse-phase chromatography. The proteins were reduced, alkylated (cysteine residues) and digested with trypsin. The resulting peptide pool was then labeled with the amine specific ITRAQ reagents. Eight samples labeled with eight different isotope-coded versions of the ITRAQ reagent were combined into eight-plex ITRAQ mixes and analyzed as a single sample using mass spectrometry. Each ITRAQ mix is analyzed by two dimensional LC-MS/MS. ITRAQ mixes are pre-fractionated by strong cation exchange into six fractions that are further separated by HPLC. The mass-to-charge ratios are the primary data used to determine whether individual analytes are "hits" (Table 1).

### Immunoassay Multi-Analyte Profile (MAP)

Multiplexing is accomplished by assigning each analyte-specific assay a microsphere set labeled with a unique fluorescence signature. To attain distinct microsphere signatures, two fluorescent dyes, red and far red, are mixed in various combinations using different intensity levels of each dye. Each set of microspheres is encoded with a fluorescent signature by impregnating the microspheres with a unique dye combination. After the encoding process, an assay-specific capture reagent is conjugated covalently to each unique set of microspheres, creating an ELISA-like assay on each bead surface.

After optimizing the parameters of each assay separately, Multi-Analyte Profiles (MAPs) are performed by mixing up to 100 different sets of the microspheres in a single well of a 96- or 384-format microtiter plate. A small sample volume of plasma (10UL – 20UL) is added to the well and allowed to react with the microspheres. The assay-specific capture reagent on each individual microsphere binds the analyte of interest. A cocktail of assay specific, biotinylated detecting reagents (e.g., antigens, antibodies, ligands, etc.), is reacted with the microsphere mixture, followed by a streptavidin-labeled fluorescent "reporter" molecule. Finally, the multiplex is washed to remove unbound detecting reagents. After washing, the mixture of microspheres is analyzed using the Luminex 100™ instrument. Similar to a flow cytometer, the instrument uses hydrodynamic focusing to pass the microspheres in single file through two laser beams. As each individual microsphere passes through the excitation beams, it is analyzed for size, encoded fluorescence signature and the amount of fluorescence generated in proportion to the analyte. As each microsphere is encoded with a unique signature, the classification identifies the analyte being measured on that individual microsphere. As the microsphere passes through a green diode-pumped solid state laser (532 nm), a fluorescence "reporter" signal (580 nm) is generated in proportion to the analyte concentration. Analyte concentrations from the DiscoveryMAP and OncologyMAP panels were analyzed to identify hit proteins (Table 2, 3).

### SMA MAP Panel

The top hits were identified and selected for their statistical value and assay performance for use in a new SMA MAP panel with both LC/MS and DiscoveryMAP and OncologyMAP analytes. New immunoassays were created for the top 9 LC/MS 9 analyte hits: Tenascin XB, Cartilage Intermediate Layer Protein 2, CD93, Cartilage Oligomeric Protein, Lumican, Cadherin13, Dipeptidyl peptidase 4, Peptidase D, and Thrombospondin-4 (THB4 example in Figure 2).

