NATIONAL PROGRAMME ON TECHNOLOGY **ENHANCED LEARNING** (NPTEL)

Applications of Interactomics using Genomics and Proteomics technologies

> **Course Introduction by** Prof. Sanjeeva Srivastava

> > **MOOC-NPTEL**

Applications of Interactomics using Genomics and Proteomics Technologies

Lecture-13 **Applications of Protein microarrays in Malaria Research-11**

> Dr. Sanjeeva Srivastava Professor **Biosciences and Bioengineering IIT Bombay**

(Refer Slide Time: 00:27)

MOOC-NPTEL

Applications of Interactomics using Genomics and Proteomics Technologies

LECTURE - 13

Applications of protein microarrays in Malaria **Research-II**

Dr. Sanjeeva Srivastava

Professor **Biosciences and Bioengineering IIT Bombay**

MOOC-NPTEL Applications of Interactomics using Genomics and Proteomics Technologies IIT Bombay Welcome to MOOC course on applications of interactomics using genomics and proteomics technologies. We are discussing about different microarray based platforms and how to perform some biological applications on these chips.

(Refer Slide Time: 01:36)



In last lecture Ms. Apoorva Venkatesh showed you how to perform a microarray experiment using serum samples obtained from patients who are suffering from falciparum or vivax malaria.

Today we are going to continue the demonstration and also show you the ways to do data normalization and how to do a microarray data analysis, specifically if you go worse to look for a biological question of question of interest.

In this case we are going to talk about several ways of how to make meaningful data from the patients who are suffering from malaria using protein microarray based platforms. So let's have this lecture and demonstration session today.

Welcome to the MOOC NPTEL course on applications of interactomics using genomics and proteomics technologies. I'm Apoorva Venkatesh, your TA for this course, and today we are going to talk about a microarray data normalization and analysis.

In the last lecture we're trying to profile humeral responses of malaria positive patients using microarray technology, so we are going to start from there, what we are going to do today is to see how to normalize microarray data using excel, so what we'll do is we'll start with the raw file you get from the microarray scanner, right, so once you take a slide and you scan it in the scanner you will extract the raw data and here is the excel sheet you see,

Excel File Edit View Insert Format Tools Cells Window P Help	@ 💭 ① ¥ 😜 🖓 №% (૨) Mon 456 pm Q 📰
P Side Alax	
) 🖀 🕅 🖬 😹 🕺 🗓 🗋 🛷 😂 - 🔃 - 🗵 - 🏠 - 🏗 - 🛃 🗂 🏥 mark 🖩 🧕	Q. Constant Street
A Roma Layout Tables Charts SmartArt Formulas Duta Peulaw	A 0
Alt Fast Algebra Runsar	Format Galla Thereas
📄 👌 Fill + Caller Bodyl + 11 + Ar Ar 📟 📰 aler 😴 Wag Tet + General + 🛃 . Normal Mal	Kood NewYork Categories In . Do
The Others B / U A . A . B . A . B . A . B . A . B . A . B . A . B . A . B . A . A	tory Hout Dealer Same All
At 1 0.0 C 6 AT	
	M N O P O R S
ATE 1 C C C C C C C C C C C C C C C C C C	
2 20 40	
L So	
2 Type-were a result a Rest 1 Type-were a rest 1 Type-rest	
> Settings*	
Carrier, Apportation ine Solution and	
/ Pherson - 17	
o Waveergtha-635	
9 ImageFiles-C:\/poorva\g6_meathy\/ieathy\/	
0 Normalization/Method=None	
1 NormalizationFactors=1	
[2] Jpegtmage=C:\Apoorva\JgG_Healthy\Healthyside 6_H15_Rescan_400.100_28.5.17.jpg	
13 StdDev-Type 1	
4 RatioFormulations=W1/W2 (635/#2)	
5 FeatureType+Circular	
6 Barcode=	
7 BackgroundSubtraction=LocalFeature	
8 ImageOrigin=0, 0	
9 apegOrigi=1540, 1450	
0 Creator-GenePix Pro 7.2.27	
1 Scanner-GenePix 4000	
12 FocusPosition=0	
73 Temperature-0	
4 Lineskveraged=1	
S Commente	
6 PMTGalo-400	
7 ScanPowers 100	
The LacerPowers 1 M	
The second state and a second state of a second state of the secon	
AND	

this is the type of data you get, I'm showing you this one particular slide, so I'll just like to repeat that one slide can probe 8 patient serum, so here in this one particular excel which you see here,



(Refer Slide Time: 02:35)

we actually had data for 8 patients.

So first of all I'm going to show you is how to reorganize this data, okay, so let's see first of all what kind of parameters are exported, and you will see that all important parameters are provided in this excel, for example start with pixel size is 10, the slide was canted a wavelength of 635 nanometer, then you go down normalization method this was not normalized yet, so it says none, then if you scroll down further you can see the PMT gain which is 400 scan per 100 laser power 1.34,

		1 101100	1000	era versope		P								0 1 4	Q 80%	(1) Mon 4/5	7 pm Q
••				_				P Side1.xis									
	0 0 🖌	10 · CE ·	I - 20 -	F. 6											Q+		
A Home Layout Tabl	es Charts	Smarth	f Formula	es Data	Review												-
East.	Fant			Algement			anter .				Farmat					Carlle .	Thernes
, i fil + Calor de	dal = 11	- A- A		and a second	West Text +	General			Rormal	Ref	-		utrat [Calculation	334 .		- Art- 8
ats Ocer B /	2	4 · 4	- E.E.	31, 431, 43	Negs -	50 - 56	1 22 23	Conditional	(Deck Cell	ingle in the second	the second			liste	· Inset	Oerate Form	at Themas A
A11 100	- A scard	Power=100															
A 8	C	D	1.6	F	G	H	1	1	K	L	M	N	0	P	0	R	5
NormalizationFactors	1	-	-		-								-		_		-
Jpegimage-C:\Apoorv	al/igG_Healt	hy/Healthy	slide 6_H15,	Rescan_400	100_28.5.1	7.jpg											
StdDev=Type 1																	
RatioFormulations=W1	1/W2 (635/)	#2)															
FeatureType+Circular																	
Barcode=																	
BackgroundSubtractio	n-LocalFeat	ure .															
ImageOrigin=0, 0																	
JpegOrigin=1540, 1450)																
Creator+GenePix Pro 7	2.27																
Scanner-GenePix 400	3																
FocusPosition=0																	
Temperature-0																	
UnesAveraged=1																	
Comment-																	
PMTGain=400																	
Scatt Dwerp 100																	
Later Freer-1.34																	
Fiters- <empty></empty>																	
ScanRegion=0.0.2200	7200																
ScanArea+1																	
Supplier=																	
Block Column	Row	Name	ID	×	¥	Dia.	F635 Media	F635 Mean	F635 SD	F635 CV	8635	8635 Medi	8635 Mean	n 8635 SD	8635 CV	% > B635+1	% > 8635+2 F
1 1 1	1	I IgG mix 1	IgG mix 1	2110	1800	120	58283	55099	12719	2	3 955	955	1004	208	2	0 100	100
1 1 7	1 1	ligG mix 2	lgG mix 2	2360	1800	150	36869	35439	7842	2	2 992	992	1300	406	3	5 100	100
1 1 7	1 1	lgG mix 3	IgG mix 3	2610	1800	150	37620	35789	8873	2	4 1011	1013	1112	411	3	7 100	100
1 1 4	1	IgG mix 4	IgG mix 4	2860	1800	140	28591	27492	6328	2	3 1005	1005	1087	280	2	5 100	100
1 1 1	1	ligG mix 5	lgG mix 5	3120	1800	140	15633	14910	3772	2	5 994	994	1058	1 196	1	8 100	100
	Intellity ships	L'USA Berle	100 1 inches	to shife 2 Mill be		atthe shire 2 to	pearl, 400, 100	Healthy shield	L LYNN . Reports	100 1 100							
Northal View Read							Sum-0										
	100	0.00	-		-	with a		1. 000	00	0	0	S		10	0	1201	
	28		17		C	P 1		-			(02)	AL P		1.1	1 1		

(Refer Slide Time: 03:10)

so basically later on if you want to go back and check this slides again, if example, for instance if you forget the parameters you used you can always go and open this excel to see what you had done, right.

So now let's scroll down further you will see block, column, and row, (Refer Slide Time: 03:28)

	HH	XE	0.4	20 · (11 ·	I . 20	T. (a)	- A 1999	- 0									Q+		
Roma	Layout	Tables	Chart	SmartArt	Formula	Data	Pester												1
440.			Fart			Algoritant			rise .				Farrial					Carlls	Thank
. (*)	Fil + Cal	ibr (Bed)	6 11	 A. A. 		aber i	Wing Test +	General		1 miles	Rormal	Red	-		A 4 64	iculation.	X84 -		Ante
. 0	Gear + B	11	1.1.1.1.1	A . A .		3. 41.41	Nege -	10 - 56	3 24 23	Constiened	Check Call	- Explored to	ry	1.00	artar B	fie .	- Inter	Delate Format	Themes
Cab	1.0	0.0	fe fire							Tomating		_	_						
1	h	8		D	E	F	G	H	1	1	K	L	M	N	0	- P	Q	R	5
ScanA	rea=1	-								-									
Suppl	ier»																		
Block	Colu	ma	45	Name	ID	х	¥	Dia.	F635 Media	F635 Mean	F635 5D	F635 CV 0	8635	8635 Media	8635 Mean	8635 SD	8635 CV	% > B635+19	4 > 8635+2
	1	1		lgG mix 1	lgG mix 1	2110	1800	120	58283	\$5099	12719	23	955	955	1004	208	20	100	100
	1	2	1	1 IgG mix 2	lgG mix 2	2360	1800	150	36869	35439	7842	22	992	992	1100	406	35	100	100
	1	3	1	1 lgG mix 3	lgG mix 3	2610	1800	150	37620	35789	8873	24	1013	1013	1112	413	37	100	100
	1	- 4		l IgG mix 4	IgG mix 4	2860	1800	140	28591	27492	6328	23	1005	1005	1087	280	25	100	100
	1	5	1	1 lgG mix 5	lgG mix 5	3120	1800	140	15633	14910	3772	25	994	994	1058	298	18	100	100
	1	6	1	1 lgG mix 6	IgG mix 6	3360	1800	140	6303	6439	1208	18	975	975	1004	159	15	100	100
	1	7	1	a human la	g a human lig	3620	1800	120	65535	61423	10277	16	981	981	1080	314	29	100	10
	1	8	1	1 a-human la	çə-human iş	3870	1800	140	36126	34035	10236	30	998	998	1092	276	25	100	100
	1	9		a human ly	g a human lg	4120	1800	140	24230	23319	6712	28	1012	1012	1097	233	21	100	100
	1	10		a human la	(a-human ly	4360	1800	140	10185	9843	2642	26	991	991	1035	149	14	100	100
	1	11		a-human la	a human la	4610	1800	130	4139	4215	798	18	975	975	1001	122	12	100	300
	1	12		a human ly	a human ly	4860	1800	130	2337	2423	329	13	955	955	970	89		100	100
	1	13		1 TTB5	TTBS	\$110	1800	170	1174	1177	111	9	937	937	944	64	6	94	8
	1	14	-	1 TTBS	TTBS	\$360	1800	170	1161	1162	137	11	929	929	935	56	5	89	7
	1	15		1 1185	TTBS	5610	1800	170	1137	1130	104	9	941	941	949	60	6	88	7
	1	16		TTB5	TTBS	5860	1800	170	1359	1339	265	19	972	972	995	110	11	76	
	1	17	-	no DNA Co	r NoONA Cor	6110	1800	180	4652	4098	980	20	1017	1017	1062	367	15	100	100
	-	- 1	-	I NO DINA CO	I NOUTRA CO	2110	2050	170	4027	4061	620	15	955	955	1015	192	18	100	100
			-	I NO DINA CO	NOUNA CO	2360	2040	180	4231	4290	926	23	1031	1031	1100	242	22	100	100
		- 2	-	no DNA Co	Notifia Co	2630	2040	180	4387	4340	928	23	1029	1029	1001	230	19	100	30
		- 2	-	no DNA Co	No DNA Con	2800	2050	180	4299	4793	920	24	1026	1025	1093	299	10	100	30
		- 2	-	EBA175 O	1EBA175.01	1140	2030	120	1516	1551	316	20	081	081	1013	131	12	96	40
			-	EBA175 0	EBA175 0	3610	2040	120	1273	1277	151	20	977	977	1025	170	17	80	
	1			MIEP3 0.34	MSP3 0.34	3860	2040	130	30492	20403	2788	13	998	958	1089	364	33	100	10
			-					190											
11.7		1000	sectory side	Courd serves	and I setting	state a proc. No.	and the A se	anny snow 2, No.	ALC: NO. 100.	rearry yield	Contraction of the last	1 mar							

so this is very important, again let's go back to the slide layout, one slide can probe 8 patient sera, and one particular pad that is one pad which probes one patient sera has 4 blocks, so which means that if I keep scrolling down so every 4 blocks represents one patient data, right, so when I keep scrolling down and I go to block 5 a new patient begins, so that is what I'm going to talk to you about how to reorganize this, so if example if you see here, 4 ends here, right, (Refer Slide Time: 04:04)

