

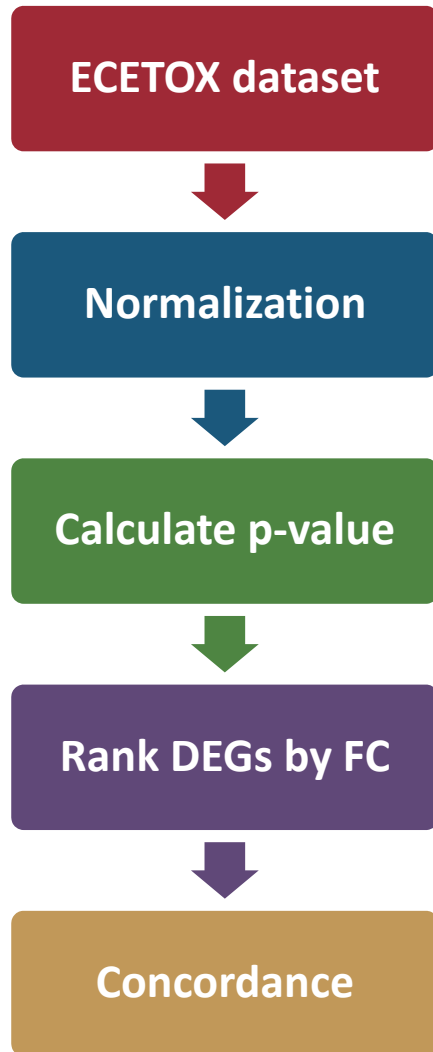
Towards Reproducible Toxicogenomics: T-test versus ANOVA

Binsheng Gong, Ph.D.

Division of Bioinformatics and Biostatistics
National Center for Toxicological Research
U.S. Food and Drug Administration

*MAQC Society Inaugural Meeting
SAS Institute, Cary, NC, USA; April 10-13, 2017*

Workflow for DEGs

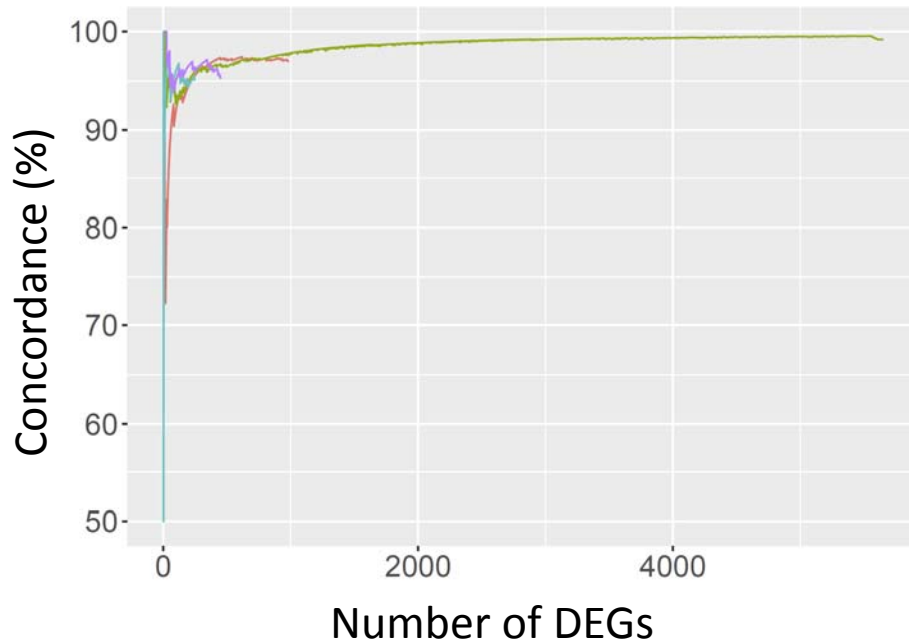


Group1	Group2	Group3	Group4
gProcessedSignal			gMedianSignal
Median-center to 1000 (raw data)			Mean-center to 9 which is derived from all the arrays with log scale
Welch	Student	ANOVA	Welch

