Article 4

From the supplied excel file “ROS.xlsx” it is evident that the authors have transformed the geometric means from the channel labelled “CFSE” from the raw flow cytometry files to the natural logarithm (evidenced in column B and D of the aforementioned excel file). This is not stated in the methodology and the reasoning behind this transformation is unclear. The summary data plot in figure 4 states “relative fluorescence activity” on the y-axis, however this is misleading as the natural log transformed geometric means are plotted without any relative transformations to control values.

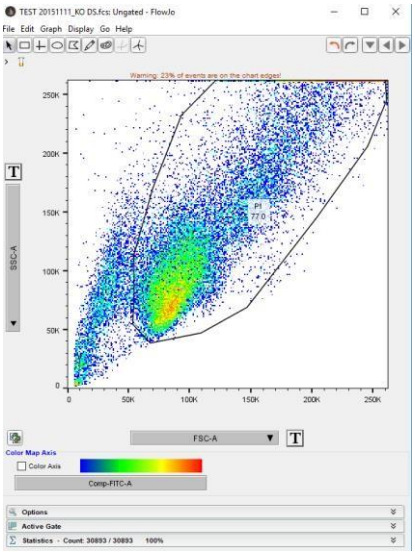
The authors have provided a single raw flow cytometry file for each condition represented in figure 4, rather than the triplicate data files that comprise the summary data. Whilst these files generally match the fluorescence intensities displayed in the representative plots of figure 4 there are some unexplained anomalies. Firstly, the representative plot for the KYSE140 line “UK5099 stain” is identical to that presented for KYSE450 line “control stain”. Secondly, the counts represented in the histogram account for approximately 1/5th of all the acquired events for each sample, even after debris have been excluded on the FSC/SSC, therefore how these histograms were stratified from the raw data is unexplained. Thirdly, the channel used for calculating the DCFH-DA fluorescence is labelled “CFSE” in the raw flow cytometry files. The emission wavelength of CFSE overlaps with that of DCFH-DA, therefore the benefit of the doubt can be given that the authors have mislabelled the correct channel with CFSE.

Metadata for each flow file is embedded in the .FCS file. This revealed the samples provided by the authors were acquired on an LSRII model cytometry rather than the “BD FACSCalibur flow cytometer” stated in the material and methods of the publication. The raw files were also acquired on the 31st August 2016, two days before the publication was received by the journal, leaving little time for thorough analysis and co-author review.

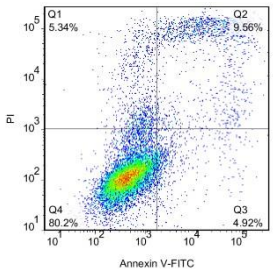
Article 5 Figure

4

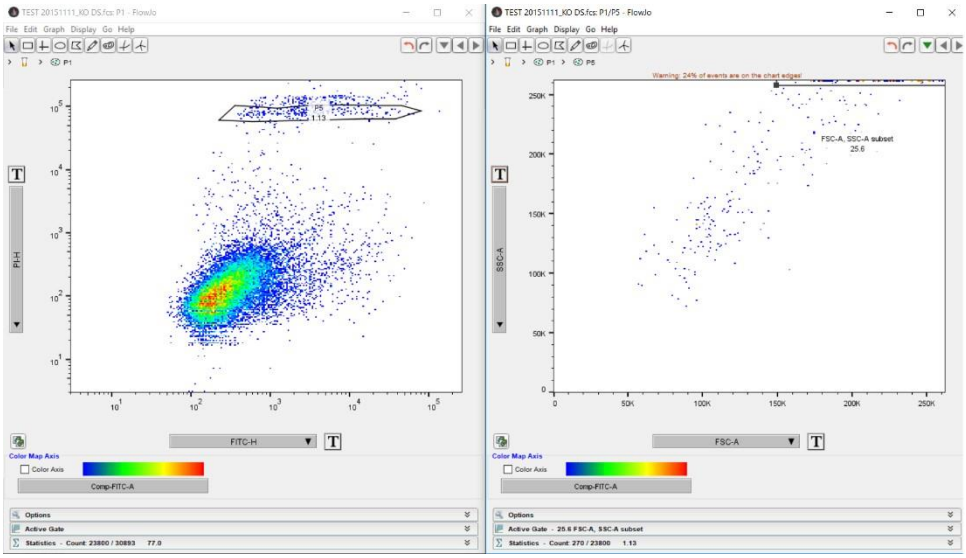
Panel C. Firstly, the voltages are not appropriately set during acquisition of the flow cytometry files.



