TECH NOTE

nCounter® Vantage 3D™ Protein Assay Normalization & Data Analysis

Introduction

The nCounter Analysis System utilizes digital barcode technology to measure a variety of nucleic acids, including mRNA and miRNA, with high levels of precision and sensitivity. NanoString has developed a method for protein analysis using DNA oligonucleotides that are covalently attached to primary antibodies specific to proteins of interest. This allows for the ability to quantify both RNA and protein in a single assay, known as 3D Biology Technology. Part of the power of the nCounter Analysis System is that the raw counts directly represent target molecules and not amplified extrapolations. Because small differences in sample input and hybridization conditions can affect the magnitude of counts, it is important that the counts are normalized to these potential sources of technical variance in order to obtain accurate estimates of changes between experimental groups. This tech note is designed to guide you through the best practices for NanoString protein data normalization and analysis.

Background Subtraction and Negative Control Antibody Probes

All of the Vantage 3D Protein Assays include negative antibody controls (e.g. MmAb-IgG1 and MmAb-IgG2A) that will not specifically bind to proteins in your sample. These negative controls are used to determine non-specific background in the assay. As discussed below (see Signal-to-Noise), this information is useful to understand the performance of the assay and it can be used to subtract background counts from the specific protein target counts. It is not, however, generally advisable to perform background subtraction from either these or the Negative Hybridization ERCC spikes*, as this can potentially inflate differences in low expressed protein. As an example, if the background levels are 20 counts from the negative control antibodies and two samples have counts of 21 and 42 (2X difference) for specific target proteins, following background subtraction (-20), the counts will be 1 and 22 (22X difference). At this relatively low level of expression, a 22-fold change likely overemphasizes the biological effects of this relatively small change in protein molecule abundance.

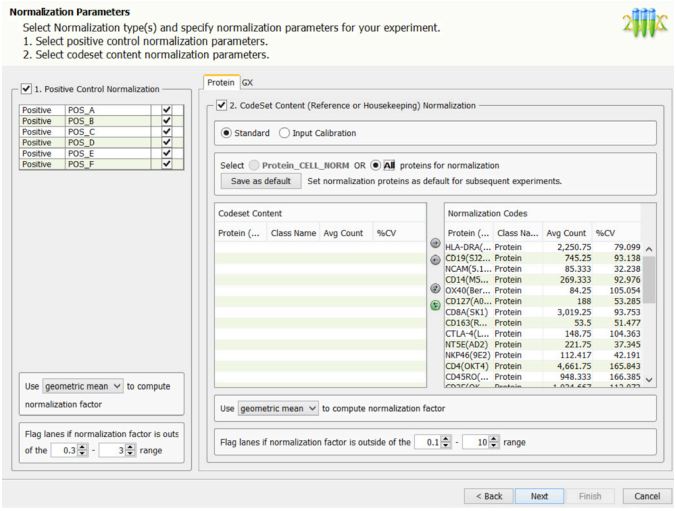
Normalization

Normalizing NanoString data removes technical variance, which allows for discrimination of small fold change between groups (~1.3). Determining the right normalization strategy depends on the type of assay used, the sample types profiled, and the underlying sample biology. Note that strong biology can likely be observed without normalization or by using several different strategies, whereas observation of small biological effects will require more careful considerations of the normalization strategy.

Protein Normalization in nSolver™ Analysis Software

For nSolver analysis, you first have the option of whether to perform "Positive Control Normalization," which uses the ERCC spikes* and is a good control for hybridization efficiency across both RNA and the protein barcodes. Performing this step is recommended. For protein specifically, you next have the options of (1) not performing normalization by deselection of "CodeSet Content Normalization," (2) selecting "Protein-CELL_NORM," or (3) selecting "All proteins for normalization" (Figure 1). In this example experiment using the Vantage 3D Protein Immune Cell Profiling Panel on PBMC, there is not an option for "Protein_CELL_Norm" as no single target has been pre-defined as a stably-expressed protein for normalization. In Figure 1, the geometric mean** of all proteins is selected for normalization and an example of the impact of this type of normalization is seen in Figure 2. In this experiment, there are four groups of biological replicate samples: A, B, C, and D. Groups A and B are unstimulated PBMC while groups C and D are PBMC stimulated with PMA. The impact of normalization on the variability of expression of these markers across the biological replicates within each group (A, B, C, and D) is very clear (Figure 2; Compare "Raw Counts" to "All Protein Geomean Normalized Counts"). While selecting "All Proteins for Normalization" worked very well in this example, it is also possible to assess and remove any proteins from the "Normalization Codes" window (Figure 1) that have levels of variability (%CV) that are outliers to the other probes, which may improve the results if such outliers are present in your data set.

Figure 1.