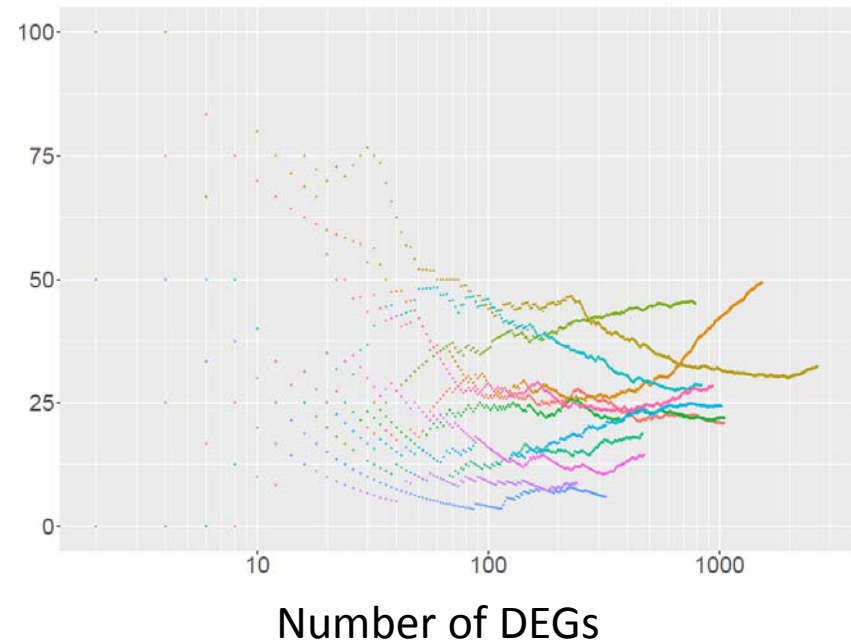


Much Lower Concordance between T-test and ANOVA than Expected

Welch's vs. Student's t-test



Welch's t-test vs. ANOVA



P-value < 0.05 Fold Change \geq 1.5

About Denominator

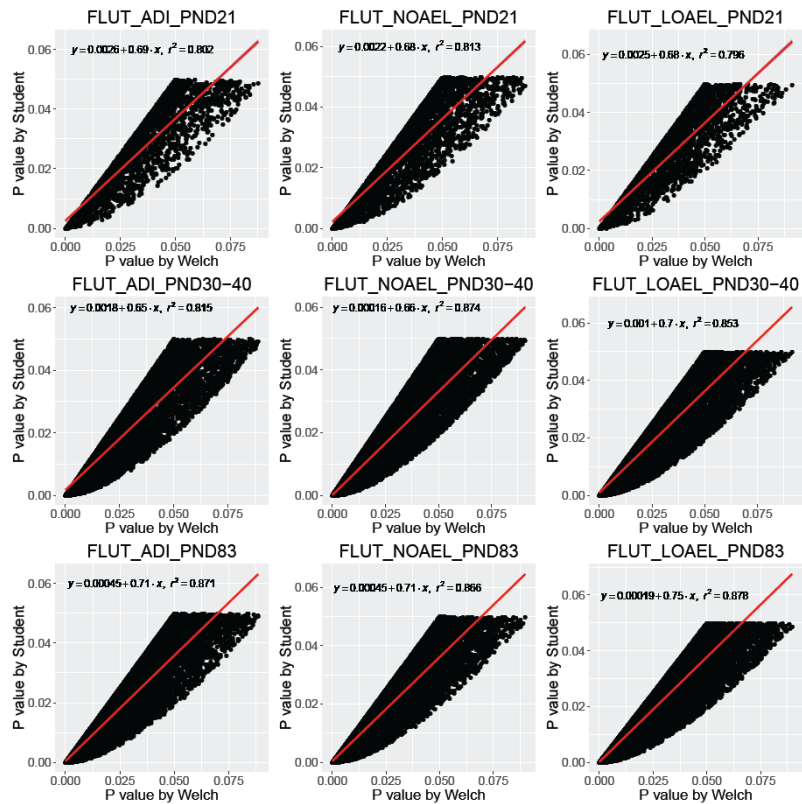
$$statistic \sim \frac{nominator}{denominator}$$

Statistical Test	Comparison	Sample variance	Variance estimation
Student's t-test	case vs. ctrl	equal variance	pooled, compared groups
Welch's t-test	case vs. ctrl	unequal variances	unpooled, compared groups
Pairwise t-test	group _i vs. group _j	equal or unequal	pooled or unpooled, all groups
Dunnett's test	group _i vs. ctrl	equal variance	pooled, all groups
Tukey's range test	group _i vs. group _j	equal variance	pooled, all groups

