Discovery of neurological biomarker signatures using SOMAscan™ Multiplex Proteomic Technology

Sally K. Nelson[1], David Sterling[1], Merit E. Cudkowicz[2], Martina Sattlecker[3], Steven Kiddle[3], Richard Dobson[3], Geoffrey Baird[4] , Nick Saccomano[1], Stephen Williams[1]

[1] SomaLogic Inc., Boulder, CO
[2] Massachusetts General Hospital, Boston, MA
[3] King's College London, Institute of Psychiatry, London, UK
[4] University of Washington, Seattle, WA

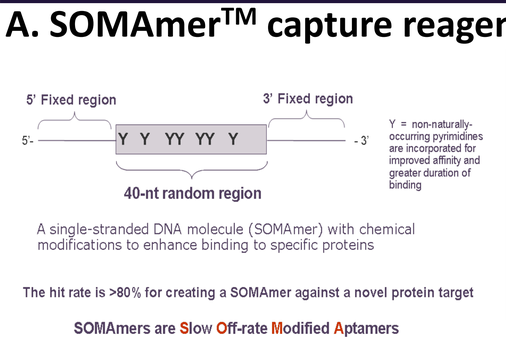
SomaLogic

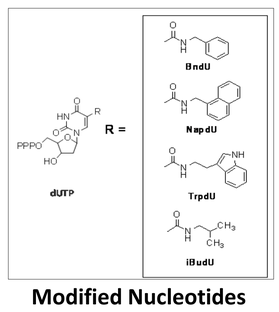
Abstract

The potential of blood-based protein biomarkers to diagnose disease early and accurately could greatly enhance medical care and offer new avenues for therapeutic discovery and development. However, the discovery of such biomarkers is compromised by the technological challenges of profiling the circulating proteome: a complex mixture of thousands of proteins that range from high (µM) to exceedingly low (fM) concentrations. Using a new high-content multiplex proteomics assay (SOMAscan), we assayed thousands of blood samples for biomarker discovery and algorithm training. The archived samples spanned >30 sites in the US and Europe, with blood and CSF samples from a wide range of central nervous system diseases that include Amyotrophic Lateral Sclerosis (ALS), Multiple Sclerosis (MS), Parkinson's Disease (PD) and Alzheimer's Disease (AD). Robust biomarkers were identified in ALS to develop preliminary classifiers which performed well on our test sets with differential expression seen between samples from disease patients and healthy controls with an estimated AUC = 0.93 (0.83, 0.97) and an associated sensitivity of 0.89 ± 0.02 and specificity of 0.8 ± 0.03. These biomarkers were specific to ALS as they were not affected by neurodegeneration-based symptom mimics such as MS or AD. We also identified 11 blood proteins that were significantly (q-value < 0.05) correlated with AD diagnosis. Multivariate analysis showed that a set of 13 proteins can predict an independent test set with a sensitivity of 67% and a specificity of 64%. SomaLogic's SOMAscan technology has been shown to replicate findings from immunoassay, gel electrophoresis and mass spectrometry proteomic studies demonstrating both the opportunities and the challenges in the validation of blood-based protein biomarkers of AD onset and progression.

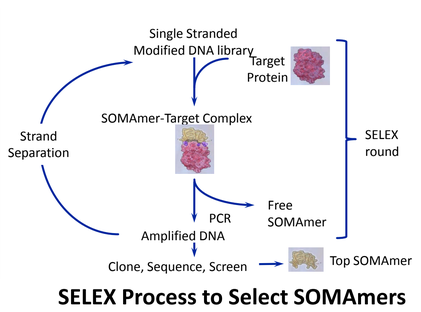
Assay Principles

A. SOMAmer™ capture reagents are made of modified DNA





Modified Nucleotides



SELEX Process to Select SOMAmers

To date, SOMAmer reagents have been selected to 1129 diverse protein targets of biological interest. Most lead SOMAmers have affinities better than 1 nM. The SOMAmer menu covers targets with a wide range of pI's, diverse molecular functions, and several known disease or physiology associations (1).

B. SOMAscan Assay Overview

The SOMAscan assay converts protein signal to nucleic acid signal, through the SOMAmer binding reagent enabling the use of conventional, fully validated DNA detection methods for detection (**Figure 1**). The latest version SOMAscan discovery assay, v3.0, is multiplexed to 1129 analytes per sample, automated using a Biomek FxP (2), and detected using custom Agilent 15,000 spot microarrays. Once discovered, a panel of biomarkers can be translated to a streamlined assay in just a few weeks with a standardized protocol. The overall assay performance and throughput are detailed in **Table 1**.

