Introduction to Proteogenomics Dr. Sanjeeva Srivastava Dr. Bing Zhang Department of Biosciences and Bioengineering Baylor College of Medicine Indian Institute of Technology, Bombay

Lecture – 59 Linked Omics (Part II)

Welcome, to MOOC course on Introduction to Proteogenomics. In the previous lecture, Dr. Bing Zhang provided you the capabilities of an online tool Linked Omics. In today's lecture, you will be exposed to the various steps involved in analyzing large dataset using linked omics. So, let us welcome Dr. Bing Zhang for his last lecture.

(Refer Slide Time: 00:51)

InkadOmice				
And the bulk of the second s	across 12 curves types			
New Andrews a Select Consor Type - Second Research property for	1717-5 Select Contair Color	at Store Starth Galacest Starth Attitude	Bataset Street Method	
sta. Nutation detergen	BIBPH : BOLECT CANCER CO HOAT D			F
 Separation (Separation) 	diar 6 Eartin		surveys in the second second second	
+ Select Search attribute, fur		Canow Tape	∧ Q Bengin suffert	<u>c</u>
+ Select Target detauet, for	Acute Mysical Legisence (LAML)		TOGA, LANK	
(RWard) delatrice	Advected laterary (ICC)		TOGA, AGE	
a Scient Statistical Method	Balder ponate ovincere (RCA)		700A, BLGA	
a case prents prost	Burtow Cale Clarat.00		100A,100	
Antippet Ann. by	Read Property Street of The		TODAJNEA	
man's Protect fails that news	Same 1 h i d finster		Terrier 1 2 3 4 5 8 Med	
r Noran	Direct Land Addates			-
	Albertatura			
	II. Anna trailion of \$1.7 and howerd	MICOLOGICAL STREET, Michael South Saltanes Burnans Central	JALLEL JOYS RECEIPS URsensby Columby of Note Galeria	
	210-bitro Report University Diseasing of August Constants	NDAMO Andensin	PhNs, Pacify Northwest National Lans	
	L'HE LINUNSTRUM NUM CANONA	VL-Mediator (intensity	WEIMA Namburgton University School of Medicine	
	Ciri-Cirvos surs	läh mäkäumina läänima Josajane mikitek Besumisessi	Selectures their	
	Piday RNA-Surran Hides 2001 RNA Socializing	Må nifklanna filter 200 nifbit Sepenary	NDA MMAALD Anderson Reverse Presse Presse Presse Come	
	Manufi Aurena Infraso Auren Obd. Maria alas 27	Methodol & Lances (Infraum Harran Oblin Methodol HSD	Dilace froms G laatine LC MEMO/ 4ND	
	DV*12-Allynamic Devenantility Survey DVP	Wester Therest (270 Dittinus Vester (2445445	West & Office, Terror UTO Distances & Terror & Eventee rybrid Summare.	26110

So, let us go to the new analysis and then there a few steps you need to follow in order to perform this analysis let us start with proteomics data. So, basically we want to ask which proteins in ovarian cancer associated with poor prognosis and which proteins associated with poor prognosis in ovarian cancer. In order to do this first you have to identify the ovarian cancer data set to do the analysis on ovarian cohort. So, you can browse, but let us do a search bearing.

(Refer Slide Time: 01:35)

