

Immunome™ Protein Array

Protocol

April 2017

A. WETLAB PROTOCOL

1. Before Starting

1. Pour approximately 200 mL cold Serum Albumin Buffer (SAB) into a slide trough/dish and keep at 4°C until required.
2. Pour 2 litres of SAB into a 2 L Duran bottle and equilibrate to 20°C for 30 min in a designated circulating water bath. Fix a 50 mL laboratory dispenser to the bottle before use.
3. Pour 1 litre of SAB into a 1 L Duran bottle and equilibrate to 20°C for 30 min in a designated circulating water bath. Fix a 5 mL laboratory dispenser to the bottle before use.
4. Label the 15 mL Falcon tubes from number 1 up to 24 and place in order in a polystyrene tube rack.
5. Pipette 4.5 mL of SAB into each tube using the designated 5 mL Eppendorf laboratory dispenser.
6. Label the Quadriperm plates with consecutive numbers corresponding to the number of samples (up to 24) on the bottom of the plates.
7. Place 24 Pap jars into a suitable rack and label each tube with consecutive numbers corresponding to the number of samples (up to 24).

2. Sample Dilution Preparation

1. Check each sample visually to ensure that each of the tubes has sufficient serum for assay. Place in the Jeio Tech shaking incubator set at 20°C to thaw for 30 mins.
2. Then, vortex mix each sample for a count of three at full speed.
3. Centrifuge for 3 minutes at 13,000 g. Disinfect the centrifuge with 70% ethanol if spillages occur.
4. Dilute the serum/plasma in SAB buffer to provide the assay solutions (1:800 dilution). Take the first sample, call out the sample ID on the tube, and pipette 5.63 µL of the sample into 4.5 mL of SAB Buffer in tube number 1.

Note: Teammate to write down the sample ID in batch record sheets. This step must be carried out in a Class II Biological Safety Cabinet. Be careful not to disturb the undiluted serum samples during movement.

5. Vortex to mix for a count of three at maximum speed. Place the vortexed tube in a different tube rack to avoid confusion with unused buffer tubes.
6. Repeat step 6 and 7 for the remaining samples.

Note: Teammate to check that the correct samples are added to the correct tubes and mark the batch records accordingly.

3. Incubation with diluted samples

1. Take out the slide dish and rack containing 200 mL cold SAB.
2. Randomly pick 12 pap jars containing 2 protein array slides.
3. Take each slide in turn from their storage buffer by gripping the array between thumb and index finger at the labeled end of the slide.
4. Drain excess liquid from the slide by touching the edge of the array on the rim of the pap jar.
5. Lift the rack from the slide dish and place the slide in slot 2 with the barcoded side facing towards slot 1. Then place the rack back in the slide dish.
6. Add each slide in turn to the rack from left to right, making sure the slides are all in the same orientation.
7. When all the arrays have been added, gently shake the rack up and down five times to aid mixing at the slide buffer interface.
8. Put the lid on the slide dish and shake on an orbital shaker at 50 rpm, for 5 minutes.

Note: Washing times exceed 5 minutes are not critical.

9. Pipette 4.0 mL of each diluted sample from the 15 mL Falcon tubes into their corresponding numbered Quadriperm chambers.
10. Place several layers of white laboratory tissue onto the bench surface and cover this with three layers of Kimwipe tissues.
11. When the slides have finished washing, remove each array in order from the rack.

Note: Always place the rack back in the buffer between removing each array.

12. Grip the slide at the labeled end between index finger and thumb and wipe the back of the slide once on the Kimwipe tissue. Then blot the long edge of the array three times on the wad of lint free tissue paper.
13. Place the slide immediately into Quadriperm chamber containing the diluted serum sample. Ensure that the array does not rest on the lugs at the bottom of the plate.
14. Set a timer to countdown for 2 hours after addition of the first array. Gently swirl the plate to cover the slide with incubation solution.
15. After addition of all slides, scan the barcode on each slide and it will automatically log into the relevant batch record.
16. Incubate all the Quadriperm dishes in the shaking incubator at 50 rpm, 20°C for 2 hours.

Note: Ensure that the dishes are kept horizontal at all times to prevent slopping of solutions between chambers. Handle the dishes very gently to prevent slopping or splashing of contents between chambers.

