somalogic™

An Evaluation of SomaScan® Data Reproducibility at Various Sites

Introduction

Proteomics usage has substantially grown in the past decade due to its dynamic nature and direct correlation with clinical outcomes. The SomaScan® Assay has become an increasingly popular platform for measuring thousands of proteins (up to half the human proteome) at once. While samples can be run at SomaLogic headquarters (Boulder, CO, USA), the Authorized Site program has enabled researchers to perform the SomaScan Assay in their own laboratories after they have completed extensive onboarding training and site qualification. This offers added benefits of allowing the SomaScan Assay to be performed in other countries where cost and shipping requirements to the United States may not be advantageous. While the initial qualification of the Authorized Site and continual monitoring of QC metrics gives confidence in the performance of the assay at these sites, a more direct comparison of clinical samples run at both sites offers additional possibilities for robust analysis.

In this technical note, we investigate the performance of full SomaScan Assay data generated at two separate locations, utilizing the same control and clinical samples; this includes examination of specific analytes that may differ between sites and whether predictive models such as SomaSignal® Tests (SSTs) will produce the same result.

We characterized the data between two laboratories, Site A and Site B, using the same samples processed at both sites. The median total CV for both sites was low and below our expectation of ~5% median total CV, which demonstrates consistently low noise. Analytes have excellent concordance, with 97% of the menu showing a correlation coefficient greater than 0.5 (and 82% showing a correlation coefficient ≥ 0.8). To illustrate the utility of predictive models at Authorized Sites, we evaluated over 20 SSTs and concordance was greater than 0.93 in all cases (with most showing a concordance ≥ 0.98). This initial two site study gives confidence in the consistent performance of the SomaScan Assay across sites.

Background

For scientific studies, reducing sources of biological and technical noise is ideal. Noise can cause unexplained variability and generally reduces reproducibility of repeated data points, which often results in an increase in the Coefficient of Variation (CV). External causes of technical noise are optimally reduced as much as possible so that biological variance of an experiment reveals itself in the data clearly. In most instances, variation can be minimized by any number of means such as:

- Isolating, processing, and stored samples consistently and as required by the assay
- Using a well-controlled experimental study
- Consistently performing an assay, utilizing the same instruments and equipment for processing samples and extracting data

These are ideal, though in some cases, not logistically possible. Increasing the sample size for a study has many benefits and is a common method to compensate for inherent noise by increasing study power (SL00000873). Subsequently, normalization methods can be used to mitigate technical variance in the data without obscuring meaningful biological patterns.

Biobanks offer an important source for large numbers of valuable biological samples, for example some population level studies can require tens of thousands of samples. Even so, the physical processing of that many samples becomes a barrier. Even with the ability of the SomaScan Assay to run hundreds of samples a day at a single location, performing the assay at more than one site at a time would speed up the experimental timeline tremendously.

When a data set is processed at multiple facilities, there is always a risk of introducing site-specific bias as different equipment is used and there could be some variance in the sample handling or assay workflow. Though the process for becoming a SomaLogic Authorized Site requires initial qualification of staff and instruments as well as ongoing monitoring of

data which ensures consistent results over time (D0004375), demonstrating the consistency of results between sites requires a direct comparison. Here, we have outlined a process for identifying site-level comparability by examining paired samples processed at two laboratories (Site A and Site B), using both clinical and control samples.

Primary Expectations of Assay Performance

The SomaScan Assay uses a set of established samples and SOMAmer® Reagents as controls with a robust normalization process to mitigate technical variation and reduce assay noise (SL00000517). Metrics like QC Percent In-Tails (accuracy) and QC replicates CV (precision) provide some information about the assay performance and consistency (D0004573). Every SomaScan Assay run include Calibrator Control replicates (for data standardization), QC Control replicates (to ensure that normalization has adjusted variance at an appropriate level), and Buffer Control replicates (to check the background noise of the assay). These are to ensure that biological and clinical samples return accurate and consistent data.

To reduce variability, a study processed at multiple facilities should include the same lot of assay Master Mix, Calibrator Control and QC Control. Minor variations to laboratory conditions, as well as sample shipping or storage, may account for some of the minor differences observed from site to site.