Barbard and State State Barbard and State Barbarba	TD Anastinani 53 Publied TM Anarose Amulto	Established Back State	Barle 🧔 Bangin Fra	ne () in ()	we three	C 2344	8 Amanun (E NIT E	epere		Citer Incommuni
Implication Impl	or? Passil (b): 21a miles Part.	Browny 1 16 5 17 42 within					Pendent 1	2.3	4 5	3 100	
Norm State Description											
Balay A Balay A Balay A Desker D Pailor D Pailor D 1000, DV 800 MMASA 100-41 900000 8000000 00000000000 000000000000000000000000000000000000										Bear Ch	
TODA, CV B0000 with Adam 16 with P162114 B Grave Produes, TMR4 spl TODA, CV B0000 B0000 B0000 D0140 D015114 B Frade Produes, TMR4 spl TODA, CV B0000 B0000 D0140 D016114 B Frade Produes, STR000, Strained TODA, CV B B004 D016104 D016114 B Frade Produes, STR000, Strained TODA, CV B B004 D016104 D016104 B Frade Frade STR000, Strained TODA, CV B B004 D016104 D016104 B Grade Frade STR000, Strained TODA, CV B B004 D01610 D016104 B Frade STR000, Strained TODA, CV B B004 D01610 D016104 B Frade STR000, Strained TODA, CV B B004 D01610 D016104 B Frade STR000, Strained TODA, CV B B004 D016104 D016104 B Frade STR000, Strained TODA, CV B STR000 Strained Strain		Ballar A Barrate O buttate	O Date type	0 Peters	Ú Pele	0 милия	C Analysia tevel	0	Pipelins	0	
TODA, OP B 500A 2004 B P04214 B Proved Proved <td< td=""><td></td><td>TODA. OV BODGC</td><td>se@MdSeq</td><td>16+ill</td><td>01080219</td><td></td><td>Germ</td><td>Frahman N</td><td>Star NDM</td><td></td><td></td></td<>		TODA. OV BODGC	se@MdSeq	16+ill	01080219		Germ	Frahman N	Star NDM		
TODA, CP B DOA DP46 PEBITIN Peak Peak Peak Peak Peak Peak TODA, CP D D TODA, CP B DOA DP46 PEBITIN B Conv Peak EXECUTION TODA, CP B DDA DP46 PEBITIN B Conv Peak EXECUTION TODA, CP B DDA DP46 PEBITIN B Conv Peak EXECUTION TODA, CP B DDA DP46 PEBITIN B Conv Peak EXECUTION TODA, CP B DDA DP46 PEBITIN B Conv Peak EXECUTION Resonance PEBITING PEBITING PEBITING PEBITING PEBITING EXECUTION PEBITING EXECUTION Resonance PEBITING PEBITING PEBITING PEBITING PEBITING EXECUTION EXECUTION EXECUTION PEBITING PEBITING EXECUTION EXECUTION <t< td=""><td></td><td>705A.0V BI</td><td>80%A</td><td>DIP40</td><td>01002216</td><td></td><td>Food</td><td>Forlins G</td><td>enca</td><td></td><td></td></t<>		705A.0V BI	80%A	DIP40	01002216		Food	Forlins G	enca		
TODAL OF BOA DPGA DPGA <thdpga< th=""> DPGA DPGA DPGA</thdpga<>		TODA,OV B	ADA	01/160	01082216		Food	Postus, D	Parlance_Street		
Opp, rol IDSA 2014 Del SETS ID Order Production IDDOTS ID		0 1004,0V B	80M	549.60	01052216		Owne	Fishes, St	ing:		
Elevery 15-5/17 Hollins		1006,0Y B	RDA	89918	0100276		Care	Fedna, S	sticl_instell	-	
Biology 16 (27) Morea Paymention			V	-							
Massestime Marketing Production of Efficiency and commenced and any approximation back disclosed Marketing Production and commenced and any approximation back disclosed Marketing Production and commenced and any approximation back disclosed Marketing Production and commenced and commenced		droeing 1 to 5 of 10 writes	Systematione					PRIMA		4 Kert	
Manuality Instrume Performance Performance <t< td=""><td></td><td></td><td>Mutellan</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></t<>			Mutellan								
Description spectra del finanzia Description spectra del finanzia del finanzia Description spectra del finanzia del finanzia Description spectra del finanzia del finanzia del finanzia Description spectra del finanzia del finanzia del finanzia del finanzia Description spectra del finanzia del fi		Attractations	Indextra	And Designed	and the same is	and the state of t		-		•	
Jackson supervisionsy if with the intervision of the intervision o		D-Droad involute of N/T and I should	Exhang UPPs	Prin Gerung	/IDC0Comen	Water Diversity	Conserve a ware o	arra			
VECtowards where Counting International Count		and open interesting and the set of the set	(CNA scientifica		Phile chains have	Sand Radional Ta	-				
Collinations (double configuration Angean (2014)). Solid line in Section 2014 (double configuration and a section		LIEG Cristerily of Nov@ Carolina	W-REDWON University		41,04 Hadress	University Sales	ul (Effedure				
Hitse Wat, Sarran Killer (20) MAX INF And Sarran Hitse 2011 HIMA Singurang I KAN KIMA SA Jadpana Kances Hasa Helen hing San Bearting Kan Alberta Manak Kance Kana I Kance Kila Kance Alberta Mana Mana Mana Mana Kila Alberta Mana		Co-Onnual cases	GA mit-Bunita General A Solutioning	Values of the	GABullyring GA	4					
No. 49 Second Michigan and Page Machine Second Max. Allocat Pages A Second Pages		NUTLING HERAT COLUMNIAL AND AND ADDRESS AND AD	NS YOR I SHOW HERE 200	I with A Simplement	NEA WHAT ALL A	rdinses Kourts	Phase Preton Artig	Cites .			
Kang yan J		Ne of Ferning Erfring to ray ONA	Machille Lawrence Million P	scrar DSA	00xi/1tem 0	Dealton LOARS!	NS (7549				
BAY 51, Myneter Senser Mars Genere Mars Genere (Str. 1994) TMG TMG TMG LCC Ontonio Mars LC Ontonio UCC Ontonio Mars LC Ontonio Mars LCC Ontonio Mars A Thomas Official And Andre Official		INPEC Mynotocianisme Neter Amerika Shift	Wes Perre LTG Orthoad	West 12 MILITIE	Union & Olivert The LOADING INVAD	orma L/G Catalita	o izran & therew Q	Examine Pylord	Curringers Dife		