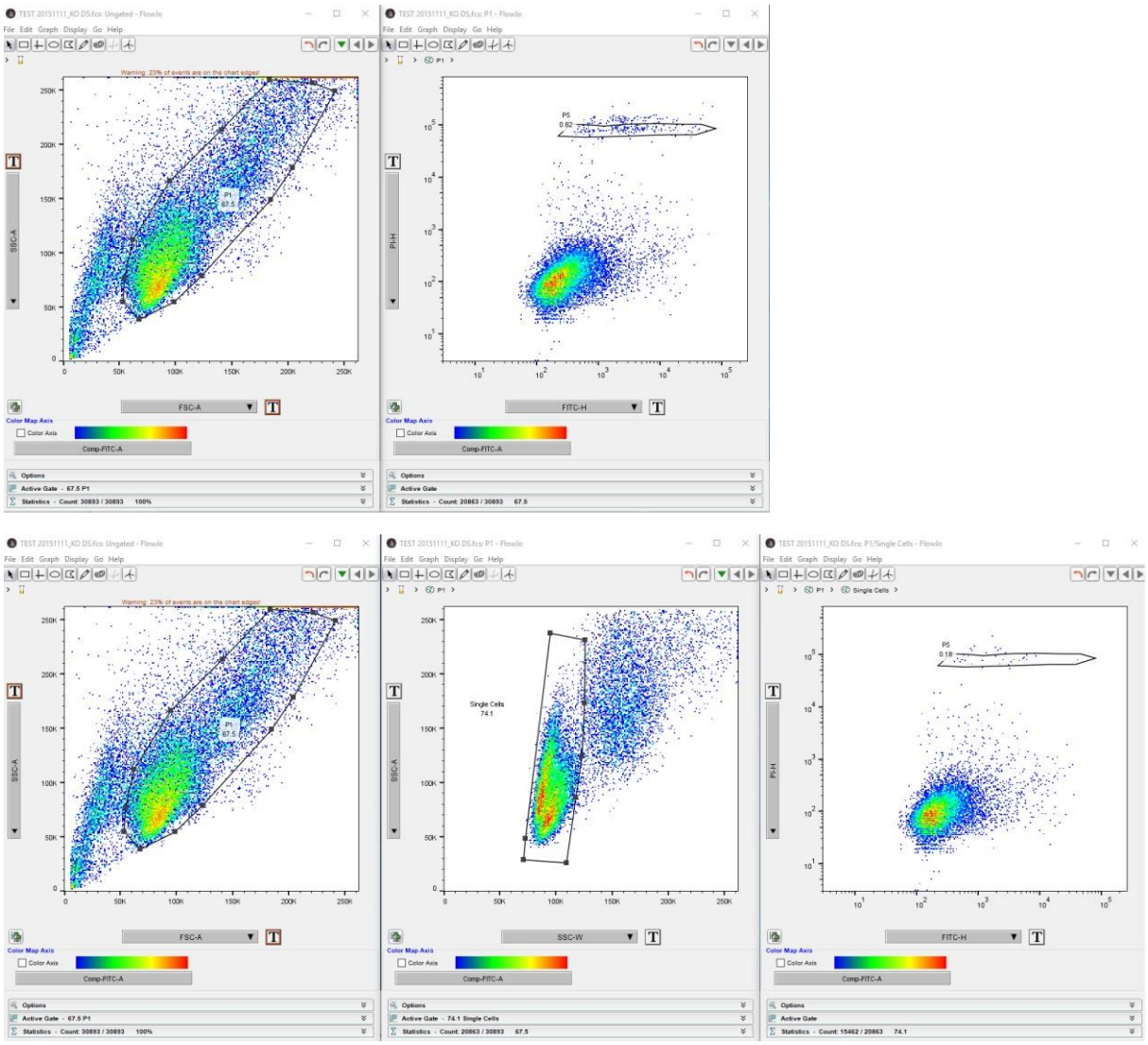
As illustrated above from the raw file labelled “TEST 20151111_KO DS.fcs” 23% of the events are located on the chart edge, due to improper setting of the SSC channel in particular.