## Results

Table 1: LC/MS Plasma Protein Hitlist

PROTEIN	DESCRIPTION	Q-VALUE
CLIP5	CARTILAGE INTERMEDIATE LAYER PROTEIN 2	<0.001
COMP	CARTILAGE OLIGOMERIC LAYER PROTEIN 2	<0.001
TNXB	TENASCIN XB (DISTINCT ISOFORM)	<0.001
CLEC3B	C-TYPE LECTIN DOMAIN FAMILY 3, MEMBER B	<0.001
ADAMTS4	ADAMTS, HE 4	<0.001
TNXB	TENASCIN XB (DISTINCT ISOFORM)	<0.001
DPF4	DIPEPTIDYLPEPTIDASE 4 (CD26, ADENOSINE DEAMINASE COMPLEXING PROTEIN 2)	<0.001
THBS4	THROMBOSPONDIN 4	<0.001
CDH13	CADHERIN 13, H-CADHERIN (HEART)	<0.001
OMD	OSTEOCALCIN	<0.001
CRPAC1	CARTILAGE INTERMEDIATE LAYER PROTEIN 1	<0.001
F13B	COAGULATION FACTOR XIII, B POLYPEPTIDE	<0.001
CIQR1	COMPLEMENT COMPONENT 1, Q SUBCOMPONENT, RECEPTOR 1	<0.001
LUM	LUMICAN	<0.001
PEPD	PEPTIDASE D	<0.001
APCS	AMPHILOD P COMPONENT, SERUM	<0.001
COL2A1	COLLAGEN, TYPE II, ALPHA 1 (PRIMARY OSTEOARTHRITIS, SPONDYLOEPHYSEAL DYSPLASIA, CONGENITAL)	<0.001
NOV	NEPHROBLASTOMA OVEREXPRESSED GENE	<0.001
C2B	BISED COMPLEMENT C2B	<0.001
COL6A3	COLLAGEN, TYPE VI, ALPHA 3	0.001
CRP	C-REACTIVE PROTEIN, PENTAMER/RELATED	0.002
VTN	VITRONECTIN (SERUM SPREADING FACTOR, SCAMTOMEDIN B, COMPLEMENT SPROTEIN)	0.002
GSN	GELSULIN (AMPHILODIN, FINNISH TYPE)	0.003
VTN	VITRONECTIN (SERUM SPREADING FACTOR, SCAMTOMEDIN B, COMPLEMENT SPROTEIN)	0.003
MB	MYOGLOBIN, BETA	0.003
ALPL	ALKALINE PHOSPHATASE, LIVER/OSTEOCYTE	0.003
APOA4	APOLIPOPROTEIN A IV	0.004
FCGR3A	FC FRAGMENT OF IGG, LOW AFFINITY B1, RECEPTOR (CD16A)	0.004
NEO1	NEODOMIN HOMOLOG 1 (CHICKEN)	0.004
C2B	BISED COMPLEMENT C2B	0.004
LRG1	LEUCINRICH ALPHA2-GLYCOPROTEIN 1	0.004
CTSD	CATEPSIN D (LYSOSOMAL ASPARTYL PEPTIDASE)	0.006
C2	COMPLEMENT COMPONENT 2	0.010
RPS27A	UBIQUITIN AND RIBOSOMAL PROTEIN S27A PRECURSOR	0.006
IF	I FACTOR (COMPLEMENT)	0.006
IGFBP6	INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 6	0.006
INHBC	INHIBIN, BETA C	0.008
F13A1	COAGULATION FACTOR XIII, A1 POLYPEPTIDE	0.009
NCAM1	NEURAL CELL ADHESION MOLECULE 1	0.009
CP	CERULOPLASMIN (CERULOPLASMIN)	0.010
AGA	ASPARTYLGLUTAMINASE	0.012
C2	COMPLEMENT COMPONENT 2	0.013
HBA1	HENNOGLOBIN SUBUNIT ALPHA	0.016
LRG1	LEUCINRICH ALPHA2-GLYCOPROTEIN 1	0.018
SPP1	SECRETED PHOSPHOPROTEIN 1 (OSTEOPOINTIN, BONE SIALOPROTEIN I, EARLY TLR4/MYD88 ACTIVATION 1)	0.018
GM	CREATINE KINASE, MUSCLE	0.018
ILR	INTERLEUKIN 6 RECEPTOR	0.018