4. Washing after Sample Incubation

1. Towards the end of the incubation period, fill 24 Pap jars with 30 mL of SAB.
2. When the incubation time has finished, remove each array from the Quadriperm dishes and wash each array individually in a Pap jar.
3. **Wash 1:** Use forceps or index finger to lift the array at the numbered side of the dish. Grip the labeled end of the slide between finger and thumb and place the first slide into its respective number of Pap jar.

Note: Make sure that your gloves are scrupulously clean before handling arrays directly.

4. Cap the Pap jar and invert four times before placing in order in the Pap jar rack on the shaker and shake at 50 rpm.
5. Start a timer to countdown 20 min after addition of the **first array**.
6. Process the remaining slides in order and place each in the Pap jar rack on the shaker whilst shaking at 50 rpm as they are prepared.

7. **Wash 2:** After the 20 minutes of incubation has finished, take the first array and pour off the wash solution into an empty beaker then dispense another 30 mL of SAB into the tube **at the back of the array**. Invert the Pap jar four times and place in the Pap jar rack on the shaker at 50 rpm. Start the timer to countdown 20 min.
8. **Wash 3:** When the second wash step is nearly finished, prepare a slide staining box with rack and add 200 mL of SAB. When the second washing has finished, take the first Pap jar and pour off the buffer. Take the array between index finger and thumb and place in the slot 2 of the slide rack with the barcoded side facing towards slot 1. Place the rack back in the SAB.
9. Start the time to count down 20 min and add the remaining arrays sequentially until all slides have been transferred. Ensure the slides are all in the same orientation and order. Replace the slide rack in buffer between the additions of each array.
10. When all the arrays have been added, gently shake the slide rack up and down five times to aid mixing. Place the lidded box on a shaker for the remainder of the incubation time at 50 rpm.

5. Incubation with Cy3-Anti-Human IgG

1. Measure 200 mL of **SAB** at 20°C into a volumetric flask and add 200 µL of Cy3-antihuman IgG. Mix well by repeated inversion. Pour the solution into a fresh slide staining box (without the rack) and cover until required.
2. Place a wad of Kimwipe tissues on top of the working bench. Ensure that the benches do not become contaminated with buffer.
3. After the third wash is finished, lift the rack of arrays from the wash solution and place on the wad of Kimwipe tissues.
4. Bang the slide rack gently on the tissues five times to remove excess wash buffer. Immediately place the arrays in the Cy3-antihuman IgG solution.
5. Shake the rack up and down five times to help mixing of the probing solution at the surface of the arrays. Be careful not to shake the arrays out of the racks. Set a timer to countdown for 2 hours.

6. Lid the box and shake in the shaking incubator at 50 rpm, 20°C for the remainder 2 hours of incubation time.

6. Washing after Cy3-Anti-human IgG incubation

1. Towards the end of the incubation period, fill a slide staining box with 200 mL of fresh SAB buffer (20°C).
2. **Wash 1:** When the incubation has finished, lift the slide rack from its incubation solution and place into the fresh SAB wash solution.
3. Shake the rack gently up and down five times. Replace the lid and shake for 5 min at 50 rpm at room temperature.
4. **Wash 2:** After wash 1 has finished, lift the slide rack out of the dish and pour off the buffer into a beaker. Pour in 200 mL of fresh SAB buffer (20°C).
5. Shake the rack gently up and down five times. Replace the lid and shake for 5 min at 50 rpm at room temperature.
6. **Wash 3:** After wash 2 has finished, lift the slide rack out of the dish and pour off the buffer into a beaker. Pour in 200 mL of fresh SAB buffer (20°C).
7. Shake the rack gently up and down five times then replace the lid and shake for 5 min at 50 rpm at room temperature.
8. When the third wash has finished, lift the slide rack out of the dish and pour off the SAB. Fill the box with high purity water.
9. Place the slide rack in the water and shake gently up and down five times. Then pour off the high purity water. Repeat this step three times to ensure the buffer components are washed away from the slide rack and arrays.
10. Place a wad of Kimwipe tissue on the clean bench and also in the clean and dry staining box.
11. Remove the slide rack from the dish and bang gently five times on the wad of Kimwipe tissue to remove excess water.
12. Place the slide rack in the fresh staining box on top of the Kimwipe tissues.

7. Drying down the slides

1. Lid the box and dry the arrays by centrifugation for 4 minutes at 400 g.

Note: Add a balancing box if necessary.

8. Scanning

1. Place the slide in the slide holder with the barcoded side facing upward. Close and lock the cassette lid.
2. Place the slide holder into the Agilent slide carousel.
3. Scan the slides at 10 μ m resolution, 16-bit.

