Micro Array Confidence Intervals; a more detailed look

Protein Interaction Map Integration & Visualization

Dr. Werner Van Belle werner.van.belle@gmail.com werner@onlinux.be University Hospital Northern Norway

Poker Table Distributions

Take 85 poker tables; what is their height distribution ?

#	Height
10	69
20	70
25	71
20	72
10	73



Poker Table Distributions





Furniture Distributions

50 pieces of the same furniture

#	Height
5	65
7	66
15	67
13	68
10	69



Furniture Distributions







- If we take 4250 random combination of such table stacks:
 - How many times will we find a specifically sized table
 2 ?



- For each group of a specific table 2 sizes:
 - how many times will the 1st table occur with given size ?

	65	66	67	68	69
	425	595	1275	1105	850
69	50	70	150	130	100
70	100	140	300	260	200
71	125	175	375	325	250
72	100	140	300	260	200

Of the 4250 stacked tables ? How many will be 136 cm in total ?

- A: 415

65

66

67

68

69



Distribution of tablestack height





Confidence Intervals

If we pick a random combination:

4199

• What is the probability that its size >=135?

134	50
135	170
136	415
137	705
138	924
139	895
140	660
141	330
142	100

0.99

Confidence intervals

Probability that its total size is <=141 ?

134	50
135	170
136	415
137	705
138	924
139	895
140	660
141	330
142	100

0.98

4149

Confidence intervals

• Probability that its size is >=135 & <=141?

134	50
135	170
136	415
137	705
138	924
139	895
140	660
141	330
142	100

0.96

4099

Confidence intervals

• The range [135:141] is a 96% confidence interval.

0.96

4099

134	50
135	170
136	415
137	705
138	924
139	895
140	660
141	330
142	100

- -

© Werner Van Belle, April 2007

Measurement Errors

- Every measurement has a measurement error
- If the true (but unknown) value is 'vr' then one will measure 'vr + something'

$$v_m = v_r + \epsilon \qquad |\epsilon| \le \beta$$

If we can model the error distribution, then we can say something of the real value

Probability Distributions

Based on the PD of the error, a confidence interval that will cover at least 95% of the real values can be created

If we measure r, then r will be within [m-0.4,m+0.4] in 95% of the cases.

0.9



Measuring The Error Distribution

- Micro-arrays are subject to measurement errors as well.
 - Intensity dependent
 - Often badly understood
 - 'Measurement' includes biological variability
- How can we measure the measurement error ?
 - measure the same experiment multiple times
 - provide the same sample for Cy5 and Cy3

Red should= Green

Green



Too close to error



Intensity Dependent Error

Difference between channels (red-green)





Importance of Modeling (I)





Importance of Modeling (II)





Importance of Modeling (III)





Reported

 The variability of the experiment, the variability of the cell system, the variability of the measurement technique and the number of measurements allows us to state that:

mCG16947.2 will be down-regulated at least 196.32 times in 97.5% of the same experiments.

	Expression	Probes Regul	ation Ra	tio		
ID	Scrambled	SiRna1	Least	Measured	Most	
mCG16947.2	116.3	6 0.47 down	196.32	245.83	inf	LEUCINE RICH REPEAT PROTEIN
mCG1034323	0.3	5 61.82 up	140.91	176.62	inf	urocortin 3
mCG121683.1	30.7	7 0.28 down	87.23	110.95	inf	RIKEN cDNA C030018G13 gene

Automatic Filtering (I)

Automatic Filtering (II)

Influenced by/Influences

- MK5 -> Multiple changes in gene expression
- 27000 gene expressions measured
- Those that change will very likely influence other proteins

Which proteins are likely influenced by our measured up/down regulations ?

The 'Involved' Game

 Protein change will influence nearby proteins, which in turn ...



The 'Involved' Game

Multiple proteins changes will all influence their neighbors as well.



The 'Involved' Game

- This network is iterated a number of times to expand the sphere of influence of all the altered gene expressions.
 - affected proteins will have higher numbers
- Protein Interaction key mechanism for signal transduction
 - Protein Interaction Network as published by
 - Jean François Rual et al- Towards a Proteome Scale Map of the Human
 - Protein Protein Interaction Network Nature 2005 vol 437, p. 1173-1178

Involved Proteins by Rank

PROTEIN CGI-126 (PROTEIN HSPC155)

RAD50-INTERACTING PROTEIN 1. [Source:RefSeq;Acc:NM_021930]

RHO-RELATED BTB DOMAIN-CONTAINING PROTEIN 2 (DELETED IN BREAST CANCER 2 GENE PROTEIN) (P83).

NADH-UBIQUINONE OXIDOREDUCTASE 18 KDA SUBUNIT, MITOCHONDRIAL PRECURSOR (EC 1.6.5.3) (EC 1.6.99.3) (COMPLEX I-18 KDA) (CI-18 KDA) (COMPLEX I- AQDQ) (CI-AQDQ).

CHROMATIN ACCESSIBILITY COMPLEX PROTEIN 1 (CHRAC-1) (CHRAC-15) (HUCHRAC15) (DNA POLYMERASE EPSILON SUBUNIT P15).

ADIPONECTIN RECEPTOR 2. [Source:RefSeq;Acc:NM_024551]

ODD-SKIPPED RELATED 1; ODZ (ODD OZ/TEN-M) RELATED 1.