*RCC Spikes: External RNA Control Consortium (ERCC) spikes are external RNA that are spiked into the samples prior to hybridization to control for technical variation in hybridization efficiencies. Any variation across the samples in hybridization will have an equal effect on both RNA and the DNA oligos that are conjugated to the panel antibodies.

**Geometric Mean: A special type of average where we multiply all numbers together and take a square root (for two numbers), cube root (for three numbers), etc. Geometric mean works better than arithmetic mean (regular average) for normalization purposes while working with highly variable numbers.

Alternate Normalization Strategies

Rather than selecting all proteins, as in the example above, selecting individual or sub-groups of proteins within the panel for normalization can also be a successful approach and can be done in nSolver Analysis Software by moving specific genes in or out of the "Normalization Codes" column (Figure 1). The decision about which proteins are best suited for normalization will depend on sample type and experimental design. For example, if you are working with the same cell type across a series of different stimulations, it might be appropriate to select a single protein or a few proteins that are expected to be stably expressed across all the tested conditions. Taking this a step further, if the expectation is that the stimulation will have an effect specifically on the phosphorylation status of various proteins in the panel, it would be reasonable to normalize to the total proteins in the panel as their concentrations would not be predicted to change.

In the Vantage 3D Protein Solid Tumor Assay for both lysates and FFPE, for example, Histone H3 has been pre-defined as a normalization protein, so it will automatically be used if "Protein-CELL_NORM" is selected. In this assay with this analysis setting, Histone H3 is the only protein that will be used to normalize the sample and may be used as a default approach and an alternative to the geometric mean of all proteins. For the Vantage 3D Protein Immune Cell Profiling Panel, CD45, although not pre-defined, could also be used as a single protein for normalization.

No matter the normalization method you choose, including biological replicates in your initial experimental design will be helpful in determining which method does the best job of reducing within replicate variability.

Signal-to-Noise

The negative antibody controls (i.e. MmAb-IgG1 and MmAb-IgG2A) included in all Vantage 3D Protein Assays are used to determine non-specific background in the assay. Even after extensive post-antibody incubation washes, there will be some level of background, which must be taken into consideration when assessing the quality of results from the specific antibodies in your Vantage 3D Protein Assay. Signal-to-noise is determined by dividing the counts from each specific antibody by the geometric mean** of the two negative controls within the same sample. These calculations can be easily performed by exporting the data to Excel. As a good starting point, as long as the signal-to-noise is above 3, the data can be considered comfortably above background. This does not, however, mean that values below 3 are not to be trusted. Consistency across replicates will increase your confidence even in these low abundant targets though further orthogonal validation would be suggested in this case.

Figure 2.