-	Lorent	Tables	Charles	- Amarthe	Formulat	Outs De	and						- 1	H_{Λ}	1.15	100			· / -
1.0			Eart			Algement		No. 10							and in		-		1
. 19	10. + 0	Carlor (Body)	a 11.	 AriAri 	1.1.1.1	aler 🗇 🕷	mp Text = 0	eneral		1	Rental	Red	-B.1			100	100		a.
	Ower -	BIU		6 - A -	5.8.3	41 41 1	Negs -	1.00.0	22 23 4	Conditional I	Direct Lat		100	- 1	-	2	-	21	
		00-4								prosting			- 5	× /		1			
		00.7	i ann		And Address	1000		and it is not in			-		1.22						
-	4	3	17	hypothetic	PVK 05773	16160	5830	170	5095	5083	924	18	1000	-11	N.	1.1		1 /**	
	4	4	17	hypothetic	PVX 08502	16410	5830	160	6124	6032	845	14			-				1.1
	4	5	17	Empty	Empty	16650	5820	170	4961	4861	1125	23		Contraction of the				_// '	
	4	6	17	heat shock	PFI0875w-4	16920	5830	170	4622	4646	978	21			-	66 12	6-		
	4	7	17	erythrocyte	PFD1235w	17170	5830	160	5698	5645	1129	19	_		n .	10.000			
	4		17	blank	blank	17410	5840	170	4863	4939	896	18	1087	1087	1112	144	12	100	
	4	9	17	amino acid	PVK_11457	17670	5830	170	4985	4991	976	19	1098	1098	1129	153	13	100	
	4	10	17	Falstatin, p	PVK 09903	17910	5830	160	5273	5263	3045	19	1089	1089	1124	152	13	100	
	4	11	17	ubiquitin C	PVK_09148	18160	5840	160	5365	5427	1203	22	1095	1095	1132	167	34	100	
	4	12	17	hypothetic	PVK_00453	18410	5830	170	5430	5471	987	18	1109	1109	1145	171	34	100	
	4	13	17	Empty	Empty	18650	5820	170	4803	4821	1124	23	1122	1122	1179	243	20	100	
	4	14	17	erythrocyte	MAL6P1.1	18910	5830	170	5219	5183	906	17	1133	1133	1181	221	18	100	
	4	15	17	erythrocyte	PF11_0521	19160	5830	160	3966	4113	901	21	1081	1081	1123	179	15	100	
	4	16	17	Empty	Empty	19400	5820	170	4903	4864	1218	25	1044	1044	1084	186	17	100	
	. 4	17	17	Blank	Blank	19650	5830	170	982	985	51	5	997	997	1040	166	15	0	
	db.	1	1	ligG mix 1	igG mix 1	2090	10800	120	65326	56473	15582	27	972	972	1025	230	22	100	_
	- 45	2	1	lgG mix 2	IgG mix 2	2340	10810	340	39073	37378	7767	20	3014	1014	1096	334	30	100	
	5	3	1	lgG mix 3	IgG mix 3	2590	10810	140	33724	31295	8292	26	997	997	1105	398	36	100	
	5	4	1	IgG mix 4	igG mix 4	2840	10800	140	23712	22371	5342	23	980	980	1066	328	30	100	
	5	5	1	IgG mix 5	igG mix 5	3090	10810	140	11951	11455	2602	22	963	963	1023	196	19	100	
	5	6	1	IgG mix 6	IgG mix 6	3340	10810	150	4585	4652	1013	21	947	947	582	218	22	100	
	5	7	1	a-human la	(a-human lg	3590	10800	120	65535	61227	11580	18	958	958	1064	314	29	100	
	5	8	1	a-human lą	a-human lg	3840	10800	140	28818	26853	7737	28	960	960	1027	198	19	100	
	5	9	1	a-human la	(a-human lg	4090	10800	140	18224	17484	4939	28	966	966	1024	184	17	100	
	5	10	1	a-human la	a-human ig	4340	10800	140	6985	6849	1758	25	959	959	991	142	34	100	
	5	11	1	a-human la	a-human lg	4590	10800	340	3302	3185	635	19	935	935	952	90	9	100	
	5	12	1	a-buman la	ca-human lg	4840	10810	130	1841	1833	240	13	925	925	938	75	7	100	
	5	13	1	TTBS	1185	5080	10800	170	1119	1125	102		927	927	938	80		87	
5T		10 2 mail	ty shake 1.	UCLQ Residen 4	00 / Healthy stu	ie 2 Mill Bencart,	ADD / Health	shile 2, Report.	400.100. 10	natting winds 3,17	NAL Rescar, HOD	2 Pear					-		

blank and you will see that is IgG1, IgG mix 1 which starts again, so this is basically your new patient, so what you are going to do is we are going to first reorganize this, but before this let

me tell you which are the columns which are important for us, so now I'm going to scroll back up and we are going to go through the columns which we have on this excel.

STH H	XD	0.0	100 · 100 ·	I . to.	E	-										Q.		
the Layout	Tables	Charl	to Smarthe	Formate	Data	Faultra												
Aut		. Aust.			Algement				-			Farmer.					Cerie .	The
1 FR + CH	ter Bedy	1 10 11	A A	10.00	10 ANT 10	Wang Test -	General		100	Barringt	Bed	-		ear can	alatine .	. 20.	- (E)·	A
Our. B	1.5	1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1	A	: 新清:	1.61.61	Lings -	S. 1. 16	2 24 4	Congitional	Cherry 140	and training	inger and		etter her		inset.	Osiste Roman	There
40 2 0	0.1	fe Bor	8															
	100.00		1	Acres Arrest								1000						
vesAveraged=1																		
mment-																		
ATGain=400																		
aniPower=100																		
serPower+1.34																		
terso-(Empty>		-																
annegori-u.u.	1200,72	50																
anseres-1																		
ock Colu	ma l	how	Name	iD	×	v	Dia	F635 Media	1635 Mean	F635 50	F635 CV	8635	B635 Media	B635 Mean B	635.50	B635 CV	% > 9635+11	L > 863
1	1		1 lgG mix 1	lgG mix 1	2110	1800	120	58283	\$5099	12719	23	955	955	1004	208	2	100	1
1	2		1 lpG mix 2	lgG mix 2	2360	1800	150	36869	35439	7842	22	992	992	1100	405	3	100	- 1
1	3		1 lgG mix 3	lgG mix 3	2610	1800	150	37620	35789	8873	24	1013	1013	1112	413	37	100	1
1	4		1 lgG mix 4	lgG mix 4	2860	1800	140	28591	27492	6328	23	1005	1005	1087	280	21	1 100	1
1	5		1 lgG mix 5	IgG mix 5	3120	1800	140	15633	14910	3772	25	994	994	1058	198	10	1 100	1
1	6		1 lgG mix 6	lgG mix 6	3360	1800	140	6303	6439	1208	1 18	975	975	1004	159	15	300	1
1	7		1 a-human ly	e a tellar le	3620	1800	120	65535	61423	10277	16	981	981	1080	314	25	9 200	3
1	8		1 a-human ly	e a húMán ly	3870	1800	140	36126	34035	10236	30	958	958	1092	276	25	1 100	1
1	9		1 a-human ly	g a-human lig	4120	1800	140	24230	23319	6712	28	1012	1012	1097	233	21	1 100	1
1	10		1 a-human ly	g a-human lg	4360	1800	140	10185	9843	2642	26	991	991	1035	149	14	1 100	1
1	11		1 a-human li	g a-human ig	4610	1800	130	4139	4215	798	18	975	975	1001	122	1.	1 100	1
1	12		1 a-human li	g a-human ig	4860	1800	130	2337	2423	329	13	955	955	970	89		100	1
1	13		1 1185	TT85	\$110	1800	170	1174	1177	111		937	937	944	64		94	
1	14		1 1185	TTBS	5360	1800	170	1161	1162	137	- 11	929	929	935	56		. 89	
1	- 15		1 1185	1185	5610	1800	170	1137	1130	104		941	941	949	60			
	- 16		1 1185	1185	5860	1800	170	1359	1339	265	19	972	972	995	110	1	76	
1	17		1 NO DNA CO	I NOONA CO	6110	1800	180	4652	4055	980	20	3017	1017	1062	267	1	1 100	
			2 NO DRA CO	e Noonia Cor	2110	2030	1/0	4027	4091	620	10	900	930	1015	192		1 300	

So now apart from block, column, row, the name, and the ID,

basically this is your protein ID, we don't need any of the other columns except for column AH, so you see column AH here, (Refer Slide Time: 04:44)

	1 DOI 1	New Inser	Formal	Tools Da	a Window	9 110	rp.							0 4	C (101	(O) Man	4:59 pm Q
				-	-	-	And		Dide' All								
1 07 H 1	× × 10	10.4	10 · 111 ·	I . 50 .	下・国	1 Mb	- H 10								Q+		
Logo	ut Table	e Charles	SmartArt	Formula	Date	Faultra											
240		fast			Algement			10.00	the second second			Farmer				Garle	
1 FR -	Caller Bed	1 1 1	- A. A.	10.00	10 ANT	A state por -	Gerend	1		Record	Bar	Seed	Becke	Catcolonian			1 - AC
O Dear *	B 1	Q	· A -	(后:景)	3. 49.49	tings -		2 34 6	Constitute	them:cat	in the second	n inger		Rate	100	et Osiate	former There
1008 B	00	fu .															
U	V	W	Х	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH A	, AI	AJ	AK	AL
in states	and a state	Martine of	Married Bar	Bartine (B.)	Sec. Bude 1	Barn 83 (10)	N.F. Brech	a birele	See Sector	S	6	and the local division of the		CONT NAMES	CONT NAME	ALC: 1 1	See.
100000	100000	mecian of	mean or na	inactes so 1	rage natio (and a set in	3. F Piken	9798	Localerity 100	547118	SALAA	16.61	635 Median - 053	FATAA	4411858	340.077	rings O
100000	100000					-	0 100	1018	100	3/3/8	34447	16.61	37368	34447	0011838	84 570	
100000	100000						0 156	1000	100	35677	34776	16.61	35877	34776	5583073	81.964	0
100000	100000						0 156	954	100	22586	26487	16.61	77584	26487	4788797	94 304	0
100000	100000					-	0 156	540	100	14539	13916	16.61	14539	13016	2125458	45.54	
100000	100000						0 156	958	100	5378	5464	16.61	5328	5464	1004556	34.182	
100000	100000						0 120	759	100	64554	60442	16.61	64554	60442	7323712	192.175	0
100000	100000				0	-	0 156	994	100	35128	33337	16.61	35128	33037	5309401	119.359	0
100000	100000	0	0	0	0		0 156	958	100	23218	22307	16.61	23218	22307	3637695	95.373	0
100000	100000	0	0		0		0 156	946	100	9194	8852	16.61	9194	8852	1535508	59.114	0
100000	100000	0	0	0	0		0 120	842	100	3164	3240	16.61	3164	3240	505782	26.344	0
100000	100000	0	0	0	0	1	0 120	828	100	1382	1468	16.61	1382	1468	290760	16.326	0
100000	100000	0	0	0	0	1 0	0 208	1246	100	237	240	16.61	237	240	244727	3.641	-50
100000	100000	0	0	0	0		0 208	1198	100	232	233	16.61	232	233	241788	4.054	-50
100000	100000	. 0	0	0	0	1 1	0 208	1200	100	196	189	16.61	196	189	235127	3.017	-50
100000	100000	0	0	0	0	1 (0 208	1148	100	387	367	16.61	387	367	278417	3.127	-50
100000	100000	0	. 0	0	0	1 1	0 256	1474	100	3635	3681	16.61	3635	3681	1202661	21.772	0
100000	100000	0	0	0	0	1 (0 208	1215	100	3072	3106	16.61	3072	3106	844671	15.865	0
	10.000	Theorem where	USED Bentan	and 2 married	and the second	the second second		to see the set									

so this is basically your F635 medium minus B635 which is your basically your background signals right, so this is our the column which we actually extract and use for an analysis, and we don't need any other column here, so what I'm going to do, I'm going to first delete all unwanted columns to make this excel less complicated, so let's delete all of this, and then we go and delete this as well.

We also don't need these parameters for the analysis, I'm also going to delete this, so finally this is what you get.

1 09 1	XING	00	-	10-00-	I . 20 .	T	0											
	Lapout Tal	-	Charts	Smarthet	Parendas	Data Review												
.649.			Aut			Algertant	No.					Farmat					Cartes	The
(A) #	e + Cenerd	with l	4.11	- A- A-	1.0.00	aber 🔂 Weg Bet + Gav	-		100 a	Barrial	Red	tion of	Real Property lies	Carcolomo		Se . 3	. m.	(Ash-
120	- B /			A . A .		1 41 61 Flame - 80	- 16 - 2 - 3	26. 20	Congitioner	Constanting of		Fight	president.	Rate	- 0			and the second
	100		Barris .		a concentration	and the state of t		M.M.	Fernating	hanness of the local division of the local d		ball of the			_		tes core	
		- /	-	1.000	Long states						And in case	Street Course	and a second second	and the second second		100 A		
lock	Column	Roa	_	Name	1D	F635 Median - 8635			_	-	-	-		-		-	-	
	1_1_	1	1	leG mix 1	leG mix 1	57128												
	152	2	- 1	lati mia 2	lati mix 2	35877												
	1	1	- 1	laG mix 3	lgG mix 3	36607												
	1	4	1	leG mix 4	laG mix 4	27586												
	1	5	1	lgG mix 5	lyG mix 5	14639												
	1	6	1	lgG mix 6	lgG mix 6	5328												
	1	7	1	a-human la	a human la	64554												
	1		1	a-human la	a human la	35128												
	1	9	1	a-human la	a-human la	23218												
	1 1	0	1	a-human la	a-human la	9194												
	1 1	1	1	a-human la	a-human la	3164												
	1 1	2	1	a-human la	a human la	1382												
	1 1	3	1	TTBS	TTRS	237												
	1 1	4	1	TTBS	TTBS	232												
	1 1	5	1	TTB5	TTBS	196												
	1 1	6	1	TTBS	TTBS	387												
	1 1	7	1	no DNA Co	r NoDNA Cor	3635												
	1	1	2	no DNA Co	NoONA Cor	3072												
	1	2	2	no DNA Co	NoDNA Cor	3200												
	1	3	2	no DNA Cel	NOONA Cor	3358												
	1	4	2	no DNA Co	NoDNA Cor	3070												
	1	5	2	no DNA Co	NoONA Cor	3256												
	1	6	2	EBA175, 0.	EBA175, 0.1	535												
	1	7	2	EBA175, 0.	EBA175, 0.	295												
	1	8	2	MSP3, 0.3n	MSP3, 0.3n	19494												
	1	9	2	MSP3, 0.1/	MSP3, 0.1n	12887												
	1 1	0	2	Rh1 PEP8, 0	(Rh1 PEPB, 0	376												
1		7 1944	-	URA Beniat	ADD / Handha	claim 2 Mill Research 400 2 manhhor sh	in 2 Annat. 40	1.100. 2	Hard By shift 2.	THE Resear. AUD	1 194		_					
1	al line day	4	-				-	ALC: NOT THE R.										

(Refer Slide Time: 05:25)

So now you keep scrolling down and then you arrange all patients side by side, so when you do that this is what you get, so you also see that there are additional columns here, this is what you get from your gal file, and when you scroll right you will see that all the patients are now place next to each other, right, so this is that kind of excel you get first.

Now what I'm going to do is that I'm going to reorganize this excel to make it easier, (Refer Slide Time: 05:57)

63 E	1.5	1 26 1	0.0	4 13	3 · Z . 20. T	- A - A				Q
	Layers	a	-	Charts	SmartArt Formulas	Data Review				
an				Auri	and the second se	Algement	Series .		format	Daris Transa
6 FF	*	Anal		10	· A- A- = =	ale + Wap Ret + Garlers	# (z) 2	a hered he	d Good	Never Aller - Marine and - Aller -
200		B /	12/1	10.014	• ▲ • 프 프 프	40.40 Emp - 50-	56 1 24 24 25	ideara (abularas	sectors (speciality)	Input Deate Farmer Termer All*
	1	0.0	- fe				10	internal second		
1000			1.0				G	н	1	1
-	-	-					1			
A	-									
y		Arrey	Spot	Spot	and the second second					-
IN PO		Column	Now	Column	ADI spot ID (gal file)		CHEF	PlasmoOB_ID	Ore-Fragment	Description
1	1	1	1	1	igG mix 1		NIA	NA		IgG mix 1
2	1	- 1	1	2	igG mix 2		NA	NA		IgG mix 2
3	-1	- 1	- 1		BO WK 2		NIA	NA		IgO mix 3
۰.	-31	- 1	- 1	- 1	igG mix 4		NA	NA		igG max 4
2	-34	- 1	- 1		igo mix 5		INIA	NA		igo mix b
	-31	- 1	- 1		що так в		INA	NA		igo mote
1	-31	- 1	- 1		a-numan igo s		IN IA	NA		ano-numan igu 1
۰.	- 21	- 1	- 1		a human igo z		NA	NA		anti-Human igo z
	- 24	1	- 1		a numan igo a		NA	NA		anti-Human igo a
10	- 21		- 1	10	a numan igo 4		NIA	NA		ano-numan igo 4
11	-31		- 1	40	e-human igu s		NA	NA		anti-Human igG 5
14	-21		- 1	14	e numan igu e		nen.	NUM.		ano-ruman igo e
1.0	-21	- 1	- 1	10	1103		NUA	NIA		1100
1.5	-2	-	- 1	14	1100		NUR	N/A		7795
	1	- 1	- 1	10	1100		NUR.	NA		7790
	- 2	- 1		10	NoONA		NUR	NA		anDNA Control
	-2	-		1/	NuChia		NIA	NA		anDNA Control
	1	- 1			No.001A		NUR	NA		money control
20	- 2				No.0818		AUR .	NIA		anDhia Control
10.0	+ +1	Cambin		Compile	d Raw (2) Combined Raw (3)	Cambred Rev (K) Minus IVTT or	entrol / Legit HOC / metado	ata / Minus NTT control - sortest1	1.2	

this is combine data for all your patients and now we will reorganize the proteins, so as you will see that proteins are present in the same order as they are present in your slide, so what we first do is now that this is common data for all patients we have already put them together, we are going to bring the IgG mix from all four blocks together. So I'll repeat this slide layout once more, now you have 6 IgG mix here this is present in your block 1.