Study Design

For this experiment, commercially purchased EDTA Plasma samples were processed by the SomaLogic Assay Services (Site B) and distributed across 5 different plates, while the samples run at the Authorized Site (Site A) were processed on a single plate. Site B processed the data set on the SomaScan Assay approximately 2 years before Site A. All data has been fully normalized according to SomaLogic's standard procedure for EDTA Plasma (SL00000442). The final analysis compares the Relative Fluorescence Units (RFUs) for 7,289 human SOMAmer Reagents (analytes) from 58 samples. General information regarding the sample number and distribution across plates are included in Table 1 below:

Site Location	Plate ID	Buffer	Calibrator	QC	Sample
Site A	Plate 1	3	5	3	58
Site B	Plate 1	1	5	3	7
	Plate 2	3	5	3	1
	Plate 3	2	5	3	5
	Plate 4	6	5	3	24
	Plate 5	6	5	3	21

TABLE 1 Summary of each plate across the two sites included in the analysis.

Assessment of QC Control Replicates within Site A and Site B

The quality of data generated at Site A and Site B meet the performance expectations as defined by SomaLogic (D0006601). This is guided by the QC Control replicates' accuracy after standardization, or how closely each measurement aligns with an expected reference value, called the QC Percent In-Tails (Table 2). For the QC Percent In-Tails at least 85% of analytes are expected to lie within 20% of the reference value. All plates included in this study are well below this threshold, where the plate from Site A resulted in a 2.01% QC In-Tails, and the 5 plates from Site B ranged from 0.35% to 1.66%.

Site Location	Plate ID	QC Percent In-Tails (%)
Site A	Plate 1	2.01
Site B	Plate 1	0.36
	Plate 2	0.99
	Plate 3	0.58
	Plate 4	1.41
	Plate 5	1.66

TABLE 2 Percentage of analytes where the ratio of the QC measurement deviates greater than 20% of the expected reference values. At least 85% of measurements are expected to lie within the expected range, 0.8-1.2.

Additionally, the precision, or assay reproducibility, is measured by categorizing the Coefficient of Variation (CV) for technical replicates within a plate (intra-plate CV) or across all plates (Total CV). Total median CV, across 5 plates, for Site B is 4.14%, and the total median CV for Site A is 3.52%. Therefore, 50% of analytes have a total CV of less than 4.14% and 3.52%, respectively. This aligns with the expectations of the SomaScan Assay, where the median CV is approximately 5% for v4.1 (SL00000873).

Site Location	Percentile	Intra-plate	Total CV
Site A	50%	3.51	3.51
	90%	7.31	7.31
Site B	50%	2.78	4.13
	90%	5.95	8.76
Total Across Both Sites	50%	2.88	4.49
	90%	5.89	9.18

Differential Analysis of Non-Control Samples by Site

Minor differences were observed between individual measurements of the 58 samples processed across the two sites. Principal Component Analysis (PCA) was performed as a high-level assessment of the differences in data generated at Site A and Site B. Figure 1 shows the first and second principal components colored by site, where the signals in the data are very similar between the two sites. Two samples in the fourth quadrant cluster separately from the rest, indicating that the proteome of these samples are distinct from the others; Although these samples differ, the proteomic signatures are replicated across Sites A and B (examples circled in red in the plot below).

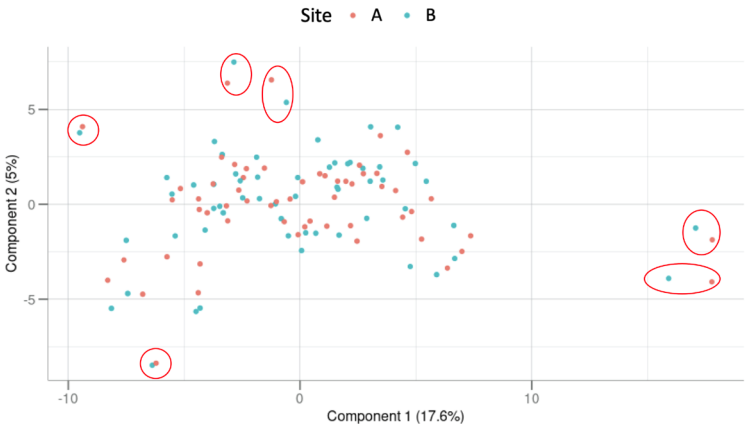

FIGURE 1 Principal component analysis (PCA) plot colored by site. Samples show similar proteomic signatures at each site with clustering of samples processed at each site. Examples of samples that showed distinct characteristics from the population, but that clustered together by site are highlighted in red, showing that performance was similar across both sites.