PRG4	PROTEOGLYCAN 4	0.020
MRC2	MANNOSE RECEPTOR, C TYPE 2	0.021
CA2	CARBONIC ANHYDRASE II	0.025
QSOX2	QSOXIN 2	0.029
PROCR	PROTEIN C (INACTIVATOR OF COAGULATION FACTORS VA AND VIBA)	0.037
PKR1	PHOSPHOLIPASE C KINASE 1	0.037
CST6	CYSTATHIONINE	0.037
S100A4	S100 CALCIUM BINDING PROTEIN A4 (CALCULIN, CALVALCIN, METASTASIN, MURINE PLACENTAL HOMOLOG)	0.037
SERPINA10	SERRIN PEPTIDASE INHIBITOR, CLADE A (ALPHA1 ANTITRYPSIN, ANTI-TRYPAIN, MEMBER 10)	0.039
ENG	ENDOGLIN (DLER/RENGU/WEBER SYNDROME 1)	0.042
PARK7	PARKIN (DOLER/RENGU/WEBER SYNDROME 2)	0.042
VCMAM1	VASCULAR CELL ADHESION MOLECULE 1	0.042
MET	MET PROTEINOSINASE (HAPTOCYTE GROWTH FACTOR RECEPTOR)	0.055
PEBP4	PHOSPHATIDYLETHANOLAMINE BINDING PROTEIN 4	0.058
ORM2	OSIOMUCOID 2	0.060
CNDP1	CARNOSE DIPEPTIDASE 1 (METALLOPEPTIDASE M20 FAMILY)	0.066
FAP	FIBROBLAST ACTIVATION PROTEIN, ALPHA	0.066
CEP1	CEPHEIN 1 (HOMOCYSTEINE)	0.067
VNN1	VANNIN 1	0.067
CA1	CARBONIC ANHYDRASE I	0.069
PTPRG	PROTEIN TYROSINE PHOSPHATASE, RECEPTOR TYPE, G	0.069
MB	MYOGLOBIN	0.069
ND1	NEODOMIN 1	0.069
TPH1	TRYPTOPHAN 5-HYDROXYLASE 1	0.069
PF4	PLATELET FACTOR 4 (CHEMOKINE [CXCL12] MOTIF) LIGAND 4	0.072
PCP2	PLASMIN GENITRINIC CARBOXYPEPTIDASE	0.077
S100A8	S100 CALCIUM BINDING PROTEIN A8 (CALCAKRANIN A)	0.081
GAPDH	GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE	0.081
CNTN4	CONTACTIN 4	0.081
SHBG	SEX HORMONE BINDING GLOBULIN	0.081
TAGLN2	TRANSGLUTININ 2	0.081
AFM	AFAMIN	0.081
NP	NUCLEOSIDE PHOSPHATASE	0.081
HP	HAPTOGLOBIN	0.081
DAG1	DYSTROGLYCAN 1 (DYSTROPHIN-ASSOCIATED GLYCOPROTEIN 1)	0.081
PROCR	PROTEIN C RECEPTOR, ENDOTHELIAL (EPCR)	0.082
ACTN1	ACTIN, ALPHA 1	0.082
CSF1R	COLONY STIMULATING FACTOR 1 RECEPTOR, FORMERLY MCDONOUGH FELINE SARCOMA VIRAL (VMS) ONCOGENE HOMOLOG	0.082
PRDX2	PEROXIREDOXIN 2	0.082
GSS	GLUTATHIONE SYNTHETASE	0.082
S100A9	S100 CALCIUM BINDING PROTEIN A9 (CALCAKRANIN B)	0.082
TNFR	TNF RECEPTOR	0.082
SOX1	SUPEROXIDE DISMUTASE 1, SOLUBLE (AMPHOTROPIC LATERAL SLENDIN 1 LIGAND 1)	0.082
PRDX6	PEROXIREDOXIN 6	0.083
SERPIN1	SERRIN PEPTIDASE INHIBITOR, CLADE D (PARFAN COFACTOR), MEMBER 1	0.090
MIRN2	MIRN2	0.090
CHAD	CHONDROHERNIN	0.090
CACNA2D1	CALCIUM CHANNEL, VOLTAGE DEPENDENT, ALPHA 2DELTA SUBUNIT 1	0.096
ORM1	OSIOMUCOID 1	0.109
AOC3	AMINE OXIDASE, COPPER CONTAINING 3	0.131