Similarly you have the same spots present in all four blocks of the same pad, so what I'm trying to do here, I'm trying to get all the IgG mix together so we'll have 24 such spots one after the other, we will also do the same thing for your anti human IgG mix.

And similarly we also going to do the same thing for your next spots, so when we rearrange our excel, this is how your excel will look, what I have done here, I've put all the 24 IgG mix of 1 pad together, right, we will go through all the columns once more, for example these are all just your spot details,

(Refer Slide Time: 07:09)

-	E	1.0	1.02%	P 41	1.00	a state of the	A LAND TO A LAND TO A LAND				0.	-
83	11.00	17	4.0		-		T. To. In. (R) - all costs in				A. Contraction	
•	Layout		Tables	Charl	64 . J	SmartArt	Pormulas Data Review					
	-			1910			and the second second	1 (#2)	Tarna	-	Carlo Thamas	
2			-	1010		00.00	The second second second second		Rorna Red	Sand Review	- · · · · · · · · · · · · · · · · · · ·	
ć	Dieser *	B	1.5	1000	110	114.11	E E E E 41 41 1 1 Map * 19 * %	> % 22 Conditioner	Calculation (NeckTor)	Exploratory	Insert Delete Format Themes A	3*
	\$	0.0	5.0	fe								
2	the state of the s	31	6	Ð	46	1000	€ŭ	н		1		
			Arra .									
			<i>C</i>	Array	80	et Sect						
	Ind	ex	Row	Column	R	w Column	ADI spot ID (gal file)	ORF	PlasmoDB_ID	ORF_Fragment	Description	
	1	1	1	1	1	1 1	ligG mix 1	N/A	NA		lgG mix 1	_
	2 2	190	1		2	1 1	IgG mix 1	N/A	NA		lgG mix 1	
	3 1	679	1	1	3	1 1	ligiti mix 1	N/A	NA		lgG mix 1	
	4 8	168	1		6	1 1	IgG mix 1	N/A.	NA		IgG mix 1	
	5	2	1	1	1	1 2	2 lgG mix 2	N/A.	NA		IgG mix 2	
	6 2	191	1	1	2	1 2	2 lgG mix 2	N/A	NA		IgG mix 2	
	7 1	180	1	1	3	1 3	2 lgG mix 2	N/A	NA		IgG mix 2	
	8 8	669	1	- 4	6	1 2	2 IgG mix 2	N/A	NA		igG mix 2	
	9	3	1	1	1	1 3	IgG mix 3	N/A	NA		IgG mix 3	
	10 2	192	1	1	2	1 :	IgG mix 3	N/A	NIA		IgG mix 3	
	11 5	581	1	3	3	1 :	IgG mix 3	N/A	N/A.		IgG mix 3	
	12 (170	1		6	1 3	IgG mix 3	N/A	NA		IgG mix 3	
	13	4	1	1	1	1 4	IgG mix 4	N/A	NA		IgG mix 4	
	14 2	293	1	1	2	1 4	IgG mix 4	N/A.	NIA		IgG mix 4	
	15 1	182	1		3	1 4	I IgG mix 4	N/A	NIA		IgG mix 4	
	15 8	871	1		6	1 4	IgG mix 4	N/A	NA		IgG mix 4	
	17	5	1	1	1	1 1	IgG mix 5	N/A	NA		IgG mix 5	
	18 2	294	1	1	2	1 1	5 lgG mix 5	N/A	NIA		IgG mix 5	
	19 1	683	1		3	1 1	lgG mix 5	N/A	NA		IgG mix 5	
	20 1	172	-	ten long	£	in a	Light rais 5	N/A	ALA		InClimite 5	
		-				1 444 (A) 2 10	second statistic of conduction and with a most succession	T rule tor T newson T a	and a second second second second			

now we'll go to your ADI spot ID which is your gal file, so this is what you get from your gal file, I'll come to this in a minute, before that let's talk about ORF, so this column here is basically your ID, this is also going to give you details about the fragment which has been printed on the chip.

So let's go to plasmo DB ID, now if you look at plasmo DB ID these are all basically each and (Refer Slide Time: 07:35)

					1 - 11-1-10	the rate manager & rath	C Bland along			A. D. I. C. A. MALED. MOR. STOPPIN
-	-		-		-		Drown Area			14
80.1	н л	5 10	10.1	- 14	909 · C	Z + 20 + 36 + 100 □ 100 tors in 100				Q.
-	Layers		dates .	Charts	Smarth	rt Parmulas Data Paview				
lak				feet		Algoriant Bottom	-	Farmat		Eals Trans
1.00		Artel		+ 10	12 Ac. A	aler Wenter Gorest	2 P. 4	Normal Ref. 1	and factor	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
20	- 100	B /	1.12		de - A	- 2 2 3 4 4 6 1 way - 10 1	1 di Constante	Calumpton (Decision) in	ingut	Treat Date Tarne Alt-
		0.6	- 14							
A		8.1	C	D	EIF	G	H	International Solution	1	
	52	24	1	1	2	7 EBA175.0.1mg/mL	MAL7P1 176	PF3D7 0731500	7	erythrocyte binding antigen175 (EBA175), pur
	53	601		3	2	6 MSP1, 0.3mg/mL	PFI1475w	PF3D7 0930300		merczoite surface protein 1 (MSP1), purified p
	54	802	1	3	2	7 MSP1, 0.1mg/mi.	PFI1475w	PF307 0930300		merozoite surface protein 1 (MSP1), purified p
	55	890	1	4	2	6 MSP2, 0.3mg/ml.	PFB0300c	PF3D7 0206800		merozoite surface protein 2 (MSP2), purified p
	56	891	1	4	2	7 MSP2, 0.1mg/mL	PFB0300c	PF3D7 0206800		merczolle surface protein 2 (MSP2), purified p
	57	25	1	1	2	8 MSP3, 0.3mg/mL	PF10_0345	PF307_1035400		merczoite surface protein 3 (MSP3), purified p
	58	26		1	2	9 MSP3, 0.1mg/mL	FF10 0345	PF3D7_1035400		merczoite surface protein 3 (MSP3), purified p
	59	894	1	4	2	10 Pf CSP, 0.3mg/mL	MAL3P2.11	PF3D7_0304600		circumsporozoite protein (CSP), purified prote
	60	895		4	2	11 Pf CSP, 0.1mg/ml.	MAL3P2.11	PF3D7 0304600		circumsporozoite protein (CSP), purified prote
	61	605		3	2	10 Pf LSA1, 0.3mg/mL	FF10_0356	PF3D7 1036400		liver stage antigen 1 (LSA1), purified protein (
	62	506		3	2	11 Pf LSA1, 0.1mg/mL	FF10_0356	PF3D7_1036400		liver stage antigen 1 (LSA1), purified protein (
	63	316		2	2	10 Rh1 PEP1, 0.3mg/ml.	PFD0110w	PF307_0402300		reticulocyte binding protein homologue 1 (RH)
	64	317	1	2	2	11 Rh1 PEP1, 0.1mg/mL	PFD0110w	PF307_0402300		reticulocyte binding protein homologue 1 (RH)
	65	27	1	1	2	10 Rh1 PEP8, 0.3mg/ml.	PFD0110w	PF3D7 0402300		reticulocyte binding protein homologue 1 (RH)
	66	28		1	2	11 8h1 PEP8, 0.1mg/mL	PFD0110w	PF3D7_0402300		reticulocyte binding protein homologue 1 (RH)
	67	314	1	2	2	8 Rh2, 0.3mg/mL	PF13_0198	PF3D7_1335400		reticulocyte binding protein 2 homologue a (R
	68	315		2	2	9 8h2, 0.1mg/mL	PF13_0198	PF3D7_1335400		reticulocyte binding protein 2 homologue a (R
	69	892	1	4	2	8 Pvivax AMA1 Eoto monomer prep2, 0.3mg/ml.	EU395600.1	NA		apical membrane artigen 1 (AMA1) Ectodoma
	70	893	1	4	2	9 Pvivax AMA1 Eato monomer prep2, 0.1mg/mi,	EU395600.1	NA		apical membrane antigen 1 (AMA1) Ectodoma
	71	603	1	3	2	8 Pvivax AMA1, 0.3mg/mL	EU395600.1	NA		apical membrane artigen 1 (AMA1), purified p
	72	604		3	2	9 Pulvax AMA1, 0.1mg/mL	EU395600.1	NA	A	apical membrane artigen 1 (AMA1), purified p
	73	279	1	1	17	7 PFL_0008_CIDR2	PFL_0008	NA	CIDR2	erythrocyte membrane protein 1, PfEMP1 (VA
	74	37	1	1	3	3 PFA0110we2s2	PFA0110w	PF3D7_0102200	Exon 2 Segment 2	ringinfected erythrocyte surface antigen (RES
	75	50	1	1	3	18 PFA0125ce1s2	PFA0125c	PF3D7_0102500	Exon 1 Segment 2	erythrocyte binding antigen181 (EBA181)
	76	991	1	4	8	5 PFA0175w_207	PFA0175w	PF3D7_0103600	Exon 2 of 7	conserved Plasmodium protein, unknown fund
	77	708	1	3	8	11 PFA0360c_2o2	PFA0360c	PF3D7_0107300	Exon 2 of 2	probable protein, unknown function
	78	81	1	1	5	13 PFA0410w-s1	PFA0410w	PF3D7_0108300	Segment 1	conserved Plasmodium protein, unknown fund
	79	370	1	2	5	13 PFA0410w-s2	PFA0410w	PF3D7_0198300	Segment 2	conserved Plasmodium protein, unknown fund
1.0.0		Com	tread Raw	Carrie	test fam (2)	Combined Raw [31] Combined Raw (6)] Minus (1/17 control) Log	i HOC _ metalata / Minut	NTT LINES - MOTHER		

every protein has a unique plasmo DB ID so that's what is mentioned in this column here, if you go to the next column which is ORF fragments, so this will explain your ORF your column H better, if you go here you will see that this specifies which exon segment is printed on the (Refer Slide Time: 07:52)

0 11 3		40.40		4.9 * * * * *	T . To . A.	· 1981 [] 1982 (1994						(de
		-	And	-	- Paradas	Algertant	Sector 1			Farral		Date Theread
84 -	-		+ 10	+ A- A		aler. I the fact	Text	- B.	harmat	Bal	Sand Server	Ye. 20. (1). Ab. 20.
Date -	8.4			4 - A		di di Lines -	10-16-1 1	2 Contrary	Calendaria	Construction of the	Automatical State	Ar an Ar
				-	Carlo Carlos	Children and and	A Real Property lies	Farmating	Are a construction of the second	- Research	Part of the second s	and the rouge three the
	100	6	0		1	6					Terrare Statement	
20	893		4	2	O Public AMA	Lets monomer non	iniant 0 fe	EU1005600.1	2014	-	and the owner where the owner w	Anical membrane artistes 1 (AMA1) Echatron
23	601			-	B Puinter AMA	1 /3 Reserved	bar or south out	E1/395600 1	NUA			apical membrane artigen 1 (AMA1) suited a
72	804		- 1	2	9 Pulvar AMA	1.0 1me/ml		EU395600.1	NIA			apical membrane artigen 1 (AMA1) mutiled p
78	279	-	÷	17	7 PEL 0008 C	082		PEL DOCH	NA		CIDR2	endbrocke membrane protein 1. PEMP1 (VA
74	37			3	3 PFA0110we	142		PFA0110w	PEND	0102200	Exon 2 Sflament 2	incidented enthropyte surface antigen (RES
25	50			3	16 PEAD125cm1	42		PEA0125c	PF307	0102500	Excent Semant 2	andbrocke binding antigen181 (EBA181)
76	991	1	4		5 PEA0175w 2	207		PFA0175w	PF307	0103500	Exon 2 of 7	conserved Plasmodium protein, unknown fun
77	708		3	8	11 PEA0360c 2	n2		PFA0360:	PF307	0107300	Expn 2 of 2	probable protein, unknown function
78	81		1	5	13 PEADEIDW-S	1		PEAD410w	PF307	0104300	Segment 1	conserved Plasmodium protein, unknown fun
79	370	1	2	5	13 PEAD410w-s	2		PFA0410w	PF307	0108300	Segment 2	conserved Plasmodium protein, unknown fun
80	947		4	5	12 PEAD410w-s	3		PFA0410w	PF307	0108300	Segment 3	conserved Plasmodum protein, unknown fun
81	45	.1	1	3	11 PFA0430ce1	*1		PFA0430c	PF307	0108700	Exon 1 Segment 1	secreted ockinete protein, putative (PSOP24)
82	804	1	4	3	3 PFA0430ce1	12		PFA0430c	PF307	0108700	Exon 1 Segment 2	secreted ockinete protein, putative (PSOP24)
83	669		3	6	6 PEADEROW	lol		PFA0490w	PF307	0110000	Exon 1 of 1	conserved Plasmodium protein, unknown fun
84	917	1	4	3	18 PFA0510we1	142		PFA0510w	PF307	0110500	Exon 1 Segment 2	bromodomain protein, putative
85	628	1	3	3	16 PEAOSLOwe	ls3		PFA0510w	PF307	0110500	Exon 1 Segment 3	bromodomain protein, putative
86	639	1	3	4	10 PFB0010w (r	renamed)		PFB0010w	PF307	0200100		erythrocyte membrane protein 1, PfEMP1 (VA
87	713		3	8	16 PF80100ce2	s1		FFB0100c	PF307	0202000	Exon 2 Segment 1	knobassociated histidinerich protein (KAHRP)
88	418		2	8	10 PFB0106c_2	02		PF80106c	PF307	0202200	Exon 2 of 2	Plasmodium exported protein, unknown funct
89	617	1	3	3	5 PFB0115we3	ls2		PFB0115w	PF307	0202400	Exon 1 Segment 2	conserved Plasmodium protein, unknown fun
90	667	1	3	6	4 PFB0120w_1	101		PFB0120w	PF307	0202500	Exon 1 of 1	early transcribed membrane protein 2 (ETRA)
91	53	1	1	4	2 PFB0150ce2	13		PF80150c	PF307	0203100	Exon 2 Segment 3	protein kinase, putative
92	392	1	2	7	1 PFB0170w_1	101		FFB0170w	PF307	0203600	Exon 1 of 1	conserved Plasmodium protein, unknown fun
93	681	1	3	7	1 PFB0250w_1	101		PF80250w	PF307	0205600	Exon 1 of 1	conserved Plasmodium protein, unknown fun
94	65	1	1	4	14 PFB0300c			PF80300c	PF307	0206800		merczoite surface protein 2 (MSP2)
95	663	1	3	5	17 PFB0305c_1	02		PFB0305c	Pf3D7	_0206600.1	Exon 1 of 2	mercizoite surface protein 5 (MSP5)
96	90	1	1	6	5 PFB0305c-e3	1		PFB0305c	PF307	_0206900.1	Exon 1	mercizoite surface protein 5 (MSP5)
9.7	374	1	2	5	17 PFB0310c 1	62		PF80310c	PF307	0207000	Exon 1 of 2	mercizoite surface protein 4 (MSP4)

chip, so basically as you know spots which are printed on the chip were not purified proteins, they were IVTT spots and basically not, so what is IVTT? In vitro transcription translation, so what was expressed? The whole protein was not expressed here, only a certain segment of a particular exon of a protein was expressed, right, so basically it's not really right to say that proteins were expressed on the chip, instead it will be better to say that poly peptides were expressed on the chip, so this particular column J gives us details about the poly peptide that was expressed and printed on the chip, right, so that is how you'll get this ADI spot ID which is the unique ID for each and every protein.