Fig. 1. SOMAscan: Empirical Biomarker Discovery in 5 weeks

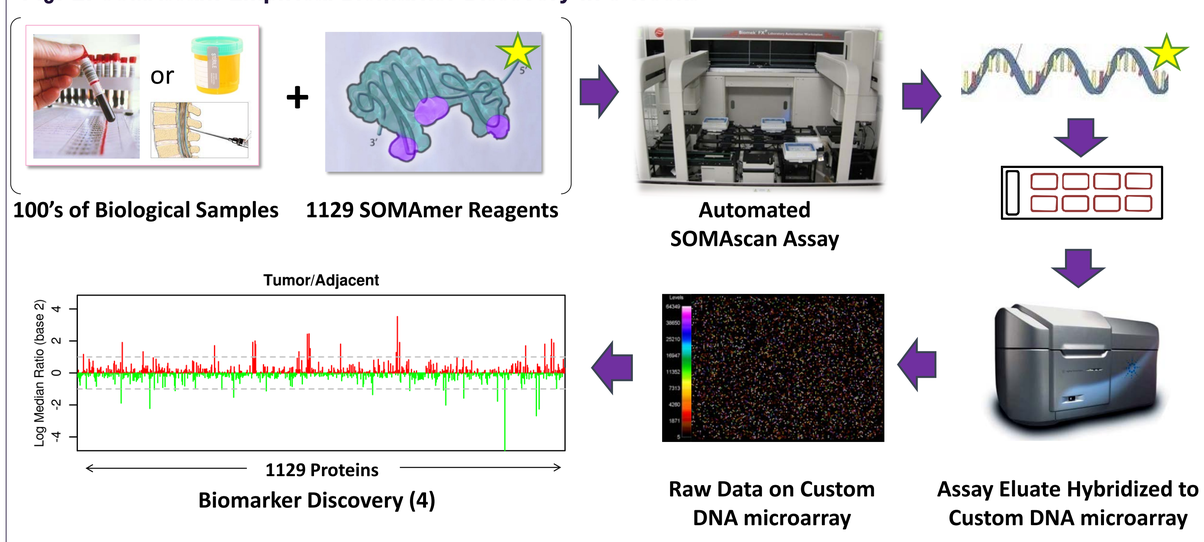


Table 1. SOMAscan™ Assay Overall Performance

Metric	Conditions	Current SOMAscan v3.0
Sensitivity	Median LOD (buffer) Median LLOQ (plasma)	38 fM or 1.6 pg/mL 100 fM
Precision	Median Total %CV	5%
Dynamic Range	Over all proteins in serum or plasma Median individual SOMAmer	10^7 10^4
Sample Volume	Per sample	50 µL
Multiplex Size	Current # targets per sample	1129
Time to result	Automated assay & hybridization detection	~30 hours
Throughput	1 robot, 3FTEs (Discovery)	168 samples/day ~1 mil analytes/week (Premium)

References

[1] Gold L, *et al.* (2010). *PLOS One* 5(12):e15004 [2] Keeney T, *et al.* (2009). *Journal of the Association of Lab Automation* 14:360–366 [3] Thambisetty et al., (2010b). Arch Gen Psychiatry. [4] Soares et al., (2012). Arch Neurol. [5] Kiddle et al., (2012). PLoS one. [6] Doecke et al., (2012). Arch Neurol. [7] Akuffo et al., (2008). Biomarkers. [8] Hye et al., (2006). Brain. [9] Thambisetty et al., (2008). J Neurol. [10] Guntert et al., (2010). J Alzheimers Dis. [11] Thambisetty et al., (2010a). J Alzheimers Dis. [12] Choi et al., (2002). Biochem Biophys Res Commun. [13] Cutler et al., (2008). Prot Clin Appl. [14] Henkel et al., (2012). J Neural Transm. [15] Hu et al., (2012). Neurology. [16] Liao et al., (2007). Proteomics Clin Appl. [17] Liu et al., (2006). Dement Geriatr Cogn Discord. [18] Mhyre et al., (2008). Neurobiol Aging. [19] Ray et al., (2007). Nat Med. [20] Ijsselstijn et al., (2011). J Protome Res. [21] O'Bryant et al., (2010). Arch Neurol. [22] Yu et al., (2003). Proteomics. [23] Zhang et al., (2004). Proteomics. [24] Maere et al., (2005). Bioinformatics. [25] Huang et al., (2009). Nat Protoc. [26] Huang et al., (2009). Nucleic Acids Res. [27] Matsubara et al., (1990). Ann Neurol. [28] Lieberman et al., (1995). Neurobiol Aging. [29] Licastro et al., (2000). J Neuroimmunol. [30] DeKosky et al., (2002). Ann Neurol. [31] Brugge et al., (1992). Ann Neurol. [32] Hinds et al., (1994). Neurobiol Aging. [33] Altstiel et al., (1995). Dementia. [34] Mcilroy et al., (2002). Int J Geriatr Psychiatry. [35] Porcellini et al., (2008). Curr Pharm Des. [36] Giometto et al., (1988). Eur Neurol. [37] Maes et al., (2006). Neurobiol Dis. [38] Thambisetty et al., (2011). PLoS one. [39] Merched et al., (2000). Neurobiol Aging. [40] Saczynski et al., (2007). Am J Epidemiol. [41] Kuriyama et al., (2008). Psychiat Clin Neuros. [42] Laske et al., (2009). J Alzheimers Dis. [43] Lorenzl et al., (2003). Neurochem Int. [44] Caramelli et al., (1999). Prot Clin Appl. [45] Schrijvers et al., (2011). JAMA.