And then you can identify this TCGA underscore OV that is the cohort we want to look at basically it is a TCGA ovarian cohort and then you click on that and then it will give you the next step select search data set. So, as I mentioned the go of the linked omics is allow you to go from any attribute. For example, you can go from the mRNA or microRNA or mutation or proteomics to any other attribute, but here we have a fixed question we are interesting survival right.

Survival is a type of clinical data. So, in within the step 2 we want to define the search data set and the search attribute.

(Refer Slide Time: 02:28)

a sustained	for the description of the line of the				4006	0 11 14	
C A not becare !	interconcercity operations			ATTAN & Amaran i	E NYT E salari		ber Backmake
TO American B PubMed TM	C I) CPORC-OCC - SIL-IN-B	mus de modes anne C		1810,880.1		_	
Anarysee Amulta	Brain Lower Grode Glorus (.00			TCGA LGG		_	
Aller 7 Parage Stri 21g Year Paralet	Browny 1 to 3 of 40 within			Pentos 1	2.3.4.5.0	9 Not	
						Internal	
	Bitte 8 artistes					an a	
	And Arrist C make			C trattata C Analysis level	O Papalina	0	
	TOTA OV ALL	Clevel Civited	01092116	B Cinius	Frahma		
	5 100 01 PV		010027-0	B Gara	Entron WHILING		
	THE OF B	ATLA 19740	01002716	D Pool	Fishous, SHITCE		
		1004 INPA	0100274	D Post	Feature, GISTICE, Scienting		
	100,07 8			B Ore	Februa Scottica		
	100,01 8						
						4 100	
	Break Pass desa						
	SUPAR PERSON PERSON AND A PERSO	ola					
	attrictly happendataces (% 11)						
	elteraty notheparticulation (N 338)						
	gendur famora (ht 521)	Colored Winness					
	A National States of States	renana je bert					
	m managori, managori no (n. sour)						
	A side 10,339						
	L'HELENARD AT ANY CANTON	W. instant investi	PL Distantine to the later	inante based of Balline			
	Elb-Circuit care	(på. edu. Roverna Garanna Anarysas mil Decorrory	NA GADUILATTY CA.D				
	In Day RNA-Surriva Hilley 2000 RMA. Nan-watering	HE-WE-INTERNATION AND AND AND AND AND AND AND AND AND AN	BARRING MEANTANES AND	ienau Revene Praia Provin Area	Gre		

So, here we want to start with clinical data let us say you for clinical data and there is only one type. Let us say you select the clinical data and now you have step 2b as an option, but not say if you are only interested in for example, a certain stage of ovarian cancer or maybe you are only interested in a certain subtype of the cancer and then you can do the analysis for the subset of tumors.

But, today let us just say we do the analysis for the whole cohort and then we select the search within the clinical data right the multiple types of clinical information as the platinum resistance data or overall survival. So, today we are going to select the overall survival.

(Refer Slide Time: 03:26)



Now, basically you define the what to use as a query attribute basically it is within the TCGA ovarian cohort I am interested in the overall survival and then I want to ask and which proteins are associated with overall survival right. So, then on the step 4 the data type let's select the proteome. So, for this cohort both John Hopkins university and the PNNL they both did proteomics data generation for kind of overlapping subsets of the cohort.