What I mean here is that if you go to plasmo DB ID and then if you try to look for duplicates will actually find duplicates here because it could be, that for the same protein different exon fragments were printed on the chip, so you might get duplicates here, whereas if you go to your ADI spot ID you will not find single.

(Refer Slide Time: 08:55)

83 H	10	36 1	0.0	*	100.	•	I · 20 · 3 · 🙆 🗄 · 100 · 10					Q.
	-	-	-	Charles	-	SmartArt.	Formulas Data Review					*
540				fiet			Algerant Ta			Farmal		Gala Thamas
A ##	* 3	Caliber (Bri	- 1/1	11		Ar Ar	ale - Way he - Garand	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Rental	field	Sand Neutral	1 1 . 1 . 1 . 1 . Ar. 1
(20m		B. /	42.1		de.	- A -	E E E G G G _ Nep - 10 - 14	> % 41 Constant	Calculation	Charles CARTING	Ingenerating State	inst Date fame Tame Alt-
14		0.0	·					Parage of		-		
A	Ť1	8 6	1	D	E	F	6	H En		1	1	K
80	- 1	539	1	3		6 1	PFB001.0w (renamed)	PFB0010w	PF307	0200100		endtrocide membrane protein 1, PEMP1 (VAR)
87		713	1	3		8 1	6 PFB0100ce2s1	PFB0100c	PF307	0202000	Exon 2 Segment 1	knobassociated histidinerich protein (KAHRP)
	1	418	1	2			0 PFB0104c 2o3	FFB0106c	PF307	0202200	Expn 2 of 2	Plasmodium exported protein, unknown function
85		817	1	3	- 3	1	5 PEBO115we1s2	FFB0115w	PF307	0202400	Exon 1 Segment 2	conserved Plasmodium protein, unknown functio
90		567		3	6	5	4 PFB0120w 101	FFB0120w	PF307	0202500	Exon 1 of 1	early transcribed membrane protein 2 (ETRAMP
91		53	1	1		4	2 PFB0150cm2x3	FFB0150c	PF307	0203100	Exon 2 Segment 3	protein kinase, putative
90	1	392	1	2	7	7	1 PFB0170w 101	PFB0170w	PF307	0203600	Expn 1 of 1	conserved Plasmodium protein, unknown function
93	1	581	1	3	- 1	7	1 PFB0250w 1c1	PFB0250w	PF307	0205600	Expn 1 of 1	conserved Plasmodium protein, unknown function
94	1	65	1	1		4 1	4 PFB0300c	FF80300c	PF307	0206800		mercycelle surface protein 2 (MSP2)
95		663	1	3	- 5	5 1	7 PFB0305c 1o2	PFB0305c	PF307	0206900.1	Exon 1 of 2	merczole surface proteis 5 (MSP5)
96		90	1	1	6	6	5 PF80305c-#1	FFB0305c	PF307	0206900.1	Exon 1	merczole surface protein 5 (MSP5)
97	1	374	1	2		5 1	7 PF80310c 1o2	PFB0310c	PF307	0207000	Exon 1 of 2	merozoite surface protein 4 (MSP4)
98	1	187	1	4		8	1 PFB0310c 2o2	FF80310c	PF307	0207000	Exon 2 of 2	merczoite surface protein 4 (MSP4)
99		69	1	1	. 5	5	1 PFB0310c-e1	PFB0310c	PF307	0207000	Exon 1	merczoite surface protein 4 (MSP4)
100	1	115	1	1		7 1	3 PFB0330c 2o4	FFB0330c	PF307	0207400	Exon 2 of 4	serine repeat artigen 7 (SERA7)
301	1	003	1	4		1 1	7 PFB0335ce3s1	FF80335c	PF307	0207500	Exon 3 Segment 1	serine repeat antigen 6 (SERA6)
102	1	901	1	4	2	2 1	7 PFB0340ce2s1	PFB0340c	PF3D7	0207600	Exon 2 Segment 1	serine repeat artigen 5 (SERA5)
103	1	182		4	7	7 1	3 PFB0345c 2o4	FFB0345c	PF307	0207700	Exon 2 of 4	serine repeat antigen 4 (SERA4)
304		428	1	2	9	9	3 PF80345c 4o4	PFB0345c	PF307	0207700	Exon-4 of 4	serine repeat antigen 4 (SERA4)
105	1	989	1	4	8	8	3 PFB0350c 2o4	PFB0350c	PF307	0207800	Exon 2 of 4	serine repeat antigen 3 (SERA3)
106	1	116	1	1	7	7 1	4 PF80765w_6o7	FFB0765w	PF307	0216700.1	Exon 6 of 7	conserved Plasmodium protein, unknown function
107	1	996	1	4		8 1	0 PFB0900c_2o2	PF80900c	PF307	0219700	Exon 2 of 2	Plasmodium exported protein (PHISTc), unknow
108		561	1	3	5	5 1	5 PF80910w_2o2	PFB0910w	PF3D7	0219900	Exon 2 of 2	Plasmodium exported protein, unknown function
109	1	148	1	4	5	5 1	3 PFB0915w-e2s1	PF80915w	PF307	0220000	Exon 2 Segment 1	liver stage antigen 3 (LSA3)
310	1	559	1	3	5	5 1	3 PF80915w-e212	FF80915w	PF307	0220000	Exon 2 Segment 2	liver stage antigen 3 (LSA3)
111	1	598	1	3		8	1 PFB0926c_2o2	PFB0926c	PF3D7	0220500	Exon 2 of 2	Plasmodium exported protein (hyp2), unknown f
312	1	808	1	2	7	7 1	7 PFB0930w_2o2	PFB0930w	PF307	0220600	Exon 2 of 2	Plasmodium exported protein (hyp9), unknown fi
113	1	705	1	3		8	8 PF80932w_202	PF80932w	PF307	0220700	Exon 2 of 2	Plasmodium exported protein (hyp8), unknown fi
	11	Cambie		Caria	and is	in th Fi	ambiened Raw (3) / Combinent Raw (6) / Minute (977) camput	Light FOC metallatu Min	us NTT control - as	the last		
Income in	-	freed	¥ .	-				Sam-0 *				

any single duplicate because these are unique ID's for each and every protein which takes into account the exon fragment which was printed on the chip, so that's what you will see here.

If you say for example, let's look at this particular row, $(D_{1} f_{1}, S_{1}) = (D_{1} f_{2}, S_{2})$

12	5 B (* X	0		10 · 0	1. 1	· 20 · 1 · (0) - 10 · 10 · 10 · 10					Q
-	na Lago	nut 1	Tables	Charles	- terms	rike:	Formulas Data Review					A
	Len.			Fard			Algoriant	Liter		Farmal		Carlin Themas
	a ##. *	Caller	(Peril)	10.00	i-liA-	Act		· .	Rentral	Red	Gand Revital	We . We . Wh . Make 18
	A Dawn	8	1 11	101.01	4 .	A -	15 10 10 40 40 10 10 10 1 10 1 10	a for all Constant			Instantion India	
-		1.01.1		1.44		1.1.1	WINTER GATE TO AND . CALLS	Survey of the second se				Insert Deate Former Themas Phil
ŝ		0.0	2.57	97								
-	-	634			4	10	PEROD (recommend)	DEBOO10m	06307	0206100		and troops membrane protein 1. DE MD1 A(8-2)
		713				10	PFB0010w (renamed)	PEBOUIOW	PF307	0200100	Even 2 Second 1	eryprocyse memorane protein 1, Protein 1 (Wel)
		418	- 1			10	PEDDIOLE 343	PEBOTOL	PESOT	0202000	Exon 2 of 2	Records of the sector of the s
		410			3	5	PEDDUDEC_202	PEBOILder	PENDY	0202200	Exon 1 Segment 2	constant Reported protein, unknown function
	90	007					PEDDI Mar 1o1	PEBO120m	PE307	0202400	Executed 1	and transcibed membrane projets 2 (ETRAMP)
	01	43					PEDOLOGICA INC.	PFB0150r	PENDY	0203100	Excert 7 Respect 7	sentry conscious internet promit 2 (C revent 2
	92	302		2	-	- 7	PEDDI New 1o1	PEB0170w	PENDZ	0203600	Excent a degree to a	protein knase, postwe
	93	081					PERCESSION 101	PEBO250w	PE307	0205600	Excent of 1	conserved Plasmodium protein, unknown function
	94	45	- 1		4	14	PEROBOAC	PEB0300c	PE307	0206800	Exercit of 1	mercroite surface orchein 2 /USP2)
	95	663			5	17	PEROBOSIC 1/02	PEB0305c	PE307	0206900.1	Excented 2	memorale surface protein 5 (MSPS)
	96	90		1	6	5	PEDGDOScal	PFB0305c	PF307	0206900.1	Exon 1	mercanite surface protein 5 (MSP5)
r	97	374	1	2	5	17	PFB0310c 1o2	PFB0310c	PF307	0207000	Excert of 2	mercapite surface protein 4 (MSP4)
۲	98	987		4	8	1	PIB0310c 2o2	PFB0310c	PF307	0207000	Extended 2	mercuroite surface protein 4 (MSP4)
	99	69		1	5	1	PE00310c+1	PFB0310c	PF307	0207000	Exce 1	mercizoite surface protein 4 (MSP4)
	100	115		1	7	13	PFB0330c 2o4	PFB0330c	PF307	0207400	Exon 2 of 4	serine repeat articen 7 (SERA7)
	101	1003		4		17	PF00335ce3t1	PFB0335c	PF307	0207500	Exon 3 Segment 1	serine repeat artigen 6 (SERA6)
	102	901		4	2	17	PFB0340ce2s1	PFB0340c	PF307	0207600	Exon 2 Segment 1	serine repeat antigen 5 (SERAS)
	103	982		4	7	13	PFB0345c 2o4	PFB0345c	PF307	0207700	Exon 2 of 4	serine repeat antigen 4 (SERA4)
	104	428	1	2	9	3	PFB0345c 4o4	PFB0345c	PF307	0207700	Exon 4 of 4	serine repeat antigen 4 (SERA4)
	105	989	1	4	8	3	PFB0350c 2o4	PFB0350c	PF307	0207800	Exon 2 of 4	serine repeat artigen 3 (SERA3)
	106	116	1	1	7	14	PFB0765w 6o7	PFB0765w	PF307	0216700.1	Exon 6 of 7	conserved Plasmodium protein, unknown function
	107	996	1	4	8	10	PFB0900c 2o2	PFB0900c	PF307	0219700	Exon 2 of 2	Plasmodium exported protein (PHISTc), unknown
	108	661	1	3	5	15	PFB0910w 2o2	PFB0910w	PF307	02199000	Exon 2 of 2	Plasmodium exported protein, unknown function
	109	948	1	4	5	13	PF80915w-e2s1	PFB0915w	PF307	0220000	Exon 2 Segment 1	liver stage antigen 3 (LSA3)
	110	659	1	3	5	13	PFB0915w-e2s2	PFB0915w	PF307	0220000	Exon 2 Segment 2	liver stage antigen 3 (LSA3)
	111	698	1	3	8	1	PFB0926c_2o2	PFB0926c	PF307	0220500	Exon 2 of 2	Plasmodium exported protein (hyp2), unknown fur
	112	408	1	2	7	17	PFB0930w 202	PFB0930w	PF307	0220600	Exon 2 of 2	Plasmodium exported protein (hyp9), unknown fur
	113	705	1	3	8	8	PF80932w_202	PFB0932w	PF307	0226700	Exon 2 of 2	Plasmodium exported protein (hypfi), unknown fu
	1	10.00	strend in a	all come	and from it	n 1964	shined from [10] Combined from [4] Manua (WW) commit	Look and The statement in the	NTT contain a sur			

if you say that this was the ID and this is exon 1 of 2 you will actually see the ID here, and 1O2, so this becomes your unique ID for each and every protein, why I'm telling you all this is because this is very important for data analysis for all, sometimes you might just start with an analyzing your I column and then you will figure out later that there are lot of duplicates, you

don't know what you are actually doing, so what we need to do is if you want to shortlist any antigens we need to consider the G column for analysis, right.