ANOVA + {

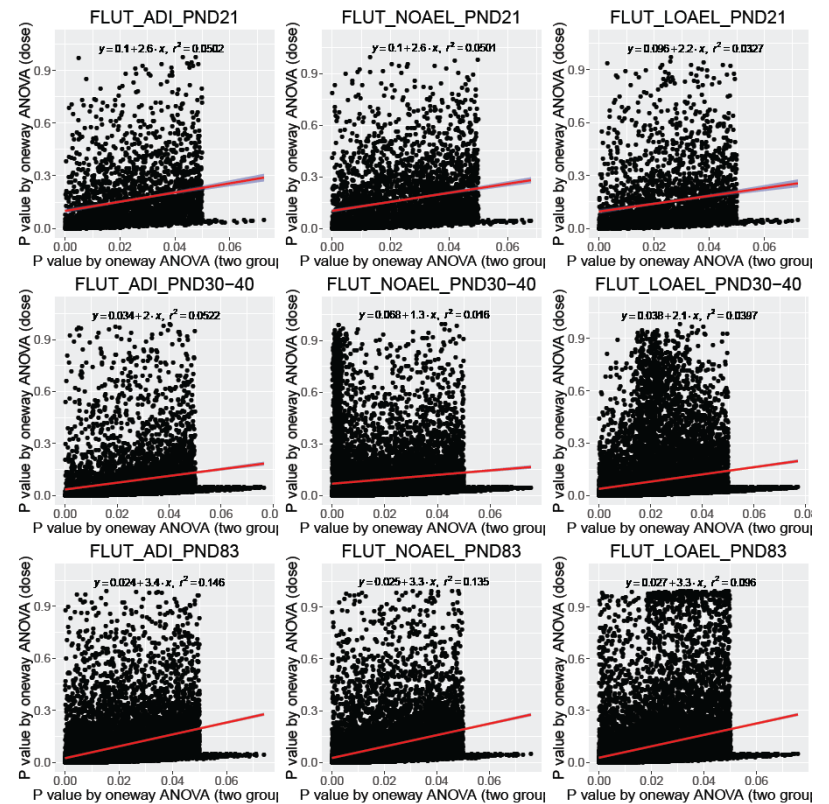
P-Value

Welch's vs. Student's t-test



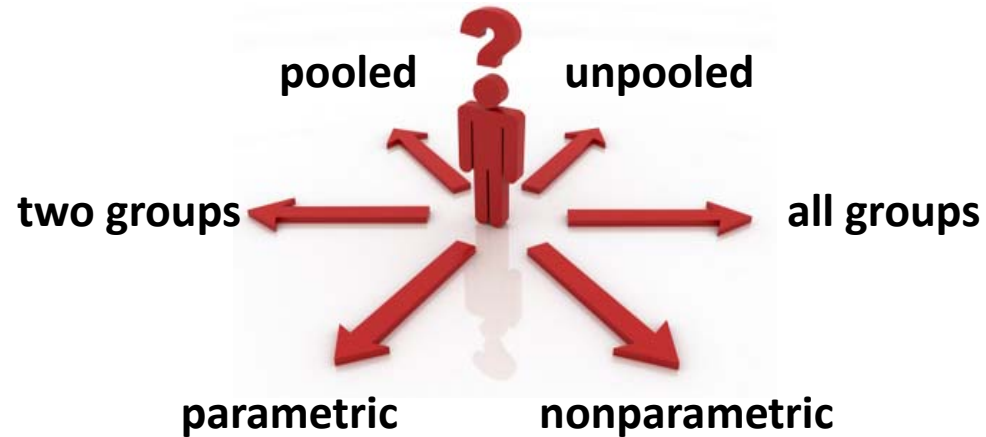
pooled vs. unpooled two groups

Welch's t-test vs. ANOVA



unpooled two groups vs. pooled all groups

Which is Correct?