But, today let us just say we choose a PNNL data. You click on the Select next to the PNNL data, but here you after we finish this proteomics analysis you can come back and then you can select for example, RNA seq and then you can correlate the overall survival to RNA seq or you can select a copy number and then you can correlate survival to copy number. But, now we are going to do the proteomics only.

(Refer Slide Time: 04:42)

TO Americania Sg. Addited TMC	U OFFAC-DOC - DU-HUB ()	tenia gle Bengin Premier () 100 ()	103 수 NS 및 C 2014 용 Anazon 온 NYT 는 cator · 는 Only	e faanerarier
	THE OF TH			
	Due 4 Coretes		tary?	
	Balton - Bangte O Hathan	O Desirger O Platform	O Sets O Institute O Analysis lived O Pagetine O	
	100A DV	Pessere Vela	91002119 Jack Geve CDIP ITMG Universitie/Name	
	0 1054.0V MVL	Pueure West Coal	andwatesh Mich, Gera COAP (TAQ Universitie)/Paris 4	
	Accounting 1 to 2 or 2 account (Shareed Yours 15 to	la francaria 🔛 ostarritaci	B B B B Freiss T Stat	
	PRIMA BULACI E VEREICAL MUNCO/	1115		
	First contract in and units			
	Conteposition		-	
	Alizarities		-+	
	Alterrities Béaultable /HT off read	BODD Canada Notes Sectores	- N	
	Althoughton (): B docal technol: Alth T and T served (): A server page as Converged Assembly of Name Device	BODDC Careadon Millione Short Derame Economic Careiro MCMM3 Andreame	y	
	Abbernations Bibliogenetics after Tard Taransed Jour Jahon Hegaras Conservery Johnson by H Hand Databas Under Conservery of New Yorkson	BODD Cywainiw Hutse Bren Gryne Bobron Crew Holdd J Arlene Wydynydd Brang	July Divers Terres January County of Not Dearins May, July Schwart, Standard County of Not Dearins May, Schwart, Space (Note)	
	Abbenations Bebaat tables of MT and Tarvard Jourges Streams Streams Sector Sector Of Colourary of Matt Sector Decilinations	BODDO Cyrusian Hildrer Bruin Gwrure Borenno Laman Boldwid Jammir Victorychell Stransji Garaffacting Gransk Angyan erfffal	July (C) Users Taxing Jawaig (Lossing of Ball Basins May, Analy Kamana Jawaig (Lossing of Ball Basins May, Analy Kamana, Sawai / Malans Sabaharang Jawai	
	Adventures Biological India APIT and Parad Any adam segues Councephones y at Net Dealer Conditional one Conditional one Conditional one	BODDO Cynaliae Nilliae Brinn Gwran Borenn Llwrae Brinn Brenn Rollwid Armani McBarlydd (1970ad) Garaffarfar (1970ad) Garaffarfar (1970ad) Hawli Igana (1970ad)	Aug. Schwart Franze Jonensky Sweeting of Hard Exercise Mall, Angle Kommunik Martin Link McCollinatory Cole in Speed of Hard Inte McCollinatory Cole in Speed of Hard Inte	
	Alexandrew Biological features of the out of sound Days alone request Sound publishing of the Continue of the Original or the Alexandrew States (2015) MAN Sound or the Alexandrew States (2015) MAN Sound Sound States (2015) MAN Sound States	BOOLOGINA MANA BALLANDA AND AND AND AND AND AND AND AND AND	And Construction and Annual Construction of the Construction of th	
	Alternatives Biological training del Tarved Day cates reports Charange Alternatives and the Cateline USEA - Income States 2016 INA- Income Cateling and Print, Income Table Print Annues States 2016 INA- Income Cateling and Print, Income Table Print Annues States 2016 INA- Income Cateling and Print, Income Table Print Annues States 2016 INA- Income Cateling and Print, Income Table Print Annues States 2016 INA-	BOOL dynamic National Series Dans Denamic Records (James National Common Automatic Common Automatic Common Automatic Common Records (Common Series Automatic Common Records (Common Records (Common Re	Activity of the second	

And, the another step 5 let us select Cox regression. So, this is for survival analysis and then Submit Query.