A Rule-Out test for Amyotrophic Lateral Sclerosis (ALS) in Subjects Presenting With Neurological Symptoms

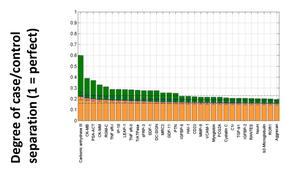
Establishing the diagnosis in people who have the disease is a slow, invasive and costly process, and establishing that an individual with neurological symptoms does **not** have ALS is equally slow, invasive and costly. The most differential expression was seen between samples from ALS patients and healthy controls with an estimated AUC = 0.93 (0.83, 0.97) and 10-fold cross-validation, giving a mean AUC = 0.89 ± 0.02, and an associated sensitivity of 0.89 ± 0.02 and specificity of 0.8 ± 0.03. These biomarkers were specific to ALS as they were not affected by neurodegeneration-based symptom mimics such as MS or AD.

Study design:

Initial discovery of ALS biomarkers (Figs 2**A-D**): 223 patients with early ALS; 100 healthy controls without ALS; 50 neurological disease mimics

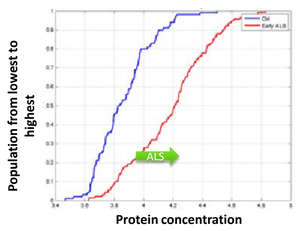
Confirmation of disease specificity of ALS biomarkers (Fig **2E**): 50 subjects with early MS; 50 subjects with neurological symptoms but not MS or ALS; 182 subjects with Alzheimer's disease

Fig. 2A. ALS vs. Healthy Controls

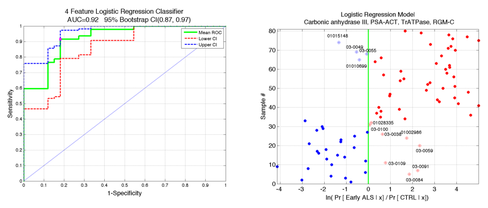


Total degree of separation (green) exceeding the false discovery rate (beige) for each of the top proteins

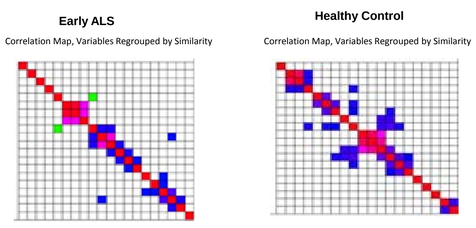
2B. Empirical Cumulative Distribution Functions for a top features used in classifier development



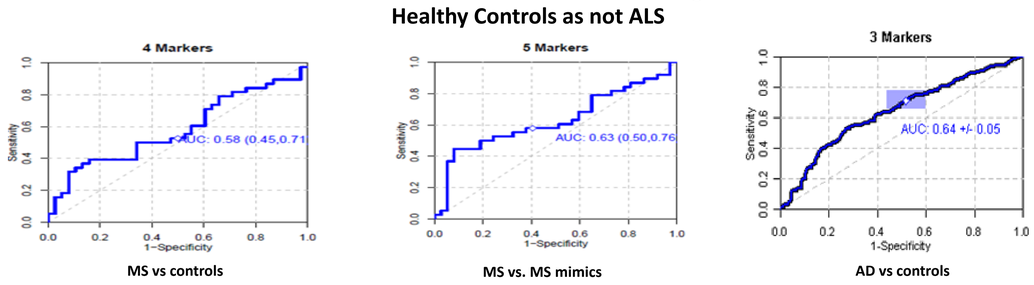
2C. ROC curve (left) and log odds plot (right) generated by four protein logistic regression model applied to the test set



2D. Finding biologic processes in ALS: identifying clusters of correlated proteins