To further compare consistency of protein measurements between Site A and Site B for 58 samples, we measured the correlation of RFU values in the paired samples measured in the SomaScan Assay for 7,289 human analytes. Correlation was measured multiple ways, including methods by Pearson (strength of the linear relationship), Spearman (rank order), and Lin's CCC (mean difference in univariate ranking). The 10th, 50th (median), and 90th percentiles of correlation

TABLE 3 Intra-Plate CV and Total CV for the 50% and 90%-percentiles for human analytes measured in the QC replicates at Site A and Site B. Median Total CV for both sites together is 4.49%.

coefficients (rho, σ) were calculated across all paired analytes (Table 4). The median correlation was greater than 0.9 for both Pearson and Lin's CCC. The median Spearman correlation was 0.87.

Percentile	Pearson	Spearman	CCC
10%	0.712	0.632	0.675
50%	0.942	0.869	0.927
90%	0.989	0.972	0.986

TABL E 4 Summary statistics of correlation coefficients for 7,289 human analytes.

Investigating Analytes with Low Correlation

There is high reproducibility in the analyte measurements for the paired samples generated at Site A and B, but a small subset of analytes did not correlate well. Internally we explored these features with low correlation coefficients, and focused on whether this might be related to: low signaling or saturated measurements (RFU), SOMAmer Reagent dilution bin (0.005%, 0.05%, or 20%), or technical variance (assay CV).

Of these three factors, low correlation coefficients appear to be weakly related to dilution group and assay CV which is further detailed below.

Impact of Dilution Group on Correlation:

SomaScan Assay to simultaneously measure proteins ranging from 10-logs of concentration in blood, the SOMAmer Reagents are divided into 1 of 3 dilution bins (0.005%, 0.05%, or 20%) and recombined later in the procedure. Human proteins typically observed in lowest abundance (femto- to pico-molar range) in blood are primarily in the 20% (1:5 dilution group). Additionally, the 0.005% (1:20,000), dilution group target proteins in highest abundance (micro-molar range) in blood, and the 0.5% (1:200) dilution group targets proteins in between (nano-molar concentrations).

Figure 2 below is a boxplot of the Pearson correlation coefficient stratified by dilution group. Most analytes have a strong correlation across all three dilution groups (correlation coefficient is greater than 0.8), but the median correlation coefficient for the 0.005% dilution group appears to be somewhat less than the other two groups, 0.5% and 20%. This suggests that the analytes in the highest abundance in the sample correlated slightly worse at the two sites.

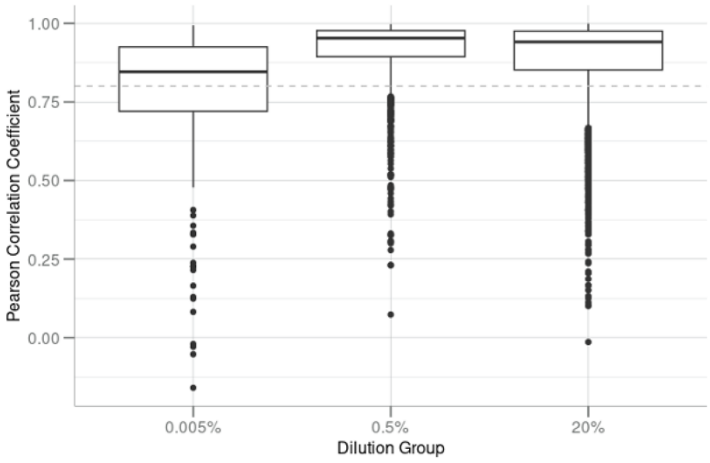

FIGURE 2 Boxplot of Pearson correlation coefficient stratified by dilution group. The dashed line represents σ = 0.8.

We further stratify the analytes into high, medium, and low bins (based on Pearson correlation coefficient) in Table 4. Note that a larger percentage of analytes that are in 'highest' abundance in blood (the 0.005% dilution bin), have a lower correlation than in the other two dilution bins.