(Refer Slide Time: 04:51)

Separat Garoor Trole Senat Garoor Trole Senat Sourch Internet, for ex, Prucesion datasese Senat Prucesion tacount	Association Result	to making!										
a Select Search attribute, for its, 1953	Show 5 Contrine	~							Noard	ĸ	C take C Belline V	
 Setters Tenger deleases, for ex, mRVA Expression (RURates), dutatore Setect Exerciser Techol 	the group g man	il I	10-17 Security Institu	. Data	Fallers	Data	training Arr	-		Barryin wine		
a Clock SLADOV horton		Search a	TESA, BY HE	Ofess	Qreat	2-024074	8 GF	nut fiernas		Autoleurent sonte Parante till	Coruma 1018.524	
annes Press per sin repo	-	depet etmant	ICON OF MAR	Parters	Win &	2.08879	res, ter	N CONF_INAND_UN	danatopisticui Desta d'for	Automotivity Protestable Televity (Ad) (20 altern 7 State		
- TANKA		Desirios/	Rollad	Oningra		8						
	Cheroper											
	Showing 1 to 1 of 1 cattle	1							Perris	ns 1 No	at .	

Any query that has been queried by other people because some other people already done this analysis and the result is saved. So, that is why it is so quick and now you get this and then you click on the view.

(Refer Slide Time: 05:05)

0 8 9 pl material CAI	x 🛊 Lisedonia x 🛊 Lisedonia x 🛊 Lisedonia Par x 🛊 Lisedonia Par x	x +
- C A Not Secure Int	atomites orginatinin pris	Faces O E
TO American St Added TMC	O CPTAD-DOD → Sinitus () San's & Sample Frances () in () WG & NG & GIRAR # Annual € NYT () spine *	Coler Backmarks
-	Genan	
	Showing 1 to 1 of 1 carries Previous 1 Next	
	NAMES AND A ADDRESS ADDRESS ADDRESS	
	Den 1 Den 1913	
	Sense C Conversition C Lagrander C An A 104 C Event, SO C Event, TO C event, TO C event, SU C Event, SO C Event, TO C	
	Yes (1958/2) 1.100-00 2201-0 230-00 230-01 230-01	
	Mar 20081 10040 2006 2000 1340 1340	
	Ver THAN AND AND AND AND AND AND	
	the 252 14840 472 676 6340 1541 5 1 -	
	Ver 1950 420ed: 50% 50% 51e0	
	the man assure the two and that I	
	Russey 1 0 6 art and (Arter Lating article) Portia 1 907	
	ten term ter	
	Edit Von & Apartie can directive adapt adapt of analysis neg art trail Edit (see Section 2 and a set ages adapt adapt of analysis neg art trail	
	Product - August memory have been realized - August - Aug	
	6 Sont all Avenue er associative A super annue advant where for sure annue Coam (2) - Summer er annues on a loger timeset advant webust for a per alives. Log(Razard ratio) (CORPH text)	
	Paulterip Constant Spontant Series (<u>Dec 1947</u>) Nagatively Constant Spontant Series (<u>Dec 1947</u>)	
	AND DESCRIPTION OF THE PARTY NAMES	
	State of the second sec	1.00
*	STATISTICS STATES STATES	
	WHAT WAT AND A PARK CONTRACTOR	
Trans. on The	estcansalet Plana *	Ben H X

And, you can also search the gene you are interested in. For example, KRAS or some other genes let us say ovarian cancer, what gene could be interesting anyone has any suggestion on which gene you want to search for? Maybe let us look at the KRAS.

(Refer Slide Time: 05:31)



It looks like it is not significant. The P-value is 0.6.

Student: TP53.

TP53 it is not significantly either, but is apparently there is little bit of trend like higher expression is associated with a poor survival.



(Refer Slide Time: 06:15)

And you can also go to the bottom of this page and this will show you the top 50 genes.

I do not know 50 or 25 genes that are most positive in your correlated with survival or most correlated with survival. So, we found this to be useful for in.

Yes, get a quick a view of the top genes and the detailed data. It is binary data it is. Subtype of one versus other subtype or and then you just have black next this two parts right. But, for survival data it is only you have test event and the still alive or dead. Right, but then you also have the time of survival or. So, let us me the first this part at the bottom shows you these are probably that. People who have died and these are still alive and then within each group and then this is a survival time. So, some most of these are censored data right most of them have not died. So, it is censored, but these are the ones which actually test events. So, basically two group of people and then within each group you have mean survival time from 0 to 9.