B. BIOINFORMATICS PROTOCOL

1. Image Analysis

Array images can be analyzed by any image quantification software including analysis software provided by scanner manufacturers, third-party software, or open source programs. The following is a brief description of how an image analysis is performed for the Sengenics Immunome Protein Array using GenePix Pro 7 software.

1. **Loading .tiff images** – Array images are automatically barcoded and saved in .tiff format after scanning. .tiff images are loaded into GenePix Pro 7 software to be analysed.

Note: GenePix only reads image data in 16-bit gray scale

Load .tiff Image

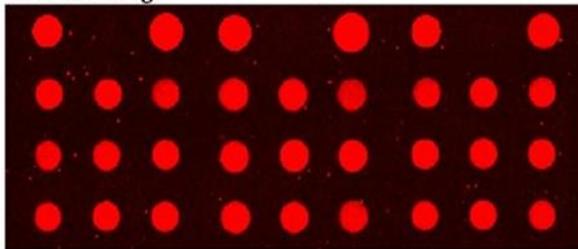


Figure A: Loading .tiff images onto GenePix Pro 7 Software

2. **Automatic Spot Detection** - A GAL (GenePix Array List) file for each array is generated to aid image analysis. Each GAL file will contain a list of proteins with respect to their positions on each array. Please note GAL files are grid files specific to GenePix software and may not be compatible with any other software. Loading the GAL file onto GenePix Pro 7 software will automatically generate grids on the array slide. Grids are then automatically aligned on each spots for data extraction.

Load GAL file and auto-detect spots

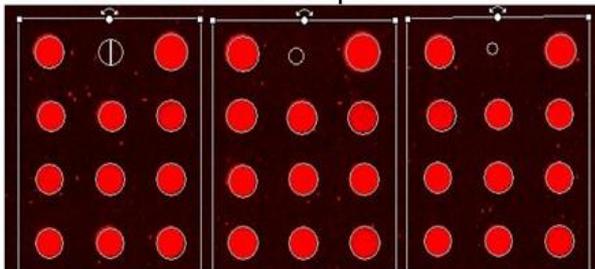


Figure B: Loading a GAL file onto GenePix Pro 7 Software

3. **Data Extraction** – Data output consist of; probe name, protein name, position, median Relative Fluorescence Unit (RFU) and background RFU for each spots. Results are generated and saved in .gpr format.

Results

Block	Column	Row	Name	ID	X	Y	Dia.	F635 Median
1	1	1	zzz_Cy3BSA_1	CTRL_A01	9.	4.	180	19305
1	2	1	zzz_iga1	CTRL_A03	1.	4.	150	219
1	3	1	zzz_igg1	CTRL_A02	1.	4.	200	20999
1	1	2	ZZ_con1	PLTA_A01	9.	5.	170	2084
1	2	2	STAT3	PLTA_A02	1.	5.	170	1944
1	3	2	ZZ_con2	PLTA_A03	1.	5.	180	1198
1	1	3	NF2	PLTA_A04	9.	5.	170	1264
1	2	3	ZZ_con3	PLTA_A05	1.	5.	180	1414
1	3	3	ACSL6	PLTA_A06	1.	5.	180	1587
1	1	4	ZZ_con4	PLTA_A07	9.	5.	170	1544
1	2	4	BCL10	PLTA_A08	1.	5.	170	2133
1	3	4	MNAT1	PLTA_A09	1.	5.	180	1704
1	1	5	ATF1	PLTA_A10	9.	6.	170	1711
1	2	5	CDK7	PLTA_A11	1.	6.	180	1468
1	3	5	ATIC	PLTA_A12	1.	6.	180	1414
1	1	6	AURKA	PLTA_A13	9.	6.	180	2390
1	2	6	SEPT9a	PLTA_A14	1.	6.	170	5375
1	3	6	ILK	PLTA_A15	1.	6.	180	1251
1	1	7	zzzz_NPM1	PLTA_A16	9.	7.	170	2094
1	2	7	SGK1	PLTA_A17	1.	7.	170	1405
1	3	7	RHOH	PLTA_A18	1.	7.	170	2280
1	1	8	CDK9	PLTA_A19	9.	7.	170	1612
1	2	8	MSN	PLTA_A20	1.	7.	170	2049
1	3	8	STK3	PLTA_A21	1.	7.	170	1819
1	1	9	ZMYND8	PLTA_A22	9.	7.	170	1696
1	2	9	TP53RK	PLTA_A23	1.	7.	180	1615
1	3	9	TRIP11	PLTA_A24	1.	7.	170	1618
1	1	10	NCK1	PLTB_A01	9.	8.	170	2134
1	2	10	ZKSCAN3	PLTB_A02	1.	8.	170	4245
1	3	10	INHBA	PLTB_A03	1.	8.	170	2227
1	1	11	POU2F2	PLTB_A04	9.	8.	170	1779
1	2	11	ISG20	PLTB_A05	1.	8.	170	2002
1	3	11	ZNF136	PLTB_A06	1.	8.	170	1378
1	1	12	STUB1	PLTB_A07	9.	8.	180	21261
1	2	12	NFE2	PLTB_A08	1.	8.	170	2408
1	3	12	MYBL2	PLTB_A09	1.	8.	170	1597
1	1	13	NR1H2	PLTB_A10	9.	9.	170	3515
1	2	13	PKM2	PLTB_A11	1.	9.	170	1554
1	3	13	HESE6	PLTB_A12	1.	9.	170	1880
1	1	14	SLA	PLTB_A13	9.	9.	170	1459
1	2	14	SPIB	PLTB_A14	1.	9.	170	1608