This is also evident with the fluorescence signal observed in the PI channel in the provided FACS plots, demonstrated above. Please note the atypical linear pattern of expression for PI as Annexin VFITC increases most prominent in the Q2 quadrant.



Note backgating of these cells using the RAW files provided (shown above) demonstrates a significant proportion of these cells reside outside the SSC/FSC border of the plot. Further, analysis of the raw facs files demonstrated approximately 25% of total events (following debris exclusion) were likely doublets (Illustrated below, in lower central panel, based on SSC-A/SSC-W parameters as these were the only ones available). Note the significant loss of PI high expression cells (demonstrated by the P5 gate below) following doublet exclusion. Together, it is apparent that quantifying fluorescence intensities for these plots is therefore highly inaccurate.



Secondarily, several of the RAW flow cytometry files supplied appear corrupt and cannot be opened ("TEST 20151123_KO G DS2.fcs" for example). Again only one data file is supplied of the 3 replicates used in the summary data.

The excel file labelled "ROS fra [redacted] i nov 2019.xlsx" containing the summary data for figure 4E does not match the presented data in the publication for the KO group (supplied data on the left, publication illustration on the right):

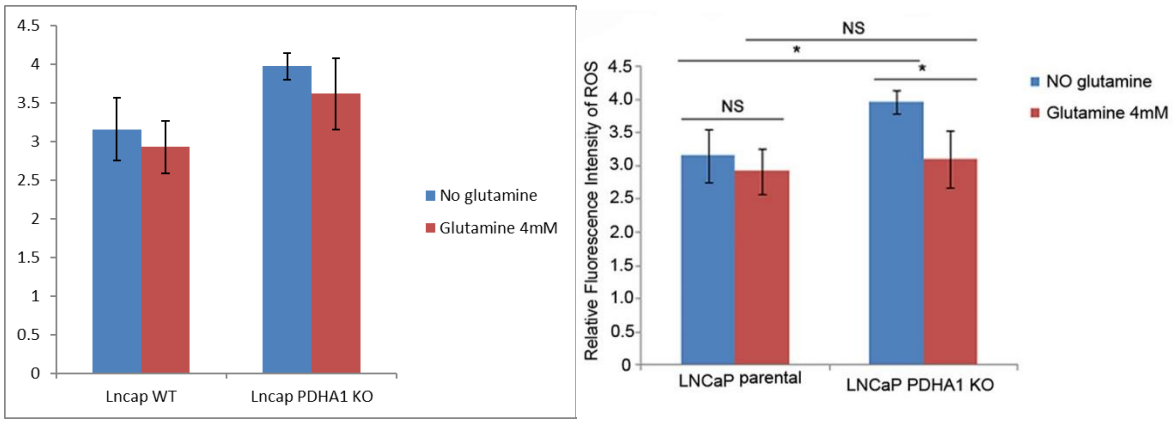


Figure 5

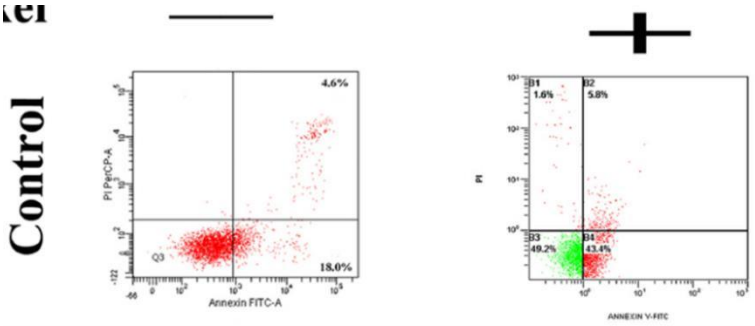
Of the 54 raw flow cytometry files that should comprise the data presented in figure 5, the authors have provided 15 files. The files provided have not been distinguished as to whether they belong to panel B, C or D, therefore the value of assessing these files is limited. The representative plot for LNCap PDHA1 KO BPTES 10uM group is different to the other presented plots. The position of the gates is altered.

From the representative plots displayed for panel 5D, it appears EGCG results in a very modest increase in ROS, compared to the marked increase induced by BPTES. However, the summary plots depict a significant increase that is comparable for BPTES to EGCG. Clearly the representative plots do not reflect the summary bar graph data. The excel file provided later by the authors (file name "apoptosis inhibitors.xlsx") includes bar graphs for figures 5B and 5C, however omits the data for 5D.