Correlation	Dilution Groups			Total N (%)
	0.005%	0.5%	20%	
High (σ ≥ 0.8)	108	966	4,937	6,011 (82.5%)
Medium (0.5 ≤ σ < 0.8)	56	112	941	1,109 (15.2%)
Low (σ < 0.5)	23	22	124	169 (2.3%)

TABLE 5 Stratification of Pearson correlation coefficient (σ) by dilution groups. The final column (Total N (%)) shows the percentage of measurements of 7,289 human targets. 82.4% of analytes have high correlation (where σ is greater than 0.8). 15.2% have medium correlation (σ between 0.5 and 0.8), and 2.4% have low correlation (σ is less than 0.5).

In a follow up experiment, we may explore the linear ranges of each reagent to understand additional factors that influence the correlation, including whether the RFU values that lie within the linear range have higher correlation across sites, or whether changes to the reagent lots, sensitivity or matrix-specific differences have some influence as well.

Impact of Assay CV on Correlation:

SOMAmer Reagent CV is calculated from replicated samples (the three QC replicates and one individual non-control sample) and the calibrator sample (five replicates) across 15 independent assay runs (plates) during assay validation. The total CV for each SOMAmer Reagent has been characterized and defined previously by SomaLogic in the SOMAmer Reagent 'Extended Annotations' to demonstrate assay reproducibility (SL00000571). Here, we investigated whether correlation coefficients are associated with the total assay CV for low correlation analytes. A weak correlation between assay CV and analyte correlations was found where some analytes with higher assay CVs seem to have generally lower correlation coefficient. This illustrates that some SOMAmer Reagents that typically are seen to have higher CVs, also demonstrate lower correlations in this comparison.

'Lifting' or 'Bridging'

Historically, at SomaLogic, "Bridging" or "Lifting" are two terms used to describe a method to shift the RFU values generated in one version of the SomaScan Assay to an RFU value that is closer to what would likely be observed in another version. This transformation is performed by multiplying or dividing the data by a scale factor. Additional details regarding bridging are described in detail in a separate document (SL00000599), but in the context of this site comparison, we used this method to determine whether the data is reproducible from site-to-site.

Our goal is to explore whether an additional scaling, or 'bridge' would be required between Site A and Site B. We found, however, that the data between the two sites were comparable without the need for any bridging. To make this conclusion, we created a reference using the 58 paired samples for potential bridging between Site A and Site B and compared it to the scale factors used to bridge between versions 4.0 (5K platform) and 4.1 (7K platform) of the SomaScan Assay. Figure 3 below shows the distribution of the scale factors needed to bridge between the sites centered around 1, with a much narrower distribution implying a lack of systematic changes between the two sites. Thus, a bridge for the analytes between the two sites is not required, as greater than ~95% of the ratios are between 0.9 and 1.1 (highlighted in Figure 3 with vertical, dotted lines).

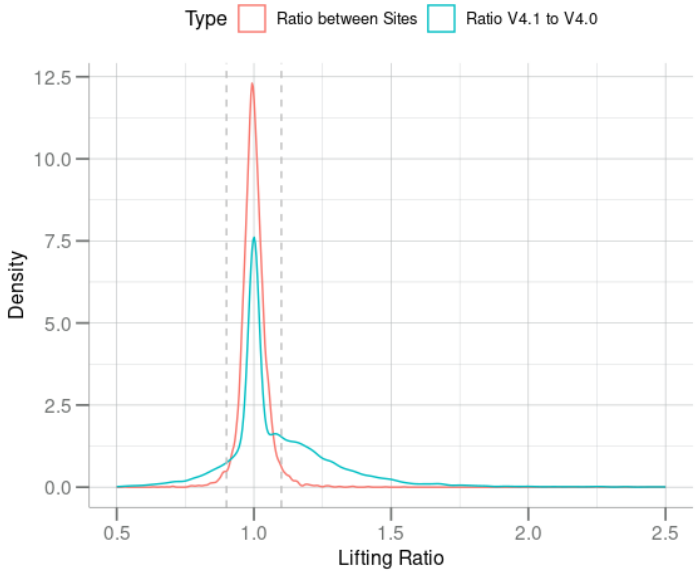

FIGURE 3 Distribution of bridging ratios for Site A and Site B compared to the bridging ratios between assay version v4.0 (5K platform) and v4.1 (7K platform), illustrating that bridging is not necessary.

Model Performance and Clinical Impacts of 'Noise'

SomaLogic's SomaSignal Tests (SSTs) were also used in this comparison to assess if the data generated at Site A and Site B result in different clinical predictions. The data comparison approach was based on simulating the differences (jittering) between the two datasets rather than using the paired samples directly. As each SST was developed for an intended use population, predictions on samples from healthy normal individuals (such as the one used in this analysis) can be overly skewed. Thus, this data comparison approach ensured a comparison between a wide range of predictions that simulate the differences in condition between the two sites.