Figure C: Data output from GenePix Pro 7 Software

2. Data pre-processing

The following steps are performed to verify the quality of the protein array data before proceeding with data analysis:

1. Calculate net intensities for each spot by subtracting background signal intensities from the foreground signal intensities of each spot.
2. Remove replica spots with RFU ≤ 0.
3. Zero net intensities if only 1 replica spot remaining.
4. Calculating the percentage of coefficient of variant (CV%) of intra-protein, intra-slide and inter-array (Equation 1) to determine the variations between the replica spots on each slide.

$$CV\% = \frac{\text{Standard Deviation of replica Spots}}{\langle \text{Net Intensities of replica spots} \rangle} \times 100\% \quad \text{Equation 1}$$

Flag a set of replica spots with only 2 or less replica/s remaining and CV% > 20% as "High CV". The mean RFU of these replica spots (i.e. proteins) will be excluded from the downstream analysis.

For proteins/controls with a CV% > 20% and with 3 or more replica spots remaining, the replica spots which result in this high CV% value were filtered out. This was done by calculating the standard deviation between the median value of the net intensities and individual net intensities for each set of replica spots. The spot with the highest standard deviation was removed. CV% values were re-calculated and the process repeated.

5. Calculating the mean of the net intensities for the remaining replica spots.

3. IgG Serial Dilution – Quality Control

The IgG dilution series acts as a positive control for assessing the binding capacity of fluorescent-conjugated secondary incubation. Accurate serial dilution quantification is used as a benchmark for ensuring that labelling efficiency and spot detection pass quality control thresholds. The experimental ratios for the dilution series of IgGs are as follows:

$$\begin{aligned}
 & \mathbf{IgG1 : IgG2 : IgG3 : IgG4 : IgG5 : IgG6} \\
 & \mathbf{x : 0.5x : 0.25x : 0.125x : 0.0625x : 0.03125x}
 \end{aligned}
 \tag{Equation 2}$$

where x is the initial concentration (i.e. EFU) of IgG spot. Figure D shows an image of all the replica IgG spots' serial dilution on the Immunome slide.

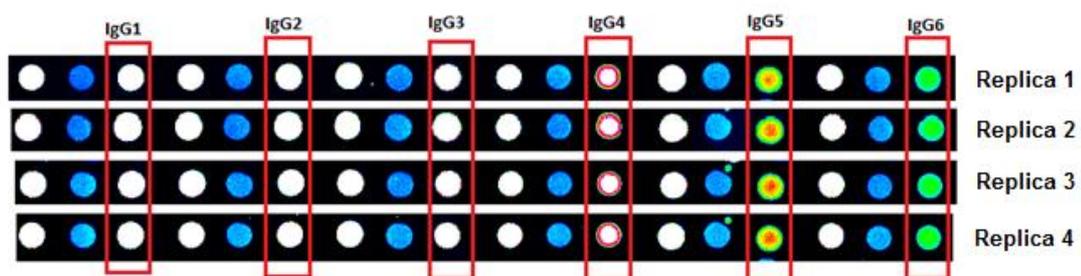


Figure D: Sengenics-IMMUNOME™ Protein Array quadruplicate images for IgG control spots showing serial dilution (Left to right).