(Refer Slide Time: 07:57)

Showing 1 to 1 of 1 estries			Prevation 1
Unifieder Uniferenter			
	- Inchese	(Analysis	
	Select Tool Gene Bet Event	mert Andysin (2010A)	
	084 - Iber	and And in	
		Al Louis Report Amount	
	Rask Oriteria (from LinkFinder Result)	P-type Data Signed on	-
	Minimum Number of Genes (Size)	,	
	Simulations	806	1
		and in case of the local division of the loc	
		Select Tool Sens to ever Select Tool Sens to ever COLA - Loose Enrichment Analysis Rank Criteris (from LinkPinder Result) Minimum Nember of Genes (Size) Simulations	Select Tool Grow End (version of Andrew 1990)

So, I think and then if you want to do some pathway analysis for this you can click on the Link Interpreter and then you can do ORA or GSEA similar to what we have done before, but with if you submit the query and basically it will give you the WebGestalt report and we have done that before. So, let us say we do not need to do that now rather I think we can go back to perform a new analysis.

(Refer Slide Time: 08:22)

The Radion	Carrier Made	ante tarta	Cup-4 Brand Dirt & Berness	
· Brill Marth Wilson, Sr	FOR BATTONET CONT			
· Breat Specie Paparities	Das 6 Horns		hert	
a Garner Statuth arritories, Tor		Course Tays	n Q Bengin valuet Q	
Visit 70 per ansar, hr	And Stated in case in (AMA)		TODA LAIK	
An Antipa Cognition	Advancement supports MED		100A.ADD	
And these times	Barrier or real or reason (FLCA)		TORADA	
Chill BLANDY LUGH	Name and Party Office of State		104,102	
Antyras feechs	Annual Property and Property of Street, or		COLUMNA .	
and Henry par bit was	Disarry 1 to 2 of Kirrson		Ferent 1/2 2 4 5 4 MM	
	-	-		
	States of the life of the local descent	RECEIPTING STATE	India Telliment Instantion and A gain (party	
	Porgens report constalitions of	NDA.W3 Anderson	RDC-Party Retreact Association	
	INCOMPANY OF MERICANING	Alarmer provid	which its approximation of the form	
	College and	lat.mit.lations discord for spins within Engineering	lation of the late	
	Citing Operatory, then \$107,784.	18 officers (the 201-the Income	MDI-MTRAM 2 Automotification Reserves Present Factors Aving Class	
	Restaur Increase advances on other (MAR Restaurceses (P)	Manager & Lowest Printers in Party State	Based Preme 2 South 6 12 Michael Will	
		The Party of Street Street Street, Str	new & Oliver, Name of Street, Name & Toronto, & Darrison, & Darris	

And, the in this time you still do the same analysis, but you ask a slightly different question you ask which copy number change is associated with survival in this ovarian cohort and the after that you do another query which RNA in the RNAseq data which mRNA abundance is correlated with survival. And, after you get results from all those three platform and I can show you how to do show you how to do the link compare.

So, you can go back to new analysis and the repeat what we have done, but just change the target set target data set.

(Refer Slide Time: 09:11)



The copy number is SCNA and another for if you filter the data type for SCNA and then you get multiple four rows right you because the results can be we have results at the both focal level and the gene level and then at the gene level number the data can be stretched coded or without stretch code. In this case we pick the one without stretch code pick the third one.

Yeah, We pick the SCNA gene level Firehose GISTIC2 that is the algorithm used to do the analysis.

Ohh, Sorry in the at the very top you select clinical, I am sorry.

We should select clinical and it is in the target data set that is so many copy number.