Table 2: DiscoveryMAP Plasma Protein Hitlist

PROTEIN	DESCRIPTION	P-VALUE
MB	Myoglobin	<0.001
SPP1	Osteopontin	<0.001
AXL	AXL Receptor Tyrosine Kinase	<0.001
APC3	Serum Amyloid P Component	<0.001
CRP	C-Reactive Protein	<0.001
MDC	Macrophage Derived Chemokine	<0.001
TM	Thrombospondin	<0.001
CALCA	Calcalcin	<0.001
LEP	Leptin	<0.001
BNP	Brain Natriuretic Peptide	<0.001
MMP2	Matrix Metalloproteinase 2	<0.001
CKM	Creatine Kinase MB	<0.001
ACE	Angiotensin Converting Enzyme	<0.001
FABP4	Fatty Acid Binding Protein Heart	<0.001
CD40	CD40 Ligand	<0.001
MF	Macrophage Migration Inhibitory Factor	<0.001
ANGPT2	Angiotensin 2	<0.001
FETUA	Fetuin A	0.001
CHFH	Complement Factor H	0.001
IL8	Interleukin 8	0.001
C3	Complement C3	0.001
PPP	Pancreatic Polypeptide	0.001
FLT1	Vascular Endothelial Growth Factor	0.002
TF	Transformin	0.002
PLGF	Placenta Growth Factor	0.002
EGF	Epidermal Growth Factor	0.002
GSTA1	Glutathione S Transferase alpha	0.002
SOD1	Superoxide Dismutase 1	0.003
VCAM1	Vascular Cell Adhesion Molecule 1	0.003
PAI1	Plasminogen Activator Inhibitor 1	0.004
M-CSF	Macrophage Colony Stimulating Factor	0.004
S100A12	S100 Protein A12	0.004
VTN	Vitronectin	0.004
FASLG	Fas Ligand	0.004
ATM	Alpha 1 Microglobulin	0.007
SOT	Serum Glucocorticoid Oxaloacetic Transaminase	0.009
ACCT	Alpha 1 Antichymotrypsin	0.01
MFI1b	Macrophage Inflammatory Protein 1 beta	0.011
SOX1	Sorlexin	0.013
TBG	Thyroxine Binding Globulin	0.014
APOA1	Apolipoprotein A1	0.015
MPO	Myeloperoxidase	0.016
B2M	Beta 2 Microglobulin	0.016
EPO	Erythropoietin	0.017
MMP10	Matrix Metalloproteinase 10	0.02
PROS1	Protein S	0.023
MMP9	Matrix Metalloproteinase 9	0.025
RAGE	Receptor for advanced glycosylation end products	0.029
IL18	Interleukin 18	0.031
CCL11	Chemokine C-C Motif Ligand 11	0.034
IgG	Immunoglobulin A	0.035
C-peptide	Proinsulin C Peptide	0.041
A2M	Alpha 2 Macroglobulin	0.043
PDGF-BB	Platelet Derived Growth Factor	0.047
CCL16	Chemokine C-C Motif Ligand 16	0.047
IL1a	Interleukin 1 alpha	0.059
APOA4	Apolipoprotein A IV	0.059
MMP9	Matrix Metalloproteinase 9	0.05