The methods used to ensure the same amount of secondary antibody used in the experiment are as follow:

1. Plots of mean net intensities of IgG 1 – 6 spots
2. Plot of CV% for IgG 1 – 6 and calculate the average IgG CV% across all samples. Threshold of CV% is 10%.
3. Calculates the average CV% between the IgG serial dilution across all samples and the experimental dilution.

4. Cy3-BSA Analysis – Quality Control

Cy3-BSA controls act as positive controls for each array on the slide. Cy3-BSA 1 – 23 markers are present on each slide and their concentration are kept constant throughout the experiment.

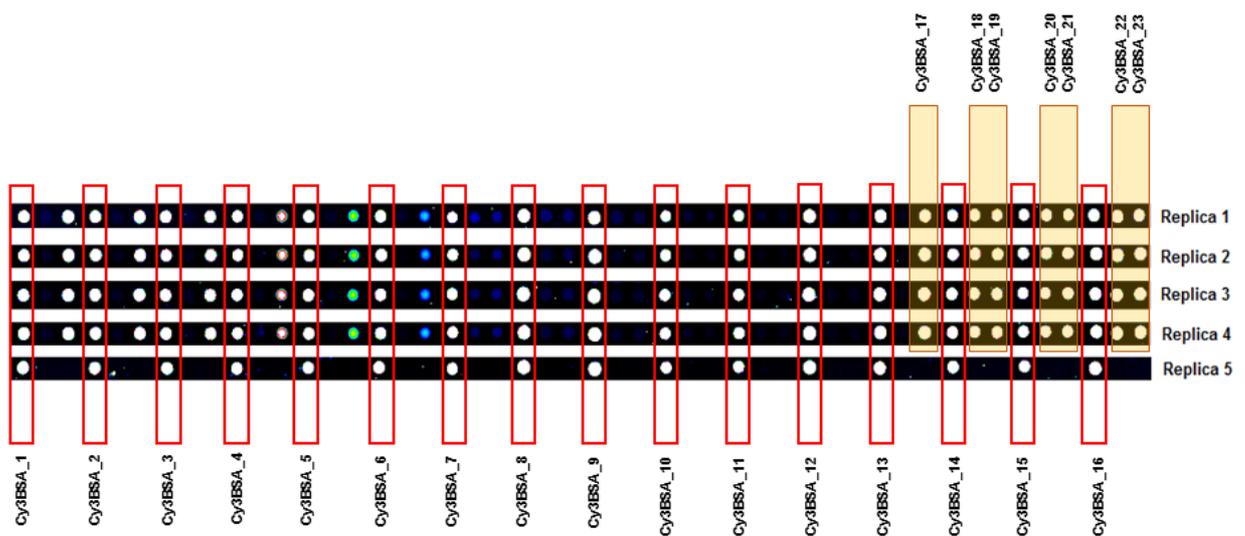


Figure E: SENGENICS-IMMUNOME™ Protein Array replica images for all 23 cy3BSA control spots showing. For slide orientation purpose, Cy3BSA 1 to 16 are spotted in replicates of 5 and Cy3BSA 17 to 23 are spotted in quadruplicate.

The methods used to ensure consistency of Cy3-BSA controls probes are as follows:

1. Plot of mean net intensities for Cy3-BSA 1 -23.
2. Plot of CV% for Cy3-BSA 1 – 23 and calculate the average Cy3-BSA CV% across all samples. Threshold of CV% is 10%.

3. Calculate the average CV% between 23 Cy3-BSA spots within each slide across all samples.

5. Data Normalisation

Composite normalisation of data using both quantile-based and total intensity-based modules [1]. This method assumes that different samples share a common underlying distribution of their control probes while taking into account the potential existence of flagged spots within them. The Immunome array uses Cy3-labelled biotinylated BSA (Cy3-BSA) replicates as the positive control spots across slides. Hence it is considered as a housekeeping probe for normalisation of signal intensities for any given study.