DNA POLYMERASE EPSILON P12 SUBUNIT (DNA POLYMERASE EPSILON SUBUNIT 4)

PROTEIN X 0004. [Source:RefSeq;Acc:NM_016301]

XPA BINDING PROTEIN 1; MBD2 INTERACTOR PROTEIN; PUTATIVE ATP(GTP)-BINDING PROTEIN

HBS1-LIKE. [Source:RefSeq;Acc:NM_006620]

HOMEOBOX PROTEIN HLX1 (HOMEOBOX PROTEIN HB24).

Involved Proteins by Rank

- NUCLEAR TRANSCRIPTION FACTOR Y SUBUNIT BETA (NF-Y PROTEIN CHAIN B) (NF-YB) (CCAAT-BINDING TRANSCRIPTION FACTOR SUBUNIT A) (CBF-A) (CAAT- BOX DNA BINDING PROTEIN SUBUNIT B).
- GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2 (GRB2 ADAPTER PROTEIN) (SH2/SH3 ADAPTER GRB2) (ASH PROTEIN).
- SERINE/THREONINE-PROTEIN KINASE NEK2 (EC 2.7.1.37) (NIMA-RELATED PROTEIN KINASE 2) (NIMA-LIKE PROTEIN KINASE 1) (HSPK 21).
- E2A-PBX1-ASSOCIATED PROTEIN; PUTATIVE 47 KDA PROTEIN.
- NEURON NAVIGATOR 1; NEURON NAVIGATOR-1; PORE MEMBRANE AND/OR FILAMENT INTERACTING LIKE PROTEIN 3.
- NEURON NAVIGATOR 3; PORE MEMBRANE AND/OR FILAMENT INTERACTING LIKE PROTEIN 1; STEERIN 3.
- NEURON NAVIGATOR 2 ISOFORM L; RETINOIC ACID INDUCIBLE IN NEUROBLASTOMA; PORE MEMBRANE AND/OR FILAMENT INTERACTING LIKE PROTEIN 2; HELICASE HELAD1.
- GRB2-RELATED ADAPTOR PROTEIN.
- DNA-REPAIR PROTEIN COMPLEMENTING XP-G CELLS (XERODERMA PIGMENTOSUM GROUP G COMPLEMENTING PROTEIN) (DNA EXCISION REPAIR PROTEIN ERCC-5).
- DOCKING PROTEIN 2 (P56(DOK-2)) (DOWNSTREAM OF TYROSINE KINASE 2).

Involved Proteins by Rank

- TYROSINE-PROTEIN KINASE FRK (EC 2.7.1.112) (NUCLEAR TYROSINE PROTEIN KINASE RAK). [Source:SWISSPROT;Acc:P42685]
- DOK-LIKE PROTEIN. [Source:RefSeq;Acc:NM_024872]
- DOCKING PROTEIN 1 (P62(DOK)) (DOWNSTREAM OF TYROSINE KINASE 1) (PP62).
- [Source:SWISSPROT;Acc:Q99704]
- TATA-BINDING PROTEIN-ASSOCIATED PHOSPHOPROTEIN (DOWN-REGULATOR OF TRANSCRIPTION 1) (DR1 PROTEIN)
- WISKOTT-ALDRICH SYNDROME PROTEIN (WASP).
- PROTO-ONCOGENE TYROSINE-PROTEIN KINASE SRC (EC 2.7.1.112) (P60-SRC) (C-SRC).
- TYROSINE-PROTEIN KINASE HCK (EC 2.7.1.112) (P59-HCK/P60-HCK) (HEMOPOIETIC CELL KINASE). [Source:SWISSPROT;Acc:P08631]
- PROTO-ONCOGENE C-CRK (P38) (ADAPTER MOLECULE CRK). [Source:SWISSPROT;Acc:P46108]
- CRK-LIKE PROTEIN. [Source:SWISSPROT;Acc:P46109]
- PROTO-ONCOGENE TYROSINE-PROTEIN KINASE ABL1 (EC 2.7.1.112) (P150) (C-ABL).
- PROTO-ONCOGENE TYROSINE-PROTEIN KINASE FGR (EC 2.7.1.112) (P55-FGR) (C-FGR).
- TYROSINE-PROTEIN KINASE ABL2 (EC 2.7.1.112) (TYROSINE KINASE ARG).

Involved Proteins Network



Involved Proteins Network

Red = Highest involvement; Blue = Lowest Involvement
Based on our lowest estimates for up/down regulation
Based on the high confidence set of protein interactions
Measured gene expressions are not listed

Involved Proteins Network

Involved Protein Network

11 precurs<mark>or</mark> (er-associated dna) P 3} (er)3) (er-associated hsp40 co-chaperone} (hd)9} (pwp1- interacting P 4}

101371 proto-oncogene tyrosine-P kinese src (ec 2.7.1.112) (p60-src) (c-src)

97007 proto-oncogene tyrosine-P binase abil (ec 2.7.1.112) (p150) (c-abi)

167193 proto-oncogene c-crit (p38) (adapter molecule crit)

101336 tyrosine-P kinese bck (ec 2.7.1.112) (p59-bck/p60-bck) (bemopoletic cell kinese)

143322 tyrosine-P kinase abl2 (ec 2.7.1.112) (tyrosine kinase arg)

Involved Protein Network

184550

173349

119655 epididymal secretory P el precursor (niemann-pick disease type c2 P) (epi-1) (

141367 clathrin heavy chain 1 (clh-17)

-javigator 2 isoform i; retinoic acid inducible in neuroblastoma; pore membrane and/or filament interacting like P 2; helicase heladi

134369 neuron navigator 18 neuron navigator-18 pore membrane and/or filament interacting like P 3

tor 1 (cellular transcription factor ()(f-1)

163545

© Werner Van Belle, April 2007

1038