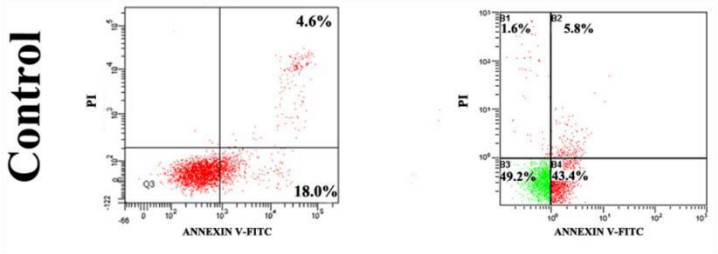
Article 9

Figure 4

The published version of this article contains a duplicate representative plot for conditions LNCaP + docetaxel and PC-3 +docetaxel. This figure has been modified by the author file named "Figure 4.jpg" (path: A9_Oncotarget_2015\Fra [redacted]\Nye data fra [redacted]) such that the plot for PC-3 +docetaxel has been replaced and the quadrant data altered.



^ Originally published Figure 4



^ updated Figure 4 (supplied by author)

The representative plots for each cell line show inconsistencies between the docetaxel – and + groups.

Firstly, the plots appear to have been captured on different instruments at different times. Note firstly the style of the presented plots in the originally published figure, the font, axis legend proximity to the x and y-axis and the abbreviations for the quadrants (Q1 vs B1) and channel name are all distinctly different. Further, the axis is not at all consistent with the docetaxel negative group beginning at -66/-122 and compressing the 1st decade (10^0), whilst the docetaxel positive group axis begins at 0 and includes the 1st decade in a fixed logarithmic scale. Some attempts to improve these inconsistencies are observed in the updated figure 4 provided by the authors upon request, note the now standardized font and axis legend text.

As the authors cannot demonstrate that the docetaxel negative and positive groups were captured in the same experiment, due to an inability to provide the raw flow cytometry capture data, it is not appropriate to compare docetaxel positive values to the corresponding negative control group. A single cytometer, FACSCalibur, is stated as the sole cytometer utilized in this publication, which appears unlikely.

No isotype control plots were provided, therefore it is also impossible to determine whether the gating strategy between these experiments is appropriate. Note the variation in signal intensities between the docetaxel negative and positive groups and the several log scale difference in the placement of the quadrants.

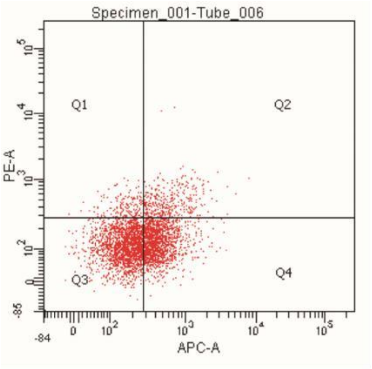
Figure 7:

- The representative plots in **A** do not match the data used in summary graphs in **B** and **C**. This is also highlighted by the author in excel file “OMS Figure 7 图 B-G-----CD44 ABCG2 数据.xlsx”

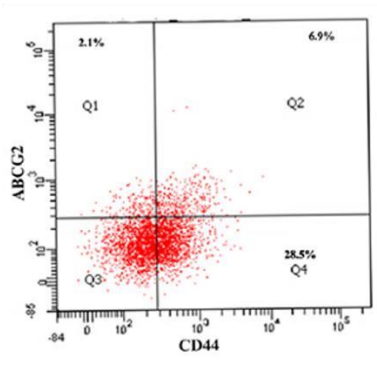
stated in cell B3: “OMS: Grønt felt betyr at verdien stemmer med eksemplene i figuren. Det mangler altså noen”. This draws the following questions:

1) Why was the data presented in the representative plots, for example for PC-3 data for IL-6, IL-11 and IL-24 conditions, not included in the analysis?
2) Why are representative plots from the data set used to analyse these conditions not included?