So now let's move on to the next column which is a description, so this we all know this basically describe what was printed on the chip, right, these are just basically the names of them, basically the names of the antigens, and the next column is your organism, so as you know you have two types of spots here plasmodium falciparum and plasmodium vivax, so basically this is going to tell you which organism does the antigen belong to,

							-		_
目 177 日 第 19 10 10 全 100 + 00 + 2 + 20 + 26 1回 10 10 10 10 H H							q.		
None Layout Tables Charts SmartArt Formulas Data Review									
Lat Fast Algement	bother		Farmal				Garia	There	
. If Fit + And + 10 + A+ A+		Normal Ref		•	and all	. 20-	· · ·	- <u>At</u>)-	22-
. ⊘Cer• B / U =• 🂁 + 🚣 + E Ξ Ξ 🗐 🗐 🗐 🖬 Napp • 📑 +	% 3 % d2 Constiened	Calculation	t Cell		iput .	Inset	Osista Form	t Themes	Aa-
U 100/-A									
K	And and a design of the local division of th	M	N	0	P	0	R	5	T
anti-Human IgG 6	NA	anti-human IgG	1375	677	473	609	413	767	532
erythrocyte binding antigen 140 (EBA140), purified protein (0.3mg/mL)	P. falciparum 3D7	Purified Protein	15197	4119	1549	631	473	1521	1362
erythrocyte binding antigen 140 (EBA140), purified protein (0.1mg/mL)	P. falciparum 307	Purified Protein	22583	5048	1484	403	331	1395	872
rythrocyte binding antigen 175 (EBA175), punified protein (0.3mg/mL)	P. falciparum 307	Purified Protein	20121	3331	1366	1280	691	1276	1202
rythrocyte binding antigen175 (EBA175), purified protein (0.1mg/mL)	P. faiciparum 307	Purified Protein	8607	1583	924	757	368	852	678
serozoite surface protein 1 (MSP1), purified protein (0.3mg/mL)	P. falciparum 307	Purified Protein	26420	16334	294	333	446	1048	925
nerozoite surface protein 1 (MSP1), purified protein (0.1mg/mL)	P. faiciparum 307	Purified Protein	13938	8283	672	474	783	863	719
nerozoite surface protein 2 (MSP2), purified protein (0.3mg/mL)	P. falciparum 307	Purified Protein	17291	17732	903	8078	9174	2071	9867
erozoite surface protein 2 (MSP2), purified protein (0.1mg/mL)	P. faiciparum 307	Purified Protein	13110	18207	727	8151	7477	1980	10175
erozoite surface protein 3 (MSP3), purified protein (0.3mg/mL)	P. falciparum 307	Purified Protein	10740	4711	4310	892	297	2992	1891
erozoite surface protein 3 (MSP3), purified protein (0.1mg/mL)	P. faiciparum 307	Purified Protein	7481	3326	2327	726	313	1995	1239
roumsporozoite protein (CSP), purified protein (0.3mg/mL)	P. faiciparum 307	Purified Protein	12416	7936	174	152	268	630	407
ircumsporozoite protein (CSP), purified protein (0.1mg/mL)	P. faiciparum 307	Purified Protein	3792	6000	176	140	94	280	566
ver stage antigen 1 (LSA1), purified protein (0.3mg/mL)	P. falciparum 307	Purified Protein	37614	34824	19824	17699	3078	13519	17585
ver stage antigen 1 (LSA1), purified protein (0.1mg/mL)	P. faiciparum 307	Purified Protein	18711	18898	9272	6790	1352	5206	9105
diculocyte binding protein homologue 1 (RH1) PEP1, purified protein (0.3mg/mL)	P. falciparum 307	Purified Protein	596	1088	189	197	210	242	270
ticulocyte binding protein homologue 1 (RH1) PEP1, purified protein (0.1mg/mL)	P. faiciparum 307	Purified Protein	198	626	163	116	175	188	271
ticulocyte binding protein homologue 1 (RH1) PEP8, purified protein (0.3mg/mL)	P. falciparum 307	Purified Protein	302	359	284	236	187	417	673
sticulocyte binding protein homologue 1 (RH1) PEP8, purified protein (0.1mg/mL)	P. falciparum 307	Purified Protein	92	180	175	172	158	183	273
eticulocyte binding protein 2 homologue a (RH2a), purified protein (0.3mg/mL)	P. falciparum 3D7	Purified Protein	32348	4005	2439	3471	4877	14303	4903
eliculocyte binding protein 2 homologue a (RH2a), purified protein (0.1mg/mL)	P. falciparum 307	Purified Protein	11257	2010	1367	1759	1897	5012	2196
pical membrane antigen 1 (AMA1) Ectodomain monomer, purified protein (0.3mg/mL)	P. vivax Palo Alto	Purified Protein	63088	43280	5127	1829	1468	49717	64027
pical membrane antigen 1 (AMA1) Ectodomain monomer, purified protein (0.1mg/mL)	P. vivax Palo Alto	Purified Protein	63048	38950	2709	1213	934	41403	63423
pical membrane antigen 1 (AMA1), purified protein (0.3mg/mL)	P, vivax Palo Ato	Purified Protein	63317	45303	5026	1817	1819	53167	63987
pical membrane antigen 1 (AMA1), purified protein (0.1mg/mL)	P. vivax Palo Ato	Purified Protein	63345	39267	3839	1716	1425	44904	64003
rythrocyte membrane protein 1, PfEMP1 (VAR)	P. falciparum	INTT	6799	5126	2785	5373	4355	9072	20407
rginfected erythrocyte surface antigen (RESA)	P. falciparum 307	INTT	13575	16656	2674	5852	3407	7356	13275
rythrocyte binding artigen181 (EBA181)	P. faiciparum 307	INTT	21760	4278	4717	7642	4723	8695	17818
Combined Raw (2) Combined Raw (2) Combined Raw (3) Combined Raw (6) Minus MTT com	woi / Logi HOC / metalatu / Mi	nus fort commai - someth				****	****		
Normal View Ready	Sam-0 *								

(Refer Slide Time: 10:08)

so you have plasmodium falciparum 3D7 here for instance and probably and if you scroll down further you will see plasmodium vivax sal-1, right, so this is going to give you details of the organism.

And then the next column which is M is going to talk to you about deep preparation of the spot, like for example you have the first few spots are basically your IgG mix right, so then the preparation is basically your IgG mix, it's not an IVTT spot, now if you scroll down further you will have similar anti human IgG, again you scroll down further you have certain purified proteins which are nothing but a controlled proteins, so our control spots are were printed as purified proteins and not as IVTT spots.

(Refer Slide Time: 10:58)

11 日 日 米 山 こ が 20・11・12・15・16 日 通 101 単							Q		
A three Land Salar Charts Bearthat Parmin Data Pastry									
han fast Alignmen	Arts		Autos				Carls .	There	
a state + land + 10 + A+ A+ = and aber = may hat + General		Annual Annual	-		-	The .	Sec. 123	A Tanta	10.
	to a ful of Constant	Columbus E			reput	A BE	Geneta Form	at Theres	Aa-
K82 1 0.0 / Je probable protein, unknown function									
K.	L.	M	N	0	P	Q	8 1	5	T
reliculocyte binding protein homologue 1 (RH1) PEP8, purified protein (0.1mg/mL)	P. faiciparum 307	Purified Protein	92	180	175	172	158	183	271
reticulocyte binding protein 2 homologue a (RH2a), purfied protein (0.3mg/mL)	P. falciparum 3D7	Purified Protein	32348	4005	2439	3471	4877	14303	490
resource/se binding protein 2 homologue a (RH2a), purified protein (0.1mg/mL)	P. falciparum 3D7	Purified Protein	11257	2010	1367	1759	1897	5012	219
apical membrane antigen 1 (AMA1) Ectodomain monomer, purified protein (0.3mg/mL)	P. vivax Palo Ato	Purified Protein	63088	43280	5127	1829	1468	49717	6402
spical membrane antigen 1 (AMA1) Ectodomain monomer, purified protein (0.1mg/mL)	P. vivax Palo Atto	Purified Protein	63048	38950	2709	1213	934	41403	6342
apical membrane antigen 1 (AMA1), purified protein (0.3mg/mL)	P. vivax Palo Alto	Purified Protein	63317	45303	5026	1817	1815	53167	6358
apical membrane antigen 1 (AMA1), purified protein (0.1mg/mL)	P. vivax Palo Ato	Purified Protein	63345	39267	3839	1716	1425	44904	6400
erythrocyte membrane protein 1, PfEMP1 (VAR)	P. falciparum	INTT	6.799	5126	2785	5373	4355	9072	2040
ringinfected envityootyle surface antigen (RESA)	P. faiciparum 307	IVTT	13575	16655	2674	5852	3407	7354	1327
erythrocyte binding antigen 181 (EBA181)	P. faiciparum 307	INTT	21760	4278	4717	7642	4723	8695	1781
conserved Pasimodium protein, unknown function	P. faioparum 307	INTT	7815	15681	2437	5393	2984	7865	3413
probable protein, unknown function	P. faloparum 307	IN/TT	1952	1155	1071	2323	1418	2280	340
conserved Plasmodium protein, yaknown function	P. faiciparum 307	INTE	15274	2058	7096	16626	4335	8980	8277
conserved Plasmodium proteid A Shown function	P. taiciparum 307	INTT	11082	6999	25270	7807	5386	9661	1872
conserved Plasmodium protein, uriknown function	P. fatcipartern 307	INTT	16322	4601	7963	12911	3655	6971	1672
aecreted ockinete protein, putative (PSOP24)	P. faiciparum 307	INTT	16505	17856	1245	10547	2972	8597	1480
secreted ackinete protein, putative (PSOP24)	P. faiciparum 307	INTT	10516	22084	7026	11930	4433	8192	1559
conserved Plasmodium protein, unknown function	P. fairiparum 307	INTT	9640	5185	3996	4905	4324	8602	1686
bromodomain protein, putative	P. faiciparum 307	INTT	10079	6023	4109	4533	3054	6110	1287
teromodomain protein, putative	P. faiciparum 307	INTT	6320	8301	2935	5636	3687	8109	1348
ervthrocyte membrane protein 1, PfEMP1 (VAR)	P falciparum 307	INTT	5251	5338	2505	7931	4222	10155	1441
knotessociated histidinarich protein (KAHRP)	P falciparum 307	NTT	6382	9450	1985	3529	2228	6062	905
Plasmodum exported protein, unknown function	P. facciparum 307	INTE	7707	3071	2250	3903	3413	5579	1227
conserved Pasmodium protein, unknown function	P. faiciparum 307	INTT	5020	4101	2374	5704	4453	7304	1501
early transcribed membrane protein 2 (ETRAMP2)	P. falciparum 3O7	INTT	23580	19799	1423	5560	3844	7071	1271
protein kinase, putative	P falciparum 307	INTT	8829	2942	2448	5228	4630	9547	1804
conserved Plasmodium protein, unknown function	P. Neisiparum 307	INTT	10091	3990	4385	5720	4867	8076	1723
conserved Paerrodium crotein, unknown function	P teciperum 307	D/TT	19527	3478	7753	6073	4412	7938	2073
and a set of the second the Constant too the Constant too the Constant too the Constant too the	Traditor I would be an	a solution of the second second							
And the second s	and a real sector of sectors and	A LOUI DESCRIPTION OF ADDRESS (_

Now if you scroll down further you will find all your other spots basically your antigens which you were trying to study are all printed as IVTT spots, so basically this entire column M gives you details about the spot preparation.

Now the all other columns here are basically your patients samples, so if I just move this a little bit, what you will see here, for example let's consider the first sample, this is basically a positive control which was part of batch 1 set 1, slide 1, and pad 1, so let me again take you back to the experiment, this experiment was performed in 4 sets of 2 batches or rather two batches of 4 sets, so you have batch 1 set 1, batch 1 set 2, then you have batch 2 set 1 and batch 2 set 2, so basically what is this telling me? This is telling me that this particular positive control was probed and batch 1 set 1 on slide 1 and pad 1, right.

Now similarly so let's go to the next one which is a real sample that was just a positive control, so this is basically probe 1 batch 1 set 1, slide 1 and pad 2, so this is going to tell me my position of the samples, so if ever I want to go back to the slides and check the real spots, right, the images of the spots then I have no exactly where to go, so for example if some samples is not behaving well and I want to go and cross check the intensity of this spots, for example some sample is giving me very high intensity signals and I want to go back and check whether its real, then I'll know exactly which file to open, because I have all the details here, so this is for the all other columns.

So this is, I hope you now understood how the excel sheet looks, right, so now what we are going to do the next thing is we are going to apply a colour gradient to this excel, right, and now I'll tell you why we are going to apply the colour gradient, so let's first do that, so for which what I'm going to do is I'm going to go to conditional formatting, I'm going to go to colour scales, more rules, and I'm going to choose it three colour scale, and then I'm going to choose number type, in case I'm going to say 0, and I'm going to assume that my entire data

falls you know in between say certain negative values and may be around 80,000 is my maximum value, so I'm just going to assume that if my data falls in this range, I'm going to split my data based on three numbers 0, then my midpoint will be say 20,000, and my highest will be 40,000, and I'm going to choose some colours here, so I say this is maybe grey, then I'm going to keep this black, and I'm going to keep this red,

1	9.8	34 3	6 43	0.4		CIP 3	E + 2++ 75 + 📷 🖽	R 10	n- 0						Q.* Dearth in Sha	
-	La	rout	Tables	Chart		lenarit, kri	Formulas Data	Pestiaw								
10 10	ra -	Care B	1.5	11.000	-	A- A-		Nag Ted -	General 		Constraint	and	1000	and Neutral planetary	Barrier Deles Farmer	A-
11		: 0	0.0	/+ Batch	1.5et1						10-01-0					
	A	8	C	D	1.6	F		G			н		1	J		
														-		
			Arra							Ne	w Formatting Rule					
			У	Array	Spot	Spot										
sort		Index	Row	Column	Row	Column	ADI spot ID (gai file)	Dryte	3-Celor Scale					ORF_Fragment	Description	
	1	1	1	1	1	1	IgG mix 1								IgG mix 1	
	2	290	1	2	1	1	IgG mix 1		Minimum		Midpaint	Mastimu	-		IgG mix 1	
	3	579	1	3	1	1	lgG mix 1	7,04	s Number		Number D	Number			IgG mix 1	
	4	868	1	- 4	1	1	IgG mix 1					Conner .			IgG mix 1	
	5	2	1	1	1	2	lgG mix 2	Varue		34	20000	*0000	_		IgG mix 2	
	6	291	1	2	1	2	ligG mix 2	Color							IgG mix 2	
	7	580	1	3	1	2	lgG mix 2					Canada .			IgG mix 2	
	8	869	1	4	1	2	lgG mix 2					Cance			lgG mix 2	
		3	1	1	1	3	IgG mix 3	1			wa.	NA .	_		kgG mix 3	
	10	292	1	2	1	3	IgG mix 3			- ,	4/A	NA			lgG mix 3	
	11	581	1	3	1	3	IgG mix 3				A.V.A.	NA			lgG mix 3	
	12	870	1	4	1	3	lgG mix 3				A.VA.	NA			lgG mix 3	
	13	4	1	1	1	4	IgG mix 4				A.VA.	NA			lgG mix 4	
	14	293	1	2	1	4	lgG mix 4				A.VA.	NA			lgG mix 4	
	15	582	1	3	1	4	lgG mix 4				A/A	NA			lgG mix 4	
	15	871	1	4	1	4	lgG mix 4				4/A.	NA			lgG mix 4	
	17	5	1	1	1	5	IgG mix 5				A.VA.	NA			lgG mix 5	
	18	294	1	2	1		IgG mix 5				4/A	NA			kgG mix 5	
	19	583	1	3	1		lgG mix S				4/A	NA			kgG mix 5	
	20	872					lati mie 5				dia.	NIA			inG mix 5	
	20	+ 6	mained	Raw , Carri	sized by		mbred Ray (3) / Combined Re	w (6) _ 16	inus IVTT cantral 2 Lag	a roc	metadata _ Minus NTT o	antral - someth			inG mix 6	
and i																

(Refer Slide Time: 14:00)

so what this is going to do? Is all my values above 40,000 are going to be in dark red, and then around 20,000 will be black, and the lowest or the least values will be grey, and those which are in negatives will be almost white, so that's how I'm going to apply a colour gradient here, so you can see in the slide here, basically what I have done is I've just minimize this excel a little bit,

(Refer Slide Time: 14:22)

2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Q.+ (^ 0·
Image: Control (Control (Contro) (Contro) (Control (Contro) (Control (Control (Control (Contro) (Dem . Dem .	Trans
	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	COMPANY OF THE OWNER

so you will be able to see all four batches at once, so I don't know if you can see a black line here, so basically this is going to split your batches, so in fact it's going to split your set, set 1 from set 2 of the first batch, and set 1 from set 2 of the second batch, so basically this is batch 1 set 1, batch 1 set 2, batch 2 set 1 and batch 2 set 2, so when you minimize this excel a little bit and you apply this colour gradient what you can see? Is that the signal intensities particularly for this batch, in fact this whole batch, but batch 2 set 1 is really high compared to the rest of the batches.