The quantile module adopts the algorithm described by Bolstad et al., 2003 [2]. **This re-organisation enables the detection and handling of outliers or flagged spots in any of the Cy3BSA control probes.** A total intensity based module was then implemented to obtain a scaling factor for each sample. This method assumes that post-normalisation, the positive controls should have a common total intensity value across all samples [3]. **This composite method aims to normalise the protein array data from variations in their measurements whilst preserving the targeted biological activity across samples.** The steps are as follows:

Quantile-Based Normalisation of all cy3BSA across all samples

(i = spot number and j = sample number)

1. Load all Cy3-BSA across all samples, j , into an $i \times j$ matrix X
2. Sort spot intensities in each column j of X to get X_{sort}
3. Take the mean across each row i of X_{sort} to get $\langle X_i \rangle$

Intensity-Based Normalisation

1. Calculate sum of the mean across each row i , $\sum \langle X_i \rangle$
2. For each sample, k , calculate the sum of all Cy3-BSA controls, $\sum X_k$
3. For each sample, k ,

$$\text{Scaling factor (k)} = \frac{\sum \langle X_i \rangle}{\sum X_k} \quad \text{Equation 3}$$

6. Biomarker Discovery – A Recommended Approach

The recommended approach for identification and ranking of biomarkers for the Immunome Protein Array studies us by using a penetrance-based fold change (pFC) method. This measures the likelihood that a given raw fold change is significant, thus increasing the relevance and reliability of results. Biomarker discovery from protein arrays is markedly different from that cDNA, RNA, and DNA arrays. The frequency of any given single biomarker is far lower in a cohort study when looking at autoantibody biomarkers compared to mRNA or DNA biomarkers.

High concentrations of an arrayed protein may occasionally give a false positive signal in serology assays because of concentration-driven, non-selective binding of an immunoglobulin to the target. This can arise theoretically due to an avidity effect: weak, non-specific immunoglobulin binding sites on a specific protein becoming coupled across multiple neighbouring protein molecules via an antibody as a result of the high density of immobilized protein, thus making the protein appear to be highly antigenic. Whenever this phenomenon occurs, it would be expected to be observed in the healthy control samples and will give rise to high intensity signals and/or signals that are close to saturation on the arrays. In Immunome, proteins such as RBPJ and IGHG1 show consistently high signal intensities across all samples.

For this reason, given a large sample number (i.e. 100 – 200 samples) and availability of sample cohort, a penetrance-based fold change (pFC) analysis method is implemented for the identification of highly expressed proteins in each case sample. This method will remove any false positive signals from the data by setting a protein-specific threshold (i.e. background threshold). This threshold is calculated on a per-protein basis via the mean + 2 standard deviation intensity of the signal for each specific protein measured for healthy control samples. A step-by-step description of this method is as follows:

1. Normalisation of data.
2. Individual fold change for case/control is calculated by dividing the RFU value for each protein in each sample, H , by the mean + 2 standard deviation of the RFU values of each protein across all the control samples (i.e. background threshold).

$$\text{Individual FC} = \frac{H_{\text{Case or Control}}}{\mu(H_{\text{Control}}) + 2 \times SD(H_{\text{Control}})} \quad \text{Equation 4}$$

3. Penetrance frequency (number of case/control samples with individual fold changes ≥ 2 fold) are calculated for each protein.
4. The Mean Penetrance RFU for both cases and controls are calculated for each protein.

$$\text{Mean Penetrance, } \mu(H_{\text{Case/Control}[i]}) = \frac{\sum H_{\text{Case/Control}[i]}}{\text{Frequency}_{\text{Case/Control}}} \quad \text{Equation 5}$$

5. Penetrance Fold Change for case samples are calculated for each protein.

$$\text{Penetrance Fold Change}_{\text{case}} = \frac{\mu(H_{\text{Case}[i]})}{\mu(H_{\text{Control}}) + 2 \times SD(H_{\text{Control}})} \quad \text{Equation 6}$$

$$H_{\text{Case}[i]} = H_{\text{Case with FC Case} \geq 2 \text{ fold}}$$

6. Putative biomarkers are identified and ranked according to the following criteria:
 - i. $\text{Penetrance Fold Change}_{\text{case}} \geq 2$.
 - ii. High Frequency % Differential (typically $\geq 10\%$).

7. References

1. Duarte, Jessica, et al. "Protein function microarrays: design, use and bioinformatic analysis in cancer biomarker discovery and quantitation." *Bioinformatics of Human Proteomics*. Springer Netherlands, 2013. 39-74.
2. Bolstad, B. M., Irizarry R. A., Astrand, M, and Speed, T. P. (2003) A Comparison of Normalization Methods for High Density Oligonucleotide Array Data Based on Bias and Variance. *Bioinformatics* 19(2) ,pp 185-193.
3. Causton HC, Quackenbush J, Brazma A. *Microarray gene expression data analysis: a beginners' guide*. 1st ed. Malden: Blackwell Publishing; 2004.