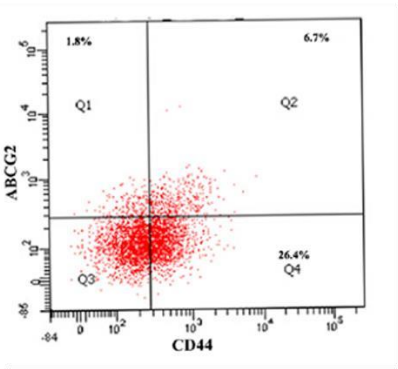
- Concerning additional supplied file, File labelled PC3 IL-24 (filepath: A9_Oncotarget_2015/Fra%20Yu/Figur%207/Fig7CD44ABCG2 文章图/PC3/CD44-ABCG2-PC3-IL24.pdf)



File labelled PC3 IL-24 (filepath: A9_Oncotarget_2015/Fra%20Yu/Figur%207/Fig7CD44ABCG2文章图/PC3/CD44-ABCG2-PC3-IL24.pdf)



Matches PC3 IL-10 illustration from figure 7A of article 9



Matches PC3 IL-24 illustration from figure 7A of article 9

The supplied file matches both panels presented in figure 7A for PC-3 IL-10 and IL-24. Note the quadrant percentages have been manually superimposed onto the plots from the original data file. Whilst these values comply with the values displayed for PC-3 IL-24 condition, they do not match the values for PC-3 IL-10. Further investigation reveals that the file named “CD44-ABCG2-PC3-IL10.pdf” (located in path: “A9_Oncotarget_2015/Fra%20Yu/Figur%207/Fig7CD44ABCG2 文章图/PC3/CD44-ABCG2-PC3-IL10.pdf”) supplied later by the authors bares similarities, but importantly differences to the presented quadrant percentages from the representative blot for PC-3 IL-10. Firstly, this plot does not appear in the representative images for figure 7. Second, the values that comprise the quadrants match only for quadrants 1 & 4. For quadrant 2 the supplied .pdf file for tube PC-3 IL-10 reports a percentage of 6.3, which does not comply with the value of 7.8% present in the corresponding excel spreadsheet (see following figure). Indeed a value of 6.3% does not appear in the PC-3 area of the spreadsheet but does appear as the 3rd replicate of the LNCaP IL-10 values. This rules out the possibility that the triplicate values for Q1, Q2, Q4 were simply not matched in order. It should be noted that alterations to the gating strategy that may have generated a value of 7.8% would also have altered the quadrant percentages for the other quadrants, therefore this value was selectively altered. The correct value of 6.3% would have increased the difference observed with IL-10 stimulation, therefore there is no obvious gain for the authors to have intentionally falsified this result as it already supports their discussion point claims of IL-10 stimulation reducing tumour promotion and stemness.

Tube: Tube_004

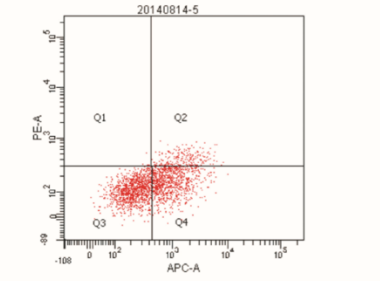
Population	#Events	%Parent	%Total
All Events	11,942	####	100.0
P1	9,682	81.1	81.1
Q1	203	2.1	2.7
Q2	610	6.3	5.9
Q3	6,110	63.1	44.9
Q4	2,759	28.5	27.7

PC-3

OMS: Grønt felt betyr at verdien stemmer med eksemplene i figuren. Det mangler altså noen

PC-3	CD44			co-			ABCG2												
KB	34.6	33.4	35.3	8.9	6.5	9.8	2.5	2.5	2.6										
IL-3	37.6	39.5	38.6	9.8	7.2	11.7	3.1	2.9	3.1										
IL-6	50	60.7	57.8	12.9	10.6	14.1	4.4	3.6	4.9										
IL-10	30.1	26.1	28.5	7.3	5.8	7.8	2.2	2.2	2.1										
IL-11	49.5	56.2	46.8	11.9	9.2	14.9	4.3	3.4	4.5										
IL-24	30.1	21.7	27.4	6.5	5.6	8.1	1.8	2.1	2.2										
PC-3	CD44			平均值	抑制率	P值	co-			平均值	P值	ABCG2				平均值	抑制率	P值	
KB	1	1	1	1	0		1	1	1	1		1	1	1	1	1	0		
IL-3	1.086705	1.182635	1.093484	1.120941	0.120941	0.017	1.101124	1.107692	1.193878	1.134231	0.011	1.24	1.16	1.192308	1.197436	0.197436	0.001		
IL-6	1.445087	1.817365	1.637394	1.633282	0.633282	0.004	1.449438	1.630769	1.479592	1.519933	0.001	1.76	1.44	1.884615	1.694872	0.694872	0.006		
IL-10	0.869942	0.781437	0.807365	0.819582	-0.18042	0.002	0.820225	0.892308	0.795918	0.83615	0.005	0.88	0.88	0.807692	0.855897	-0.1441	0.004		
IL-11	1.430636	1.682635	1.325779	1.479683	0.479683	0.011	1.337079	1.415385	1.520408	1.42429	0.001	1.72	1.36	1.730769	1.60359	0.60359	0.008		
IL-24	0.869942	0.649701	0.776204	0.765282	-0.23472	0.021	0.730337	0.861538	0.826531	0.806135	0.008	0.72	0.84	0.846154	0.802051	-0.19795	0.009		