So you will know that mainly from the IgG signals here, so this particular line which you see here are your IgG mix, and this particular line which you see here is your anti human IgG, so basically this is your control which is going to tell you whether you need to rescan your slide or not, so if this is very high then all your signals by default for this particular batch will also be high, right, so that is going to screw up your results a little, screw up your results later because all the patients in this batch are going to show high signals which will be, which is not correct, so this IgG mix printed on this chip is going to basically help you in deciding whether you need to rescan your slides at different PMT settings and PAR settings, right.

So what we will do here is we will rescan the slides once more, bring these settings down a little bit and bring, these settings not as low as this, but a little lower because this is also a little high compared to this if you see, right, so later on we realized this is because of the membrane thickness of the slides, there could be other issues also which you might encounter later, so to avoid this you need to first bring down the signals and then any changes after that will be corrected by normalization, okay.

(Refer Slide Time: 16:23)

What can be observed from this heat map is that the signal intensities of IgO control spots across different batches are

Heat map generated from data output files after rescan

What can be observed from this heat map is that the signal intensities of IgG control spots across different batches ar comparable

The minor differences will be adjusted by normalizing the data with 'No DNA controls' which will eliminate background signals

So now having rescanned all the slides as you can see in the slide, the settings look pretty uniform though a still not very uniform and you will still feel that batch 2 set 1 has higher signals, but overall it's okay because this will then be taken care of by normalization.

So now what we will do is we will proceed with normalization using excel, now there are two strategies I'm going to talk to you about today, the first strategy is basically a very simple normalization method which we will use only for visualization, for example if you want to prepare heat maps,

(Refer Slide Time: 16:51)

Data normalization

Two kinds of data normalization:

Sample specific median normalization: Each raw value is subtracted from the median of its 'No DNA controls'

2. Log2 transformed FOC- Each raw value is divided by the median of its 'No DNA controls' and Log2 transformed (Used for statistical analysis)- This is called fold-over-control (FOC) normalization which reduces the variation in signals that could potentially arise between probing operations performed at different times





then we will use this the first normalization method, however if you want to perform statistical tests, then we will use the second kind of normalization which I'll talk to you about.

So let's first go through the first normalization method, so what we are going to do in the first normalization is we are going to subtract the raw values for each of the IVTT spots from the samples specific medium value of the no DNA controls, so I'm sure that this is a really confusing, so what we'll do is we will go step by step, first I'm going to show you what raw values are and then I'm going to show you what the no DNA controls are, right, so again we are going to come back to the same excel, it is colour coded and we've reach the stage, you also know that this, now in this data we have IgG mix, we have anti human IgG, we have purified proteins, we don't need any of those right now for our analysis, we are going to directly go down to the IVTT spots, so in fact what we'll do is we will probably just delete those rows to avoid confusion.

So let's start from here, I'm going to delete the first few maybe what I will do is I'll just zoom this a little bit, so I have just zoomed this a little bit, what we are going to do is we are going to delete unwanted rows right now, so we don't want IgG mix, we don't want anti-human IgG, we don't want purified proteins right now.

Again let me tell you the purified proteins basically we don't require in the analysis, but it's important when for example your slide is not worked at all and or you have not got the signals you required, you can always go back to the positive control spots to see what the signals were, right, so this is basically used for such you know analysis just as controls, so right now we are going to delete those rows and we are going to only keep rows which are IVTT mix, right, that's what this is.

EACH FIG COI	view insert Pormal	nom una minore y nep			C \ 100 ()	0 Mon 5/0	Ar pm (
		No. The second s					
	0 C 🔮 🚥 • 🖽 •	X • 20 • 75 • 🙆 🗄 📾 🕬 😣			Q		
uma Layrut Tata	ies Charts SmartArt	Formulas Data Review					
Edit.	Fast.	Algeneri Burba	Famal		Carls .	Theres	
Fill * Calibri (be	ent - 11 - A- A-	- main aler Weg Set + Gerend -	Roomal Red Social No.	10 X X		Arts 1	- T
Our B /	U 🗠 - 🛆 -	· 문 문 권 4월 6월 _ Negs - 행 - % 3 옷 4월 4월 Constant	Catalanan Deck Left Exploredpry	NA IN	art Danata Formar	Theres A	La-
4 100	· /r 21	to and					
					Batch1_Set Ba	tch1_Set Ba	atch1_5
					Slide 1	Side 1	Slide 1
					1		. 3
				> Rav	v data (Co	mbine	2d
				raw)			
				10,00)			A
asmoD8_ID	ORF_Fragment	Description	Organiam	Preparation	Batch1_Set Ba	tch1_Set Ba	nch1_5
A	CIDR2	erythrocyte membrane protein 1, PfEMP1 (VAR)	P. falciparum	INTT	6799	5126	271
307_0102200	Exon 2 Segment 2	ringinfected erythrocyte surface antigen (RESA)	P. falciparum 307	INTT	13575	16454	267
F3D7_0102500	Exon 1 Segment 2	erythrocyte binding antigen181 (EBA181)	P. falciparum 3D7	INTT	21760	4278	473
F3D7_0103500	Exon 2 of 7	conserved Plasmodium protein, unknown function	P. falciparum 307	INTT	7816	15681	243
F3D7_0107300	Exon 2 of 2	probable protein, unknown function	P. falciparum 3D7	INTT	1952	1155	10
F3D7_0106300	Segment 1	conserved Plasmodium protein, unknown function	P. falciparum 3D7	INTT	15274	7058	70
F3D7_0106300	Segment 2	conserved Plasmodium protein, unknown function	P. falciparum 30/7	INTT	11082	6999	252
F3D7_0106300	Segment 3	conserved Plasmodium protein, unknown function	P. falciparum 307	INTT	16322	4601	796
F3D7_0106700	Exon 1 Segment 1	secreted ookinete protein, putative (PSOP24)	P: falciparum 307	INTT	16505	17858	324
3D7_0106700	Exon 1 Segment 2	secreted ookinete protein, putative (PSOP24)	P. falciparum 307	INTT	10516	22084	700
3D7_0110000	Exon 1 of 1	conserved Plasmodium protein, unknown function	P. falciparum 3D7	INTT	9640	5185	399
3D7_0110500	Exon 1 Segment 2	bromodomain protein, putative	P. falciparum 3D?	INTT	10079	6023	410
307_0110500	Exon 1 Segment 3	bromodornain protein, putative	P. faloiparum 3D7	INTT	6320	3301	299
307_0200100		erythrocyte membrane protein 1, PfEMP1 (VAR)	P. falciparum 307	INTT	5251	5338	250
907_0202000	Exon 2 Segment 1	knobassociated histidinerich protein (KAHRP)	P. falciparum 3D7	INTT	6382	9450	196
P3D7_0202200	Exon 2 of 2	Plasmodium exported protein, unknown function	P. faloiparum 307	INTT	7707	3021	225
507_0202400	Exon 1 Segment 2	conserved Plasmodium protein, unknown function	P. faloparum 3D7	IVTT	5020	4103	237
907_0202500	Exon 1 of 1	early transcribed membrane protein 2 (ETRAMP2)	P. feloiparum 3D/7	INTT	23540	19799	342
507_0203100	Exon 2 Segment 3	protein kinase, putative	P. faloiparum 3D/7	INTT	8829	3942	244
Compo	Farrys 1 of 1	Combined Raw (B) Combined Raw (4) Minus IVIT control / Logil FOC / metadota / Minus IV	P. fairinarum 9017	NTT	10091	2027	434

(Refer Slide Time: 18:45)

So now we are going to have this way 500 plasmodium falciparum IVTT spots, and 515 plasmodium vivax IVTT spots, so we are going to go down, you have deleted unwanted rows, there are few more rows below which we don't need, so after these 1015 spots (Refer Slide Time: 19:03)

Excel File Edit V	few insert Formal	t Tools Data Window 🛊 Help			Q 80% (t)	Mon 5:00 p	pm Q
••		C Bide	. without				
1 I I I I X II	E & 10-11	Σ • 20 • 35 • 🕼 🗄 🚓 τοτι 🗮 😣			Q	_	
A Hone Layout Tables	Charts Smarth	Art Formulas Data Review					
Let	Red	Algoriant Runtee	Farma		Galla	Thursda	
- B Fil + Calbr Both	6 1 1 1 A A	🚛 💷 📰 aler 🔂 Weg Tee + Garand 🔹 🛊	* Roomal Med Social No.	antal Xam .	B	Ar. 20	-
Q Cour - B / 1	U A-A	- 15 15 15 41 41 41 1 Marg - 10 - 56 + 94 42 Control	Catulation Constant Paperson In	e la		Dans Ar	-
4 1 0 0 /	4. 11	the second se			Dees rorat	119108 1100	
	J# 73				100		
15 PVX 122995	Exect of 1	drup/metabolite exporter, drup/metabolite (reparatorier	P view Sall	INTT	2020	3633	2014
15 PVY 123040	Exce 1 of 1	hunorhalical nucleis, conservat	P view Salt	DALL	4721	2726	2441
Z PVX 123105	Exce 2 of 6	hypothetical protein, conserved	P. vivas Salt	DALL	6232	5183	2210
B PVX 123350	Excent of 1 Sec	men homodomais postais, containe	P view Salt	DUTT	10064	4177	1212
9 PVX 123440	Excent of 2	hundhatical ontain, conserved	P viva Salt	INTT	1261	3668	1881
0 PVX 123505	Excent of t	hundhalical protein, conserved	P view Salt	DUTT	2245	1635	4313
1 PVX 123510	Exen 1 of 1	S4. putative	P view Salt	INTT		8778	4332
2 PVX 123520	Exert for t	DNAhinting changement sutative	P vives Salt	INTT	8377	1982	4429
3 PVX 123665	Exce 1 of 3 Sec	men hundhelical potein, conserved	P. vives Salt	INTT	5320	1771	1757
4 PVX 123705	court of a beg	hundhelicel protein, conserved	P view Salt	NTT	6164	3435	2035
5 PVX 123745	Excn 1 of 1	andopiasmin precursor, putative	P. vives Sal1	INTT		21592	2853
6 PVX 123810	Exon 2 of 2 Sec	men hunoihelical protein, conserved	P view Salt	INTT	1712	2163	1608
7 PVX 123845	LAURE OF LONG	networkensdatabiling northein, nutative	P vices Salt	INTT	2961	2907	2379
L8 PVX 123855	Excn 2 of 2	Chromatin assembly protein (ASE1), putative	P view Salt	INTT	5567	11544	2983
Q PVX 124015	Exce 2 of 3	hunorhatical protein, conserved	P vices Salt	INTT	10785	2129	6015
0 PVX 124140	Exon 10 of 11 S	eom hupothetical protein, conserved	P. vices Salt	INTT	4324	2604	2230
1 NA	200111011110	TTBS	N/A	TTBS	1018	1115	974
2 N/A		TTRS	N/A	TTBS	635	562	551
T NA		TTBS	NIA	TTBS	572	785	522
4 N/A		TTBS	N/A	TTBS	605	588	771
S NA		TTBS	NIA	TTBS	56	122	140
IS NA		TTRS	NA	TTBS	11	97	143
7 N/A		TTBS	N/A	TTBS	-14	212	179
S NA		TTBS	NA	TTRS	113	154	194
9 NA		TTRS	N/A	TTBS	782	1062	940
IO NA		TTBS	N/A	TTBS	260	917	712
11 N/A		TTBS	N/A	TTRS	304	743	458
12 N/A		TTBS	N/A	TTBS	118	391	518
		and a second sec					

there are few more like TTBS which is nothing but your buffer spot, where only buffer is spotted and then you have some empty spots data, then we have data for blank, so this is also unwanted we are going to delete that as well.

So now what we have are 1015 IVTT spots and 24 no DNA control spots, so now what are these no DNA control spots? (Refer Slide Time: 1:31)

	🐠 🐯 - 🖽 - 🗴	- 20 - Te - 100 - 100 H 10		9	 Conservation 		
A mone Layout Tables	Charle Smartart	Formulas Data Review					~ 0
ligh .	fast	Algered Series	Ternal		and the second second	Thursday	
	+ 10 + A+ A+	m m aler washer Galand	Rental Red South No.	···· 30 3		Ar. 20	-
20m - B / U	7. 4.4.	2 2 2 4 42 62 mm - 10 - 54 1 44 42 Contract	Catagories (Designation) (Sec.			man Apr	
No. 100.00	and a second sec	terretry	Accessed in the second se		and there	CORPORE THE	
		6		And and a subscription of the		0	
12 850 123526		Dillhinding chapagene subline	P view Salt	D/TT	8377	3082	4479
13 PVY 122655	Even 1 of 3 Second	kunometrical particle, posterior	P store Sait	DUTT	5220	1721	1757
14 PVX 123205	Excit i or 5 degrees	hypothesical protein, conserved	P. view Sait	INTT	6164	3495	3035
15 PVX 123745	Excent of 1	androllasmin process polative	P view Salt	NTT	0.000	TAXABLE INC.	2053
16 PVY 123810	Excert 2 of 2 Sectored	hunorihalizat outlate conserved	P view Salt	DATE I	1722	2163	1618
17 PVX 123845	cours or a only raise	anivarian visitating nonlain, no dative	E vices Salt	D/TT	7963	2907	2375
18 PVX 123855	Excep 2 of 2	Chromatin assembly protein (#SE1), putative	P view Salt	INTT	5567	THURSDAY.	2983
19 PVX 124015	Exm 2 of 3	hundrating protein growthing protein	P vices Salt	INTT	10285	2129	6035
20 PVX 124140	Exce 10 of 11 Secret	Nurvihalical protein, conserved	E view Salt	INTT	4324	2604	2210
TEL NA	exten to or the begins	acDNA Control	N/A	IncONA.	8826	3136	3439
22 NA		acDNA Control	N/A	mONA	8748	3389	3128
2 E N/A		ncDNA Control	N/A	mONA	8503	3026	2690
24 N/A		acDNA Control	N/A	noONA.	8971	2873	2636
25 NA		noDNA Control	N/A	mpDNA.	6453	2736	2627
25 N/A		noDNA Control	N/A	noONA.	9074	2941	2834
27 N/A		noDNA Control	N/A	moONA.	12212	6601	5098
28 NA		ncDNA Control	N/A	opDNA.	6580	3050	3151
29 N/A		neDNA Control	N/A	mONA.	6563	2632	2717
30 NA		noDNA Control	N/A	-Lavon	7721	2642	2722
ST N/A		ncDNA Control	N/A	TOONA	7247	2918	2915
32 N/A		noDNA Control	N/A	moONA.	7524	2716	2685
33 NA		ndDNA Control	N/A	A/Gon	7429	2824	2940
34 NA		ndDNA Control	N/A	moONA.	6554	2780	2872
ISS NA		ncONA Control	N/A	NOONA	6158	2435	2326
36 NA		ndDNA Control	N/A	APCON	6647	2518	2420
37 N/A		noDNA Control	NA	AVCon	5640	2521	2651
38 NA		noDNA Control	NA	AVCON	6277	2600	2774
39 NA		ncDNA Control	NA	NOONA	7913	2218	1533
Company of the st The Company Name	/ Combined Raw (2) / Com	times fam (3) Combined Ray (4) Minus NYT control Logi FOC metadate Minu	a NTT control - softwill T con	a	Contraction of the local division of the loc		
Surrey Line Bank		hand •			-		

So basically these spots have the entire IVTT mix except the plasmid, so basically what you expect here is no expression because you don't even have the plasmid here, whereas the IVTT spots have the entire IVTT machinery just like no DNA but they also have the plasmid where you are going to express your gene of interest, whereas you don't have that here, so what is this going to provide? This is going to provide your background signal, so what we are trying to do in the first type of normalization is we are subtracting a raw signals from background.