(Refer Slide Time: 10:17)

	Interdomics, org 10 perior print			4 0 0 0 0 0	Contrast
The Assessment St. Autors The		Barr de dampefrance (* 10 (*	1 10 \$ 10 8 B LERE & Avenue	E MAL D COMME >	Colum Res
					No.
	anne that a new				
	NAME OF TAXABLE PARTY AND DESCRIPTION.	Olana			
		0.97.0			
	anic series				
	UNDA BELET WHETDHARTU				
	Dur 4 Harm				Servit.
		O Beatres O Parters	O Barrs O breaklans O Analysis break	O Paster	0
	men - and	and meres	11-12-17-16 H FILM	Fratras, DRITES	
				Forma Settica President	
	and the m	Ena Deris		Anton Service	
	0 ania e	NAA DIPIN		Andrew Solitify reserved	
	COLOY B	EDA INTE	FARTS & STA	The second second second	-
	and the second second	1 10m		-	
	Sector Se	IN PROC.			
	REPAY NAMES & POST OFFICE ADDRESS OF				
	- Exception on our pro-				
	And the second s			-	
	E fosilullo JP.Todivisi	BOSC Carson Notae Serie Denne Notaes (sere	All a features a consequences of the la	Cantra	
	All steel ingene conceptions of a	MULTINA IN MILITARY	PAR, Aprile Institution Reservations		
	170 Street of Sale Desice	V. Market County	Red Million Concession and Provide Street of Party of Street of St		
	Chillenge of the local division of the local	GALERARIA DE LES ACESES ANDRE	Delectures in is		

It is step 4 step 4 Select copy number.

And, then you can also view the results for copy number the same way ok.

(Refer Slide Time: 10:43)

T C A Not Secure	Enterdomics.org	(alter pro										-			Online Designments
ti Annairean Sg Addres 1 1917 Frank III. Station 1919	ve ili omie i	xx + 1		ICELOV	URG .	(* un metares Consepte		de sus de sus	9 0294	B Antorio Fisher, R	E NIT BAX_NE (WARD OF DOWN				
			a section of	-											
			-	-	-	-	Pattern D	a 1		· Partes	- 24	-			
			-			-	-	-	e Crise	Farms		and and	-	-	205-04
	0	0	dateset Target			-		-		Papers,		LAC 147%			er .
			-			-				0	eries of Larg	m+3724			
			Parallel Te	-		Derepe									
			Datest Sal	10 1454	-	Patter	- 241	Autors.	and Part			Bangin Ka			and the second second
			Aunt 10	4,5Y HE	-	Ores	12023		CITER FRET			Perist		Depth	20040
			Paper and	-	-	CENT		-	tare coar	(ma),man	angrow p	Actual a	-		
											Dector of Re-	-141	12-11-12-0		
			Reason and	N.C	Lan repo	and they									
	-	Corues		-	-	-	-	-		-	-		7 0		
	Showing I to	h of h ones	8.								mos	- W	11.57		
	Bar 10	(and a state	-									-			
	and a	Juny' 1	Non-	A44 7	ä		Patter	14	No.	N.M.	1 1	-	here.F	Borg 708	
			Matt	Alliett 7	11 11 11 11 11 11 11 11 11 11 11 11 11	2200-01	1 PERCENT	1 1000	110+6	\$38will		1.Gertl	120411	10110	2. 1. 1.
			11/101	train ?	11	1.161-10	1004	100	-	-	-	COM-ST.	123101	120.0	
			LINK	Allowed A	-	(TWell	Lawy	1.000	112418	9.00m11	No.	ALCOHOL:	Arrest	Allant	
			AFRICAN .	amout P	1	1.168.14	1.Mint	6100	-	-	80	Sure.	-	Amer	Vol 1

Let us say if you were able to get two or three results from this for example, if you get all the three RNA seq, copy number and the proteome in this OV data set associated this survival. Now, you can do the link compare and what you do is you select on this the third column this select column you select the let us say all the three of them and then you click on compare. Now, you have this table with all the results.

(Refer Slide Time: 11:32)



So, basically in the first few columns you have the three results from the three platforms lets say this is ID 24 correspond to this ID 24. So, this is the results from RNA seq and the then you have ID 56 this correspond to the results from copy number and ID 57. This is a result from proteomics.

So, basically each of this you have the signal strengths meaning the signal ratio and then you have the P-value FDR right. And, then here we use the sumZ which is a meta-analysis method try to summarize a P-value from this three analysis and then you get to sumZ statistic and then sumZ P-value and sum FDR. So, let us sort the results based on the P-value.

(Refer Slide Time: 12:24)



And, as you can see now before for individual platform we get very few significant genes right. In proteomics we did not get any gene that passed the FDR cutoff but now we can see after you integrate straight platforms the a lot more significant genes than individual platform. This indicate although each platform do not give you the signal strong enough to pass a cut off, but maybe all of them point to the same direction that gave you more confidence. So, when you put the result together and you get the enrichment.