## Results

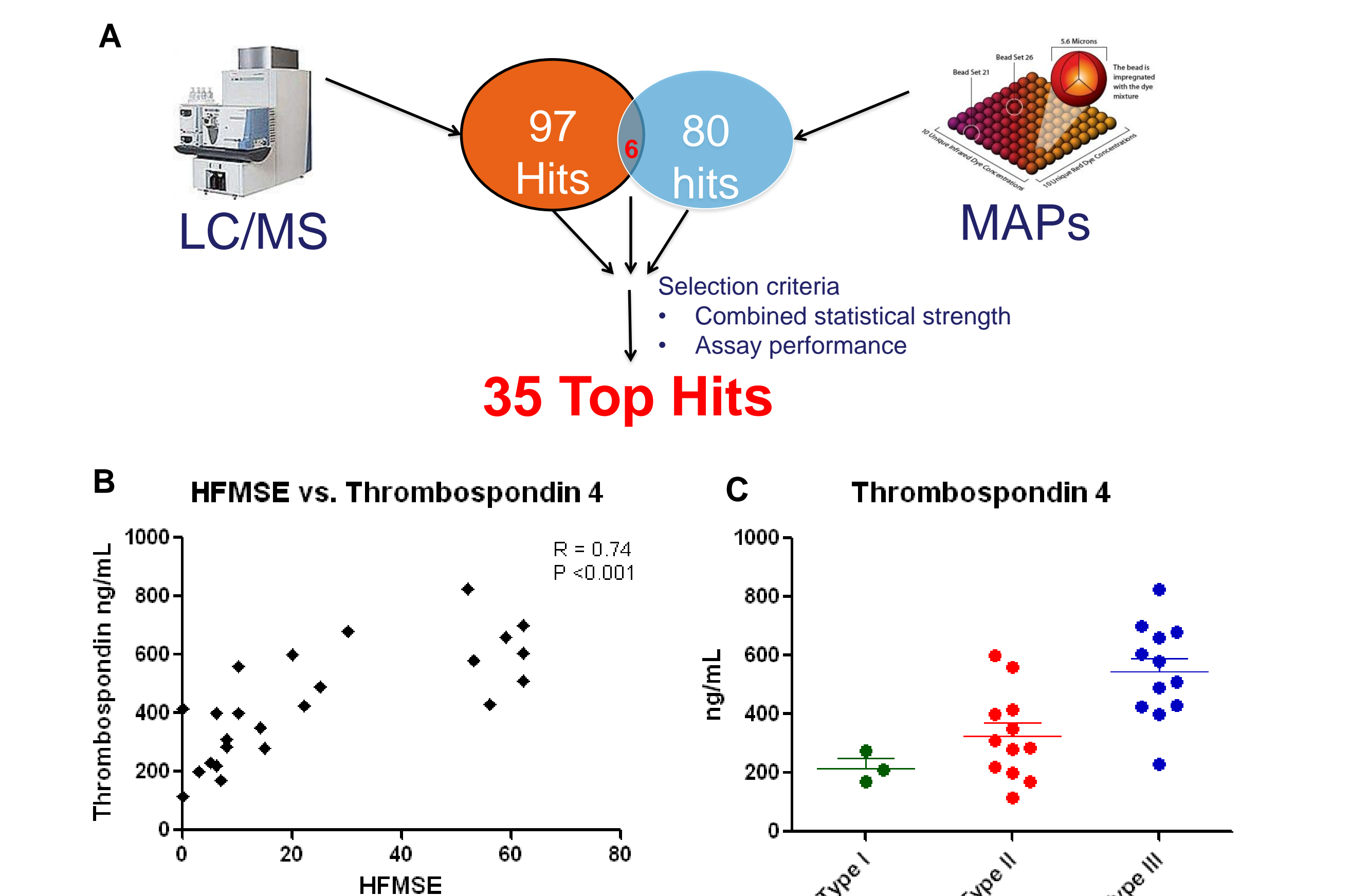
Table 3: OncologyMAP Plasma Protein Hitlist