To obtain the 'jittered' data, we first simulated the noise based on the density of differences in log10 RFU values between the paired samples and added the noise (jitter) to the original training dataset for each SST. The

jittered data should ideally mirror the output of the original (un-jittered) training data.

Then, we apply the SST to obtain a prediction, and use Lin's Concordance Correlation Coefficient (Lin's CCC) to determine whether the differences fall within what may be safely considered as assay variability.

There is a high correlation between the SST predictions on the jittered and training data, where the lowest correlation is 0.93 and 0.94 for Resting Energy Expenditure (REE) and Cardiovascular Disease (CVD) Primary tests (Table 9). Though correlation is slightly lower for these two tests, the impact on clinical result, or prediction, is minimal. Furthermore, the differences in some models that predict pre-analytic variation may have lower correlation if there are differences in sample handling between the sites.

SST Type	Model	Jittered vs Training Concordance CCC
Clinical	REE	0.93
	Primary CVD	0.94
	Alcohol Excess	0.95
	Lean Body Mass	0.96
	Percent Body Fat	0.96
	CKD Prognosis	0.98
	Liver Fat	0.98
	Lung Cancer Risk	0.98
	VO2max	0.98
	Dementia Risk	0.99
	HFpEF (1 yr)	0.99
	HFpEF (6 months)	0.99
	HFrEF (1 yr)	0.99
	HFrEF (6 months)	0.99
	OGTT Status	0.99
	Secondary CVD	0.99
	VAT	1
Pre-Analytic Variation (PAV)	Freeze-Thaw	0.96
	Time-to-Decant	0.98
	Time-to-Freeze	0.99
	Fed-Fasted	0.99
	Time-to-Spin	1

TABLE 6 Lin's CCC correlation between the jittered and training data sets using SSTs.

Conclusion

Results of this well-controlled experiment in which data is generated at two laboratories indicate that it is possible to directly merge or compare data sets. Most of the analytes have excellent concordance, where a small percentage with low correlation may have some relationship to dilution group and pre-determined assay noise.

The SomaScan Assay data generated for this site comparison are well-within the performance expectation in terms of accuracy and precision. Minor deviations in data are present plate-to-plate, but the differences fall within an accepted range of assay noise.

Approximately 98% of the SOMAmer Reagents compared across the 58 samples in this study resulted in good correlation, where the Pearson correlation coefficient is greater than 0.5. Investigation of SSTs showed that concordance is greater than 0.93 in all predictive models, with Resting Energy Expenditure (REE) and Primary CV models being the lowest still had excellent correlations of 0.93 and 0.94 respectively. The impact of the differences of the results at the two sites, from a clinical perspective, is minimal. Furthermore, any differences in the PAV models can likely be attributed to differences in sample handling between the sites.

Overall, these results suggest that most of the analytes have excellent concordance between the two processing sites, though every SOMAmer Reagent may not correlate perfectly. Site specific differences can be exaggerated when laboratory conditions (laboratory specific difference, reagent storage, etc.) and sample processing procedures are not carefully monitored. Site-to-site variation can be mitigated with adequate controls, randomizing samples among sites/plates/cohorts, and proper study design. This study is limited in that we have only compared two sites, there was an approximate two-year gap between the samples being run at the two locations, and one site ran the controls across 5 plates rather than within a single plate. Despite any limitations, we have demonstrated that only a very small percentage of the menu (2%) shows low correlation and deviation between the sites, and there is minimal impact on the SST predictions for clinical or PAV outcomes.

Additional Resources:

- Authorized Site Guidance: Control Monitoring and Performance Metrics (D0004573)
- SomaScan v4.0, v4.1, and v5.0 Data Standardization (SL00000442)
- The SomaScan Assay v4.0, v4.1, 5.0 Stability (SL00000517)
- Filtering, Interpretation and Consideration for Flagged Samples in the SomaScan Assay (D0006601)
- Ensuring Reproducibility in Proteomics: Why Coefficient of Variation is a Critical Metric (SL00000873)
- SomaScan Assay Annotated Content (SL00000571)
- Bridging Between SomaScan Assay Versions (SL00000599)

somalogic™