(Refer Slide Time: 13:07)



And then you can look at the top genes to see the minus log and meta FDR meaning this is integrated FDR and these are the top genes positive genes and the negative genes. And, then you can also have the result in the heat map.

(Refer Slide Time: 13:38)

A NOT A NOT THE OWNER.	Inustonics orgination	* 0080	
Ci thansan St Autors I		tertha @ Serie & SeegerTranse () we () and () has () All 2 All 2 Ananon () Art () terter	· · ·
		Same distance - 272.53.500 rd.51	
		Destinat Destant Dis segureren arabere	
		Generer senior battan ben Partan ber bernar ben Bartier Bartin Bartier	
		Parent Mide Or ME Once Drove D'ORN'S D Coluir Danies Restanting	Drum 15.00
	0 0	Annual TEDA ON PARE Promote Control Co	
		Beautowithstad Competence man	
	Carto		
	Managers 1 to 3 of Lon	Perims 1 Not	
	and the second		
	Inclusion	Independent	
	1110	Providence the associations	
	LinkCompare :	Comparing the associations	
		Top process and the second sec	
	10170		
		für he	-
		Andread Annual Annua	
		Sandara Marine Anderson and B	
		Main Star Units Realizing Lond 17	
		Man fine Cheris Radicions Level (29	

You can also redo the filtering and for example, you run the filter rather than setting the FDR as 0.05 let us say if you make this to 0.1 submit will give you more genes to look at. So, basically you can adjust this to get the genes that pass your cutoff. For example, now if we relax the FDR and then we get more genes to look at.

(Refer Slide Time: 14:03)



And, the all these results also can be downloaded and saved by click on the download figures.

(Refer Slide Time: 14:24)

The summer of Palaret Th	e su orac ace a hi-	a () teres de Benge France (* 100 (* 100 (*	p to g calls a most	servers and an	
	• •	Bankarderberer Bankard Bandy Bankard Bandy Bankard Bandy Bankard Kall AV HAL Bankard Kal	nation <mark>Analysis Populars</mark> 1 Consul Remons 1955, Gave COLO 2004 (Judy	Sampiralan Alifotowani asmit Material Material Material Dang of Deriver - Of (12) 80773 (A	gan com Garyes gan Gar
	Chelles .				
	Statering 1 to 3 of 3 watche Gentlements to	angelin		Prevana <u>I</u> Next	
	Stateling 1 to 3 of 2 votile	Labitopyrite : A Labitopyrite : A Lybitopyrite Lybitopyrite : Gen to Dedre	afterang Enrichtemet C Actors ant Anarycis (1985A)	Person 1 Not	
	Statesing 1 to 3 of 3 estible	control California (A Control Control Canada (See & Joon Canada (See & Joon	nathanang Zarishamant (Arazana mana Angurin (1956A)	Person 1 Nat	
	Sherving 1 to 3 of 3 vertice Lentences U	Labitationers - A Control Con	Antoney Enrichment (Antonia ana Antonesis (1986)	Person L Net	
	Starring i to 3 of 3 estile Contension to	A revenue ()))))))))))))))))))	utherer Enrichment L dents men Angel in Z Angel b havea Boherer grannl mf Enrits i	Person L Net	

So, again and the you can go to the Link Interpreter and for example, select GSEA and then now you can submit this sumZ statistic as rank metric to rank the genes and then the analysis will be perform the enrichment analysis will be performed against this summarize the statistic rather than individual ones. Again, the result will be the same I mean very similar to what we saw before for the GSEA in webGestalt

(Refer Slide Time: 14:59)

Points to Ponder

- CPTAC and TCGA datasets contain information from complimentary approaches such as Proteomics and Genomics, respectively that can be used to gain insights into various clinical aspects related to the disease.
- Linked Omics can be used for multiple cross comparisons from existing data on TCGA and CPTAC without the need of any programming skillset.



In today's session you got a demonstration about how the three modules of linked omics work using ovarian cancer TCGA dataset as an example. It also shown how the tool can be used to generate survival information from a target gene. Using data from RNA seq, copy number variation and proteome level information on the effect of target gene on survival was demonstrated.

Finally, using the third module of the tool it was demonstrated how the results appear and it can be interpreted. In conclusion, we hope that now you have a fair bit of idea how you can use these available tools and use them for your own research.

Thank you.