PROTEIN	DESCRIPTION	P-VALUE
SPP1	OSTEOPOINTIN (SECRETED PHOSPHOPROTEIN 1)	<0.001
CLEC3B	C-TYPE LECTIN DOMAIN FAMILY 3, MEMBER B; TETRAMECTIN	<0.001
IGFBP6	INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 6	<0.001
FABP4	FATTY ACID BINDING PROTEIN 4, ADIPOCYTE	<0.001
CHL1	CHITINASE 3-LIKE 1 (CARTILAGE GLYCOPROTEIN-39), YKL-40	<0.001
LEP	LEPTIN	<0.001
CTSD	CATEPSIN D	0.001
MST1	MACROPHAGE STABILIZING 1 (HAPTOCYTE GROWTH FACTOR 6)	0.001
MF	MACROPHAGE MIGRATION INHIBITORY FACTOR (GLYCOSYLATION INHIBITING FACTOR)	0.001
S100A4	S100 CALCIUM BINDING PROTEIN A4	0.001
GLO1	GLYCOLYASE 1; LACTOLYLULATHIONE LYASE	0.001
ENG	ENDOGLIN	0.001
FTL1	FMS-RELATED TYROSINE KINASE 1 (VASCULAR ENDOTHELIAL GROWTH FACTOR/VASCULAR PERMEABILITY FACTOR RECEPTOR)	0.002
ERBB2	HUMAN EPIDERMAL GROWTH FACTOR 2	0.003
NDK-8	NUCLEOSIDE PHOSPHATASE KINASE ISOFORM B	0.004
PRDX-4	PEROXIREDOXIN 4	0.01
PLAUR	PLASMINOGEN ACTIVATOR, UROKINASE RECEPTOR	0.015
IL6R	INTERLEUKIN 6 RECEPTOR	0.02
CDL2A	CHEMOKINE [C-C MOTIF] LIGAND 2A	0.038
GSN	GELSULIN	0.038
PSAT	PHOSPHOSERINE AMINOTRANSFERASE	0.039
TGFB1	TRANSFORMING GROWTH FACTOR, BETA 1, LATENCY ASSOCIATED	0.049