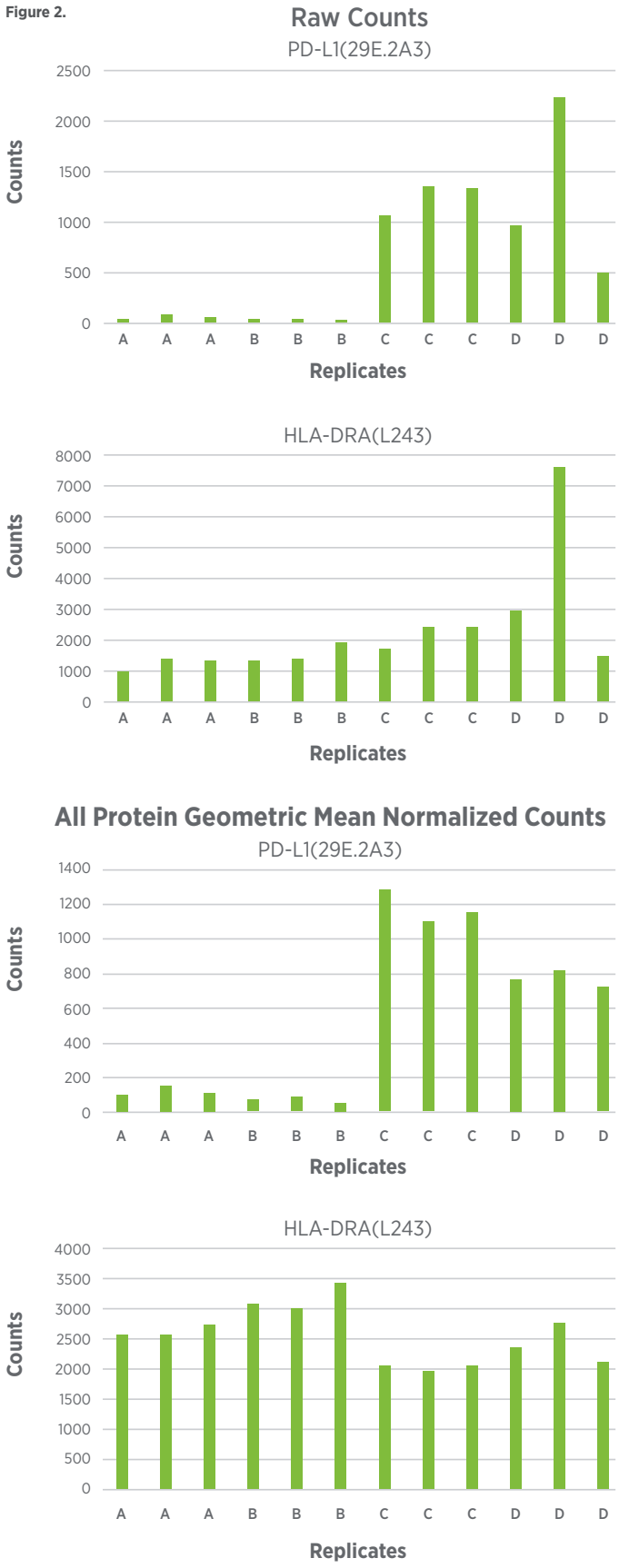


Table 1 shows the signal-to-noise values for the PBMC data described in the section on Normalization. There are two factors that are demonstrated by this data. First, the vast majority of the signal is at least 3X above noise, indicating a technically successful experiment. Second, by highlighting values above 3, it is clear the effect that stimulation has on a number of markers, thus providing confidence in the biological response to stimuli.

Learn More

This tech note is designed to provide an introduction to protein data analysis and normalization on the nCounter platform. Please contact **support@nanostring.com** for additional questions.

To learn more about our nSolver Analysis Software, visit **nanostring.com/products/nSolver**

Table 1.

	Unstimulated						Stimulated					
Protein	A	A	A	B	B	B	C	C	C	D	D	D
4-1BB(4B4-1)	2	2	1	1	1	1	27	16	15	13	13	12
BTLA(MIH26)	2	2	3	1	1	1	3	2	2	1	1	1
CD127(A019D5)	11	9	9	6	4	5	17	11	12	9	10	7
CD14(M5E2)	41	33	37	17	13	19	2	2	2	1	2	2
CD163(RM3/1)	4	3	4	1	1	1	8	5	4	2	2	2
CD19(SJ25C1)	21	15	17	12	9	10	98	52	52	45	54	39
CD27(M-T271)	55	43	46	20	15	22	500	265	255	150	181	134
CD28(CD28.2)	12	9	10	4	3	3	75	43	41	20	25	19
CD33(WM53)	46	36	42	29	22	31	6	3	3	1	1	1
CD3E(OKT3)	18	14	13	110	75	103	4	3	3	39	52	34
CD4(OKT4)	23	21	26	134	90	126	51	29	29	466	561	424
CD40(HB14)	28	23	27	8	7	10	212	120	114	60	69	56
CD40L(24-31)	1	1	1	1	1	1	3	2	1	1	1	2
CD45(HI30)	239	193	226	143	106	144	486	293	279	254	331	222
CD45RO(UCHL1)	1	1	2	30	20	29	3	2	2	101	112	96
CD68(Y1/82A)	2	2	2	1	1	1	3	2	1	1	1	2
CD8A(SK1)	78	64	73	48	34	48	365	206	198	182	223	171
CD9(HI9a)	790	618	739	154	119	156	170	95	85	20	25	21
CTLA-4(L3D10)	2	2	3	0	1	1	24	13	12	11	10	9
GITR(621)	3	2	2	1	1	1	3	2	2	9	11	9
HLA-DRA(L243)	88	75	86	63	50	62	190	110	110	123	158	113
ICOS(C398.4A)	7	6	5	6	4	5	1520	814	785	879	1094	797
IL2R(M-A251)	4	3	2	1	1	1	626	323	313	261	313	235
KIR3DL1(DX9)	4	4	3	2	2	2	5	3	2	2	2	2
NCAM(5.1H11)	8	7	6	3	2	2	8	5	6	2	2	4
NKP46(9E2)	7	5	7	7	5	6	8	5	4	3	4	3
NT5E(AD2)	26	20	23	7	6	7	16	10	11	4	5	13
OX40(Ber-ACT35)	4	2	3	1	0	0	19	11	12	2	2	2
PD-1(EH12.2H7)	3	2	3	1	1	1	57	29	26	26	29	22
PD-L1(29E.2A3)	4	5	4	2	2	1	120	62	61	40	47	39
PD-L2(MIH18)	2	2	2	1	1	1	9	6	5	3	3	3

For more information, visit nanostring.com

NanoString Technologies, Inc.
530 Fairview Avenue North
Seattle, Washington 98109

T (888) 358-6266
F (206) 378-6288

nanostring.com
info@nanostring.com

Sales Contacts
United States **us.sales@nanostring.com**
EMEA: **europe.sales@nanostring.com**
Asia Pacific & Japan **apac.sales@nanostring.com**
Other Regions **info@nanostring.com**

SEP 2019 MK1834