So now there are 24 such spots which you remember we have rearranged and that's why it's come together, grouped together like this, the first thing we are going to do is take a median of this which I have already provided you here, so this is the formula for it, I've just done this the whole thing in excel, so now we have a median value here, so the first thing what I'm going to do is for this particular sample which is in column N, (Refer Slide Time: 20:20)

	6 LO - CI -	- 2- T- 6 - M 100 - F - 2		(4	a familie t		
a Laund Tables Co	ante formatikat	Formation Data Review			-		
lait fa	4	Algenant Burriar	Faired		Carla	Thermos	
To + Caler Beat	11 - A- A-		The last last last		L. 1975.	100 B	
Omme B & HILL		of the lot of the line of the lot of the lot of	Lange Contraction of	• •	· · · · · ·	man a	
Com. DIVIN		DETATION NOTABLE TRADA . CALLARY CALLAR	Terrating	inseri (leiete Format	Themes /1	
11 100 - / 8	826				_		
		K	L.	M		0	P
494		neDNA Control	NA	NOONA	88.26	3135	343
494		ncONA Control	NA	NOONA	8748	3389	312
494		neDNA Control	NA	noONA	8503	3025	209
494		neDNA Control	NA	AVGon	8971	2873	263
404.		ncDNA Control	N/A	ArGon	8483	2736	262
404		ncDNA Control	NA	A/Gon	9074	2941	283
494		noDNA Control	NA	ArGon	12212	6601	505
414		ncDNA Control	N/A	ArGon	6580	3050	315
4V4		noDNA Control	NA	NOONA	6563	2632	271
N/4		noDNA Control	N/A	NOONA	7721	2642	27.
414		noDNA Control	NA	NOCINA	7247	2918	29
N/N		neDNA Control	NA	NOONA	7524	2716	26
N1/4		ncDNA Control	NA	noONA	7429	2824	29
4/A		noDNA Control	NA	noONA	6554	2780	283
N/A.		neDNA Control	NA	noONA	6158	2435	233
4/4		noDNA Control	NA	A//Con	6647	2518	24
AVA.		noDNA Control	NA	NOONA	5640	2521	26
N/A.		neDNA Control	NA	noONA	6277	2600	27
N/M		neDNA Control	NA	NOONA	7913	2218	15
404.		noDNA Control	NA	noONA.	8021	3684	380
N2144.		noDNA Control	NA	noONA.	8162	2288	212
N/04		noDNA Control	NA	NOONA	8687	2397	181
N/M.		noDNA Control	N/A	noDNA.	7771	2325	225
A/A		neDNA Control	N/A	A/Gon	10094	2593	24
				Median No DNA	7842	2726	2703
	Combined Ray (2) J G	mbined Raw (3) Combined Raw (4) Minus IVIT control J Logil FOC J met	talata J Minus //TT campal - sometil				
Normal View Basely		347-108279					_

this is one sample, I'm going to subtract the values for each and every IVTT spot from that particular medium, right, so each of this spots which you see here, I'm going to subtract it and that is what is my median sample specific median normalization.

So let's scroll, so probably we will do this in the same excel, okay I have kept place for that here, (Refer Slide Time: 20:50)

	# 1.04 V	new meet	Formal	cools Clat	 Window 	9 14						201	100	mater N	ON \$100 DA	P 4
						-	Side! whe									
目の日に	🗯 💥 U	0.44	a+01-)	E + 20 + 3	下・国に	1.60 199							Q-			
Mante Log	eut Table	a Charle	SmartArt	Formulas	Date	Realize										~ 5
640		Rand			Algemant		Artse		Farmat				Garle		Therea	
. In the other	Calibri (Bea	bl [#]11	- A- A-	10.00	aber.	West Test -	Gerend + M.	Rental	Ref	last	Bautid	1	in . 10	(A	. E.	
O Der.	8.1	u	4 - A -	10.00.2	41.41	intern 1	Han - the a fight of Commons	Calculation	Constanting of the	Lugherselary .	Papel			G	A	
	1. C. C. (and the second by		1901100110			Familie Familie			_		-	ser years	10.00		
10	VM.	VN I	10	10	KO	110	85		Distance of the local	1011	N.V.	L MW	N.Y.	NV.	1.17	
Rulph 7 Call	Butch? Set	Batch3 Set	Batch2 Set2	h/r	n sz	0,0	1.5		Batch1 Set I	Batch1 fast	Batch1 Set	PLAY Ellabolis 1 Kar	n.n.	Ratch1 Ca	Ratch1 C	an Ban-
Cline 6	Slide G	Girla G	Side 9						Side 1	Side 1	Slide 1	Clide 1	Glide 1	Clinia 1	Slide 1	61
Dat 5	Pad 6	Pad 7	Pad 8						Bad 1	Red 1	Rad 3	Paul A	Part 5	End 6	Bad 7	
Pieza	190.0	Pager							- 90 A	180.2	- 80 S	780 4	780.2	P 80 0	190.7	
	ind.	-	m.				> IVTT spots minu	15								
-							- it is possible		Position constru	BACK	IMIA	Dest	EHOC.	MUT	ABUT	late.
							median of IVTT cont	trol	-unite comby	rans t	Cartha	0.01	reas	100-51 L	7670	parce 1
8252_9_5	8252_9_6	8252_9_7	8252_9_8				ADI spot ID (gai file)		Batch1_Set 8	Batch1_Set	Batch1_Set	t Batch1_Se	t Batch1_Se	t Batch1_Se	(Batch1_S	et Batc
10901	15193	2769	4681				PFL_0008_CIDR2		52							
10197		3685	7068				PFA0110we2s2									
9718	12541	3287	5495				PFAG125ce1s2									
8489	12104	2988	4389				PFA0175w_2o7									
4678	7181	1786	2967				PFA03601_202									
4678 9112	7181	1785	2967				PFA0360c_2o2 PFA0410w-s1									
4678 9112 7402	7181 13015 10985	1786 7520 2553	2967 4395				PFA0360c_2o2 PFA0410w-s1 PFA0410w-s2									
4678 9112 7402 7161	7181 13085 10385 7265	1786 7520 2553 3502	2967 4395 8052				PFA0360x_2x2 PFA0410w-s1 PFA0410w-s2 PFA0410w-s3									
4678 9112 7402 7161 7192	7181 13015 10385 7265 11126	1786 7520 2553 3502 3714	2967 4395 8052 4550				PFA0360c_2o2 PFA0410w=1 PFA0410w=52 PFA0410w=53 PFA0410w=53 PFA0430ce111									
4678 9112 7402 7161 7192 86400	7181 13015 10385 7255 11126 10454	1786 7520 2553 3502 3714 2725	2967 4395 8052 4550 7147				PFA0360c_2o2 PFA0410w-s1 PFA0410w-s2 PFA0410w-s3 PFA0430ce111 PFA0430ce112									
4678 9112 7402 7161 7192 86400 5329	7181 13015 10385 7265 11126 10464 10108	1786 7520 2553 3502 3714 2726 2122	2967 4395 8052 4650 7147 6369				PFA01904, 202 PFA01104-1 PFA01104-2 PFA01104-2 PFA04304-1 PFA04304-1 PFA04304-1 PFA04304-1									
4678 9112 7402 7161 7192 86400 5329 5519	7181 13015 10385 7265 11126 10464 10108 9697	1786 7528 2553 3502 3714 2726 2122 3006	2967 4395 8052 4650 7147 6369 9822				PFA0BI00;2:02 PFA0BI00=:01 PFA0BI00=:02 PFA0BI00=:02 PFA0BI00=:10 PFA0BI00=:101 PFA0BI00=:102 PFA0BI00=:102									
4678 9112 7402 7181 7192 84420 5329 5519 6448	7181 11015 10985 7265 11128 10464 10464 10108 9697 11459	1786 7520 2553 3502 3714 2725 2122 3006 2529	2967 4395 8052 4650 7147 6369 9822 4549				PFA0380c_202 PFA0410w=1 PFA0410w=2 PFA0410w=3 PFA0430ce11 PFA0430ce112 PFA0430ce112 PFA0430ce112 PFA0510we113									
4678 9112 7402 7161 7192 85429 5329 5519 6448 9929	7181 13015 10385 7265 11128 10464 10108 9697 11459	1786 7820 2553 3502 3714 2726 2122 3006 2520 3620	2967 4395 8052 4650 7147 6309 9822 4549 6523				PFA0580c_2h02 PFA0580c=1 PFA0580c=3 PFA0580c=13 PFA0580ce313 PFA0580ce313 PFA0580ce313 PFA0580ce313 PFA05810we133 PFA05810we133 PFA05810we133									
4678 9112 7402 7161 7192 5329 5519 6448 9929 7533	7181 13015 10385 7265 11128 10464 10108 9697 11459 20889 10441	17880 78800 2553 3502 3714 2726 2122 3006 2520 3620 2554	2967 4395 8052 4550 7147 6369 9922 4549 6521 4481				PFA0800_202 PFA0810w-31 PFA0810w-32 PFA0810w-32 PFA08100w-33 PFA08000w-101 PFA08000w-101 PFA09100w-102 PFA09100w-102 PFA09100w-102 PFA09100c211									
4678 9112 7402 7161 7192 8427 8519 6448 9979 7533 8509	7181 13035 10385 7255 11124 10454 10108 9607 11459 10941 10941	1788 7820 2553 3502 3714 2726 2122 3006 2520 3620 2554 3226	2967 4395 8052 4650 7147 6369 9922 4549 6521 4549 6521 4481 4073				PFA0380c_2o2 PFA0410w=1 PFA0410w=2 PFA0410w=3 PFA0430ce11 PFA0430ce112 PFA0430ce112 PFA0430ce112 PFA0430ce113 PFA0510we113 PF805106v(renamed) PF805106v(renamed) PF805106v(renamed) PF805106v(renamed)									
4678 8112 7402 7163 7192 5329 5519 6448 9979 7533 8569 6525	7181 11015 10185 7265 11124 10464 10108 9607 11453 15960 10941 14987 5813	1788 7828 2553 3502 3714 2725 2122 3006 2520 3620 2554 3225 3225 3225 3225	2967 4395 8052 4650 7147 6369 9922 4549 6523 4481 4071 4153				PFA08100_202 PFA08100=1 PFA08100=2 PFA08100=3 PFA08100e13 PFA0800c111 PFA0800c111 PFA08000_101 PFA08100e133 PFA08100e133 PF000100e131 PF000100e132 PF000100e132 PF000100e132									
4678 8112 7402 7161 5329 5519 6448 9979 7533 8569 6521 9934	7181 13035 10385 7265 13128 10464 10064 9697 13458 10441 10441 10441 10441	1788 7828 2553 3502 3714 2725 2122 3006 2520 3629 2554 3226 2554 3226 2545 3183	2967 4395 8052 4650 7147 6369 9922 4549 6521 4481 4071 4153 6473				PFA08100_202 PFA08100=1 PFA08100=2 PFA08100=3 PFA08100=10 PFA083000_101 PFA083000_101 PFA0831000113 PFA081000(renumed) PF6011000(renumed) PF6011000122 PF6011200_101									
4678 9332 7402 7192 8549 5519 6448 9979 7533 8569 6525 9934 7372	7181 13015 10385 7255 13128 10464 10008 9697 11859 158600 158600 158600 158600 15860000000000000000000	1788 7828 2553 3502 3714 2725 2132 3006 2520 3620 2554 3226 3254 3226 3183 3306	2967 4395 8052 4650 7147 6369 9822 4549 6521 4481 4071 4153 6473 4727				PFA0580c_302 PFA0510v=1 PFA0510v=3 PFA0510v=3 PFA0580ce31 PFA0580ce31 PFA0580ve133 PFA0580ve133 PFA0580ve133 PF60010ce31 PF600150ce33									
4678 9112 7402 7181 7192 5329 5519 6448 9979 6548 8569 6523 9934 8569 6523 9934	7381 13015 10385 7265 10464 10464 10464 10464 10441 10441 10441 10441 10441 10441	1788 7829 2553 3502 3714 2725 2122 3006 2520 3629 2554 3226 2554 3226 2554 3226 3183 31306 3356	2967 4395 8052 4650 7147 6369 9822 4549 6521 4481 4071 4153 6473 4727 4895	which far 0	Continue	ter 10 . 10	PFA0380c_202 PFA0510w-31 PFA0510w-32 PFA0510w-32 PFA0510w-33 PFA0580ce1s1 PFA0580ce1s1 PFA0510we1s2 PFA0510we1s2 PFA0510we1s2 PF80510ce2s1 PF80510ce3s2 PF80510w-1s2 PF80500w-1s2 PF80500w-1s2 PF80500w-1s2 PF8050	A 1977 Assessed to a								
4678 9113 74022 7151 7192 5519 5519 6448 9979 75519 6448 9979 75519 6448 9979 75519 6448 9979 75519 6448 9979 75519 755519 75519 75519 75519 75519 755519 755519 7	7381 13015 10385 72955 11126 10464 12008 9697 11459 25880 10441 14556 38133 10441 14556 38133 10441 14556 38133 10441	1786 7525 35512 35512 3754 2725 2122 3006 2554 3620 2554 3620 2554 3126 3130 3130 3140 3140 3140 3140 3140 3140	2967 4395 8052 4650 7147 6359 9922 4549 6523 4681 4071 4153 6473 4071 4153 6473 4153 6473	etinal ka ()	Combined	tes 10 / 10	PFA05800_202 PFA05800=32 PFA05800=32 PFA05800=32 PFA058000_101 PFA058000_101 PFA058000_101 PFA058000100 (renamed) PF601300c_202 PF601000c_202 PF601000c_202 PF601000c_202 PF60	a NT carrol - a	anagi j							

so this is called IVTT spots minus median of IVTT control, so that's exactly what we have to do, we are going to say is equal to, then we are going to go to that particular spot, so say let's take the first patient, and then we will see minus and we go to the median value which is 7842, so now because I want this row to remain constant throughout, I'm going to put it dollar sign in front of the row, so this is what we get here, and now I'm just going to drag this across as well as down, so once you drag and drop,

Kerer Sinde Time: 21:36)	## 😱 @ \$ 📢 🖓 ###(10) Man \$00 pm Q ::::
□ Didet.eine □ Di H → ×, (□ □ ⊕ 00 + 01 + χ + 20 + 37 + 20 □ H → 01 → 10 00 00 00 00 00 00 00 00 00 00 00 00	Q
Norm Norm <t< th=""><th>The set of the set of</th></t<>	The set of

this is the kind of excel you get, I have just minimized this, but if you apply a colour gradient this is how it looks overall, so this is what you can use now to make your heat maps and what I have also done is that I have sorted this based on the antigens as well as the patients who were falciparum positive and vivax positive, I've split them completely, and I have also made another excel sheet based on age, you can also split them based on age, so this is how I have sorted them.