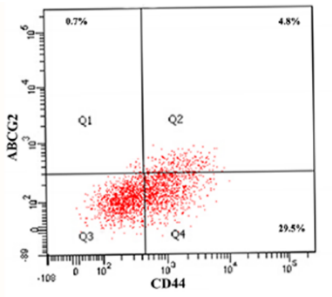
Similar to the mis-reported PC-3 representative plots, the plots for LNCaP following IL-10 and IL-24 stimulation are identical in figure 7A and match the supplied file for LNCaP IL-10. The supplied file for LNCaP IL-24 (path: "A9_Oncotarget_2015/Fra%20[redacted]/Figur%207/Fig7CD44ABCG2 文章图/LN/CD44-ABCG2-LNCaP-IL24.pdf") does however match the quadrant data on the representative plot for LNCaP IL-24 presented in Figure 7A.



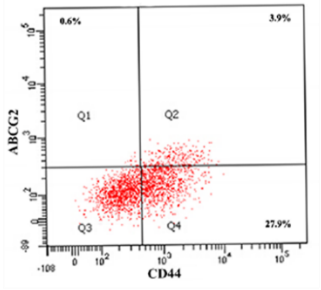
Experiment Name: Y DD 20140806
Specimen Name: 20140814
Tube Name: 5
Record Date: Aug 14, 2014 3:58:09 PM
$OP: Administrator
GUID: e1c88697-c0a7-40ff-8340-49a2c8292c75

Population	#Events	%Parent	FSC-A Mean	SSC-A Mean
All Events	9,030	####	109,908	45,749
P1	6,974	77.2	121,998	49,692
Q1	49	0.7	144,561	77,169
Q2	335	4.8	212,423	118,938
Q3	4,533	65.0	104,462	35,600
Q4	2,057	29.5	126,773	54,140
P2	52	0.7	174,483	104,105
P3	0	0.0	####	####

File labelled CD44-ABCG2-LNCaP-IL10.pdf (filepath: A9_Oncotarget_2015\Fra Yu\Figur 7\Fig7CD44ABCG2文章图\LN)



Matches LNCaP IL-10 illustration from figure 7A of article 9



Matches LNCaP IL-24 illustration from figure 7A of article 9

The omission of quadrant 3 (Q3) data, pertaining to the "double negative" population, from the excel spreadsheet would also have been informative and provided the researchers with a simply means to check the accuracy of their data input. With the inclusion of Q3 data, it is possible to summate the percentages of all populations and compare this to the only possible total of 100%.

Gating strategy should be determined by the Isotype control staining of control samples, which is not displayed. The gating strategy for placing quadrants is markedly different for the representative plot

for PC-3 IL-3 for the APC channel (measuring CD44) compared to the other plots. Do the authors have grounds for moving the quadrant gates based on the APC isotype control fluorescence signal for this experiment? Further, the authors state in the methodological section under “Flow cytometry analysis” that “Viable and single cells were gated for each sample before examination”. There is no evidence presented that supports that such gating occurred, indeed in the supplied facs plots it is evident that only a single gate named “P1” was placed, based solely on the forward and side scatter (FSC/SSC) profile of all events that neither stratifies for viable or single cells.

Facs files are not provided for figure 7, therefore the validity of the samples not provided cannot be assessed. However, the provided .pdf files evidence multiple examples of inaccuracies including incorrect extrapolation of quadrant data and mis-representation of facs plots.