### SMA Plasma Protein Biomarker Identification and Performance

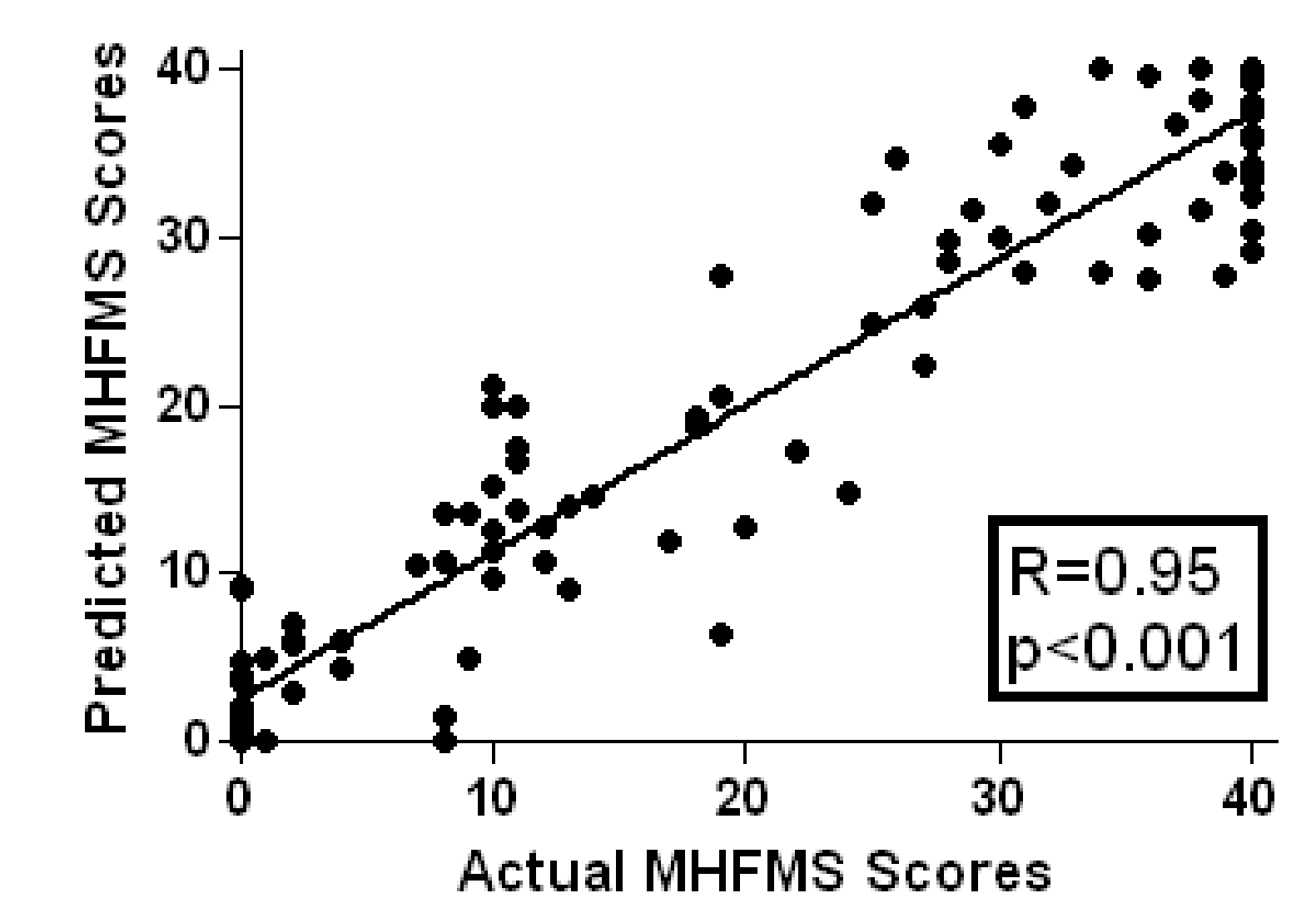
A: BforSMA plasma samples were evaluated for relationships to the Modified Hammersmith Functional Motor Scale (MHFMS) in both LC/MS and Luminex immunoassay MAP platforms. There were 45+ analytes in common between the LC/MS and MAP platforms and the results were correlated well between them (R=0.65). Of the 9 analytes that were LC/MS hits that were represented in the MAPs, 6 of them reproduced as hits. The top combined hits from both platforms were selected on a basis of statistical strength and assay performance, and many of them have either restricted expression in bones, nervous system or muscle tissue or enriched expression in these tissues. B-C: New immunoassays were generated for the top 9 biomarkers identified by LC/MS analysis. These assays were developed and quality tested with SMA samples from the PNCN Natural History Study collection and analyzed against the Expanded HFMS (HFMS). The plasma protein levels of these analytes correlated with the HFMS with R-values that ranged from 0.66 to 0.87 and had differentiable levels across SMA types - with the exception of Tenascin XB. A reliable assay for CILP2 was not developed due to lack of usable reagents.

Figure 1: Hit Generation Plan and Representative New Analyte Relationship to SMA Function and Type



**Preliminary Predictive Power of a Subset of SMA Plasma Protein Analytes**  
A group of 25 MAP plasma protein analytes were used to generate predicted MHFMS scores for the BforSMA study samples from SMA patients. The subpanel analytes were able to account for up to 95% of the variance in the MHFMS scores.

Figure 2: Correlation between Actual MHFMS Scores and Scores Predicted by a Panel of 25 Protein Hits



## Discussion and Next Steps

- Hits from the LC/MS and MAP biomarker discovery studies have been combined into a new custom multiplexed immunoassay panel for SMA
- The SMA panel is being tested with other SMA sample collections to verify the initial BforSMA results
  - Results will be available by end of July
  - Analysis will include regression against several other SMA motor scales (e.g. GMFM, HFMS) and outcome measures (e.g. FVC, CMAP/MUNE)
- This SMA panel will be available in August 2011 for further community testing as a prognostic and possibly pharmacodynamic biomarker tool
  - A streamlined version of the panel will be available in Fall 2011
- Please see BforSMA database poster #26B

## Acknowledgments

Our thanks is given to the following groups and people for their contributions to this work: The Biomarkers for SMA Study Group and the PNCN SMA Biorepository provided the plasma samples and SMA clinical data, Sarah Przedborski worked on poster design, Jason Scull of Rules-Based Medicine developed new panel assays.

All work was funded by the SMA Foundation.