(Refer Slide Time: 22:28)



So I have put all your PV positive patients together and PF positive together, and I have also sorted based on age, so you have PV positive, PF positive as well as sorted by age, so this way you can sort your excellent different ways, you can also use other software to make your heat map, but basically this kind of, once you normalize it in this way you do not perform any statistical analysis with this data.

First statistical analysis I'm going to now show you the next normalization method which is your log 2 transform fold over control normalization, so for this I'm not going to show you the entire method again because now you know how to do it on excel I'm sure you all know, for this I'm only going to show you the steps, the first thing what we do is we are going to setup floor of 100, so what we are trying to say here is that all the samples which are below 100 is going to have a value of 100, so this is going to remove all my negative values from my data, so that's the first thing and I have done it here for you and we are going to keep scrolling right.

The next step what we are going to do is to divide each and every raw value by the median of the IVTT control spots, so just like how we did previously we subtracted raw values from the median of the IVTT control spots, this time we are going to divide it so that's what is called fold over control.

So once you setup floor of 100 then you divided, and the next thing you are going to do is to convert this whole data into log values, so your log to transform this entire data, right, and that's why it's called log 2 fold over control, so once you do this, this data can be used for any statistical analysis, so this because this normalization is known to be more stringent, okay.

So now either you can use programming to do your statistical analysis or you can use different softwares which provide you statistical test, but what you need to know is which type of test you need to use which is beyond the scope of this lecture, but you can always read about what

you want to do, and you can also decide on which software you can, you want to use, for example graphpad prism is an excellent tool for preparing for graphs, it also helps you in a lot of statistical analysis, but if your data is really huge like the one we have is not very huge, but still it is not very small for graphpad prism, so for example you can have graphpad prism can get stuck in the middle if you are using data if even this size, so of course if you have bigger data sets then it's very difficult to use softwares like graphpad prism.

However if you are going to have only 10 patients or 20 patients with 40 proteins or something, graphpad prism does offer you a lot of statistical test. Apart from that there are other softwares as well, you have metabol analyst though it is metabolomics data, you still have the module called significance analysis or microarrays in it which you can explore for your microarray data analysis, but there are of course our programming and python and other things will definitely be much better for your analysis, as you will save a lot of time as well.

So what I'm going to do towards this, we're coming to the end of the lecture, I'm only going to show you very small analysis you do an, I have done on excel, so basically what my aim here is to identify most zero reactive proteins from my chip, which means that the proteins which elicit the maximum antibody response in malaria patients right, so that is my aim.

So now just to get this whole list of best zero reactive proteins, you can also do this on excel using a particular formula which I will show you now, so let's go back to the excel, this is how our excel was,



(Refer Slide Time: 26:16)

right, before we removed all of this rows which are IgG and anti-human IgG, so greeting them for now, and probably zoom this a little bit.

So now if you see that I have retained all the rows, so the first thing what we need to do of course we don't need this, but I have still retain the entire sheet from the beginning you will see that there are these 4 patients which are deliberately kept out of the analysis, (Refer Slide Time: 26:40)



for example there are so if you see there is PF + PV everywhere right, so basically these are my patients who were diagnosed with mixed infection, so I don't want any such patients in my analysis so I'm going to purely have groups which are plasmodium falciparum and plasmodium vivax, and I'm going to look at, look for their response to plasmodium falciparum antigens and plasmodium vivax respectively, so I'm not going to have any of these mixed patients, so I have kept them out, so if you want we can also delete them, right.

So we need to basically start from row number 82 that's what we are interested in, because these are the IVTT spots, so the first thing I'm going to do is I'm going to an average for each and every spot, so example let's write here average, and I'm going to say is equal to, so I get an average value and I'm going to just drag this down, so you will have an average for each and every spot for all the patients.

(Refer Slide Time: 28:08)



What I missed telling you before is that the previous sheet had many more columns here, that's because we had a lot more samples which are probed on the chips, for example we had positive controls which were nothing but samples from taken from place which is a highly malaria endemic region, so you know that those spots have to give you a signal, right, so those are my positive control samples. So don't get confused between positive control samples and positive control spot, they are totally different, so these positive control samples I have excluded them from this particular analysis,

(Refer Slide Time: 28:42)



Checking reproducibility across the batches

- In order to confirm minimal inter-batch variability, samples probed earlier were re-probed on different sets of slides
- Scatter plots of signal intensities of samples probed at different times on different slides indicates uniformity across different batches

we also had healthy controls which are basically malaria nyu individuals, means patients were not detected with malaria at the time of admission, so they were malaria negative, those patients were also taken a probe on the chip just to see there is a difference in response, such patients have also removed from the analysis.

There is also certain samples which I have probed repeatedly in probably in duplicates or four times in all, you know once in all the sets, just to check for reproducibility, so here is some scatter plots you can see where I'm showing patient to patient reproducibility, so basically I'm showing reproducibility between my batch runs, so all of these patients also have removed from the analysis.

I have basically now in this excel 200 patients, 100 plasmodium vivax and 96 plasmodium falciparum patients, and four which are mixed infection also have removed, so in this way you can choose to remove rows and columns based on what you want to study and you can make a excel less complicated, right, so that's what I have missed mentioning, but now that is done, so I have taken an average right now.

(Refer Slide Time: 29:59)

Avg Raw value/antigen > Avg [Mean (No DNA controls) + 2SD (Mean)] = Seroreactive antigens

(

Now I am going to apply this particular formula which you see here, if the average value for particular spot is more than twice the standard deviation of the mean of the no DNA controls then that particular spot, then that particular antigen is zero reactive.

So what this means is that if this is the average if this particular number is greater than this particular number which I'm going to show you right now, if you take an average of the mean + 2 times standard deviation of the no DNA spots, this is my number, so if that raw value or if any raw value is greater than this value then that spot is basically zero reactive, then that antigen is, sorry then that antigen is basically zero reactive.

So I'm going to say if this is equal to, if function this spot is greater than this, then one is zero, so I get 1 here, and what if finally get is an excel sheet like this where I have random ones and zeros, right, so all of this ones I'm going to now say are my zero reactive proteins because they are greater than twice the standard deviation of my control spots.

Now a lot of people may also use healthy control of their analysis, right, but we don't have them, so what they do is they compare the signal intensities in a malaria group versus a healthy population, but since we don't have all that I'm going to simply say that this is my, these are the list of my zero reactive proteins which I can now take forward for further analysis, so this is not a great, this is not a statistical test, this is only shortlisting my proteins and I'm here only shortlisting my proteins from 1500 to handful which I can then take forward in study.

So this is what that sheet is, now what I have done here is that I have taken this for a group of patients, but now what if I want to check this for a particular patients, so that is what is my antibody breath which you will see here, I have done this individually for every single patient maybe I'll zoom this a little bit, (Refer Slide Time: 32:50)

(Neter blue Thile, 52.50)																				
Excel	Fie Edit View	Peen!	Format T	loois Data	Window	👂 Help								- 18 G	2 @ \$	191	00% BB 55	on 5:44 pr	n Q 🔳	
Implementation of the second secon																				
P. 10 07 E	1 目的日本 X D C 4 D - C - I - S- T - 図 5 巻 m H - P														Q.					
		-	-	-																
	Layeur russes	Course .	Bearler.	- Partnards		Parties .						-					-			
101.00				-			Constant and		1 1991			_	-				100			
- (g) FR	· Celer (freig)	14144	A. A.		6 (St. 1) (St. 1)	wap he -	General			Rennel	Bef		-		Canadation		2 M . D.	• 11 •	A4	
Pasta Q Des	* B / U	100	A	1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1	43.43	Narge -	15 - 56	1 24 23	Conditional	Effective of			evit				inset Deate	Farmer	Theres All-	
And A D D C A																				
1 04		1144	100	1943	740 7	0 04		07	fiel 1	794	0.0	w	1 199	14	1.00	10	175	10	10	
103	100104-08	1.000	104		Ser. Land	A		A 1	100									-		
10.5	Change a												- 2							
104	111104 01 8															- 1		- 2		
10.0	11080.14		0														0	- 2-		
111	LINES N. B		0					0 0		1					1 1		0	- 2	1	
107	11006.34 8		0			5	2	0 0	0						0 1		a	- 2		
104 Column	12334.75 1		0		0	1		8 0		1			1	1	1 1		0			
100	11406.34 2		0				1	0 1					-	0	0 5				1	
110	6997.3 5		0		0		0	1 0	0.	0	1	0	1	0	0 0		0			
114 2000	11476.34 8		0		0		0	0 0	0.	0		0	1	0	0 5	0	0		0	
112 June 1	9675.55 8		0		0		0	0 0	0	0	0	0	0.	0	0 0	4	0	- 1	0	
110	9834.37 8		0		0		0	0 0	0	1	0	0	1		0 0		0		0.	
114	8124.01 5		0		1		1	0 0	1.	0	0.	0	0	0	0 0	0	0	1.	0	
£36946 11 (12 (ASAL) E		0		0		0	0 0	0.		0	0		1	1 0		0		1	
118 4896.5	A096.33 E		0		0	0	10.	¢ 0		2				0	0 0		0	- E	0	
111	3349.08 1		0		0		0	0 0	1	1	0	0		0	1 0		0		1	
110 4000.8	JM254 E		0		0		0	0 0							0 1		0		0	
110 10000	8126.81 1		0		0	1	0	0 0	0					0	1 0		0		1	
110-1888.8	3081.78 1		0		0		0	0 3						0	1. 0		0		8	
123 39853	9980.51 2		0		0		0	e 0	0		9			0	1 0		0			
122 3896.5	7010.04		0			-	0	1 0							1 1		0		0	
1111	1091.34 8		0					0 0									0	- 2-		
	3898.72 2		0				0								1 0		0	- 2-		
	1101.01																0	- 2		
	1385.33												- 2					- 2-		
126	5792.bd						0	a a					- 2		1 0		0	- 2		
3.24 million in	a306.17 E							8 0							8 0		0			
110 1000	4630.26 2		0				0	0 0				1		0	0 0		0		0	
111 000000	4975.18 0		0				0	0 0	0						0 5		0		3	
152 84293	4095.18 0		0				0	6 D	0	1				0	0 0		0			
115 - 4004 8	6286.55 0		0		0		0	a 0	0			0	0	0	2 5		1			
114	4121.23 5		0		0		0	0 0	0	0		0			0 0	. 0	0		1	
	a al Minus MTT or	annal - Pa J	Minus MIT Last	and - for the last	Individual Prop	aliant analy 12	Pr Antibada	Areaster (T)	Investment of the	and a statements and the state	/ for select	-	121	-	-	- 1		-		
States and	tion Ready						-	Sam-0			and the second se								10	
COLUMN STREET			-			able		100	-	0	00	0.0		- 14		0	States and states and	- 550	And in case of the local division of the loc	
		1		7 .		F	8 8) [*/				T (2)	7	2		10.00		I and	198		
		-		-					-		1000						A 4			

so what you see here is that I zoom this for every single patient so basically the previous one which I showed you was the average for a single spot for a group of patients, as well as the average for the no DNA control, here I have done it for each patient which means it is sample specific, right, so in this way if I scroll down what you will get here this, you will know the number of zero reactive antigens per patient, which means that if one patient, for example here is zero reactive to only 12 antigens, whereas there are some other patients which is zero reactive to 77 antigens, so this is basically my antibody breath.

So these are the two basic kind of analysis which I can show you in excel for now, power if microarray technology is basically the fact that you can perform this experiment very fast, probably in a day or two and then using any kind of patient data, all you need to do is map this

whole data which microarray data to each and every patient, clinical information that you have, and then you can perform any kind of statistical analysis and you can generate several results from the same single experiment, so that's the beauty of this.

I hope you have got a glimpse of how to perform data analysis and how basic statistics can be done, and how this is not the only way to do statistics at all, you can do use software and programming and I will still recommend that people do programming because if you want, even a single small change you don't have to repeat the entire analysis, also tomorrow if somebody provides you some other clinical information of the same patient population you don't have to repeat the analysis in excel, you can simply write a code for it and then in a few minutes you will get results for that as well, so that's all for now. Thank you. (Refer Slide Time: 35:00)

Points to Ponder

- Basic microarray data can be analysed using excel, however programming is highly recommended for larger data analysis
- R programming is a simple language and has been very useful for biologists
- Two types of normalization was used for the analysis: Sample specific median normalization is used for the purpose of visualization, while Log2 FOC is used for statistical analysis because it is more stringent

MOOC-NPTEL

IIT Bombay

Points to Ponder

- The use of control samples and control spots are very important for any microarray data
- IgG positive controls are used for checking the overall experiment performance. In some cases, rescanning at different PMT settings is recommended
- Multiple softwares could be explored for data analysis as well, in case programming is not an option

MOOC-NPTEL

IIT Bombay

Sanjeeva Srivastava: After going through this demonstration session and the insights of doing microarray based data analysis, you must have realized that there are many ways of analyzing and representing microarray data, of course there is no single way, no correct way of telling you what is the best way of doing it analysis, then many considerations you have to keep in mind when you are thinking about how to make meaningful information out of this high-throughput data.

There are several questions that can be answered using microarray data provided your data passes, the quality control chips and it is properly normalized, in such experiments you will control features becomes very crucial, both the positive controls and negative controls they guide you about how accurate and reload the data is, they could distinguish between real signals and background noise after proper analysis methods.

In the next class you will see another application of protein microarrays in a different application they will shift the gears to the cancer research and also the platforms. So far we have talked about self-array expression based protein microarray platform, we will now talk about how to take purified proteins printed on the chip using human proteome arrays and then apply those to investigate a deadly disease cancer, and try to talk to you about both excremental demonstrations as well as the theoretical concepts involved in performing such biological experiments. See you in lecture, thank you.

(Refer Slide Time: 36:53)

Next lecture....

Applications of protein microarrays in Cancer Research-I

MOOC-NPTEL

IIT Bombay

Prof. Sridhar Iyer

NPTEL Principal Investigator & Head CDEEP, IIT Bombay

Tushar R. Deshpande Sr. Project Technical Assistant

Amin B. Shaikh Sr. Project Technical Assistant

Vijay A. Kedare Project Technical Assistant Ravi. D Paswan Project Attendant

Teaching Assistants

Apoorva Venkatesh

Nikita Gahoi

Shalini Aggarwal

Bharati Sakpal Project Manager

Bharati Sarang

Project Research Associate

Nisha Thakur Sr. Project Technical Assistant Vinayak Raut Project Assistant

Copyright NPTEL CDEEP, IIT Bombay