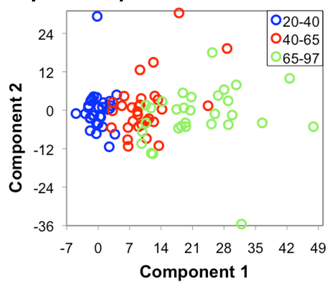


2E. ALS classifier correctly identifies other neurological diseases and Healthy Controls as not ALS

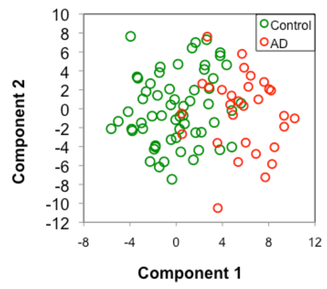


Cerebrospinal Fluid Analysis using SOMAscan™

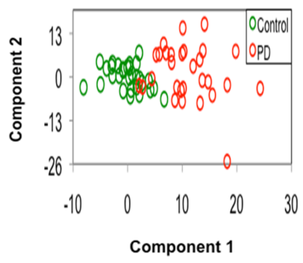
3. Principal components analysis (PCA) of CSF from healthy individuals (20 to 85 years old).



4. PCA of CSF comparing Alzheimer's Disease and age matched healthy controls



5. PCA of CSF comparing Parkinson's Disease and age matched healthy controls



Replication of blood-based protein biomarkers of Alzheimer's disease onset and progression

Protein name	This study (0.1 FDR level)	Literature discovery/panel studies	Literature candidate studies	Total number of independent replications
Alpha-1-antichymotrypsin	+ with AD.	+ with AD [14].	+ with AD [27]. + with AD [28]. + with AD [29]. + with AD [30]. + with AD [31]. + with AD [32]. + with AD [33]. + with AD [34]. + with AD [35].	11
Alpha-1-antitrypsin	+ with AD. - with left EC volume.	Increased oxidisation in AD [12]. + with AD [22]. + with AD [16]. - with brain amyloid burden [5]. + with AD [6].	+ with AD [36]. + with AD [37].	8
Complement C3	- with AD. + with whole brain volume.	+ with AD [23]. Brain amyloid burden [11]. HC volume [3]. + pre-symptomatic AD [20]. - with brain amyloid burden [5].	+ with whole brain volume [38]. + with AD [36].	6
Apolipoprotein A-I	- with AD	fast AD progression [3]. AD and MCI [15]. - with AD [17].	- with AD [39]. - with AD [40]. - with AD [41].	6
Pancreatic prohormone	+ with AD. - with HC volume and left EC volume	+ with AD [21]. + with AD and MCI [4]. - with brain amyloid burden [5]. + with AD [6]. + with AD and MCI [15].	Not found	5
Granulocyte colony-stimulating factor	- with EC volume.	with AD [19]. - with AD [21].	- with AD [42].	4
Matrix metalloproteinase-9	+ with AD.	- with brain amyloid burden [5]. + with AD [6].	+ with AD [43].	4
Apolipoprotein B	- with rate of decline.	+ pre-symptomatic AD [20]. + with AD [23].	+ with AD [44].	4
Insulin-like growth factor-binding protein 2	+ with AD.	+ with AD [21]. + with AD [6].	Not found	3
Clusterin	- with rate of decline.	- with AD [16]. Fast AD progression and HC volume [3].	+ with AD [45].	3
Hemopexin	- with rate of decline.	+ with AD [14]. With AD [22].	Not found	3
Complement C6	+ with AD and left EC volume.	+ with AD [14].	Not found	2
Complement C5	+ with AD.	+ pre-symptomatic AD [20].	Not found	2
Corticosteroid-binding globulin	- with rate of decline.	+ pre-symptomatic AD [20].	Not found	2
Inter-alpha-trypsin inhibitor heavy chain H4	- with left EC volume.	- with AD [8]. - with AD [16].	Not found	2
C-C motif chemokine 18	- with left EC volume.	with AD [19].	Not found	2
Fibronectin	- with AD.	with AD [19].	Not found	2
Creatine kinase MB	- with AD.	- with AD [21].	Not found	2
Superoxide dismutase	+ with AD.	+ with AD [6].	Not found	2
C-C motif chemokine 15	+ with AD.	with AD [19].	Not found	2

As a form of dimensionality reduction we preselected 94 candidate biomarkers from 21 discovery blood proteomics studies of AD and measured them in 687 subjects. Sixteen of the previously reported candidates were found to independently associate at the 0.1 FDR p-value level with AD-related phenotypes, suggesting a panel of proteins that show sufficient replication to be considered for further investigation as valid biomarkers in AD. Processes in AD are color grouped and underscored. (Also presented by Kiddle et al at Biomarkers for Brain Disorders: Challenges and Opportunities in Cambridge, UK)