Toxicogenomics

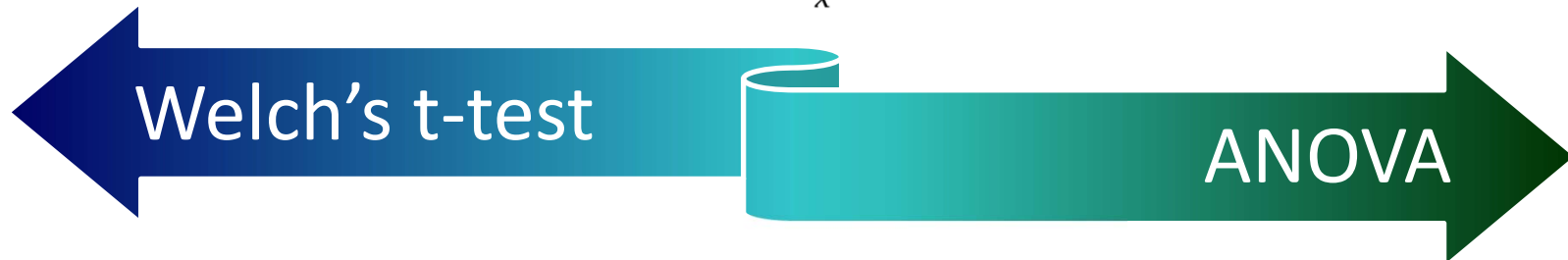
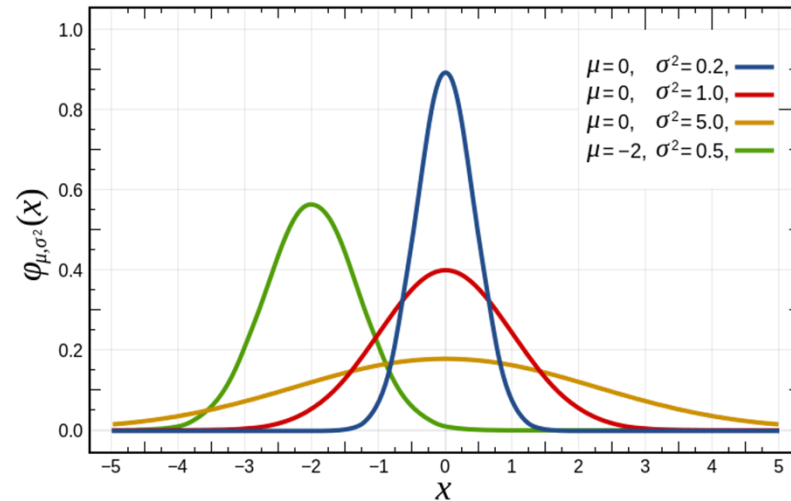
- Three replicates
- Thousands of genes
- Noisy data
- Study design

Clinical and epidemiological research

- Big sample size
- A few focused traits
- Data follows certain distributions
- Study design



Estimate the Parameters of a Distribution



- Larger sample size
- Treatment is strong
- Unequal variance

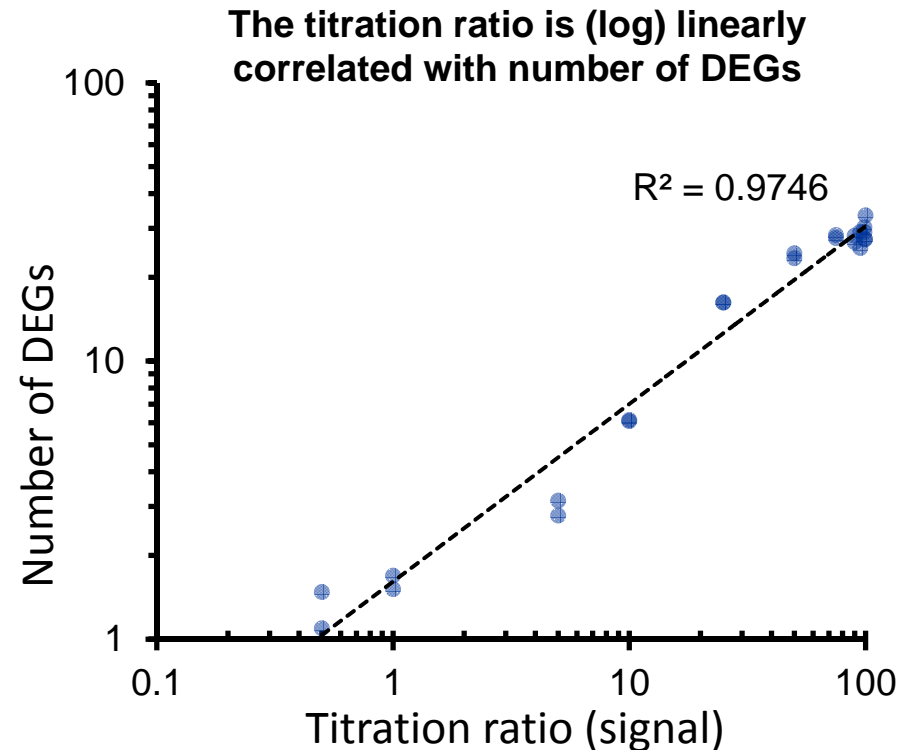


- Smaller sample size
- Treatment is weak
- Equal variance

Number of DEGs as a Measurement of Treatment Effect Size

MAQC Titration Dataset

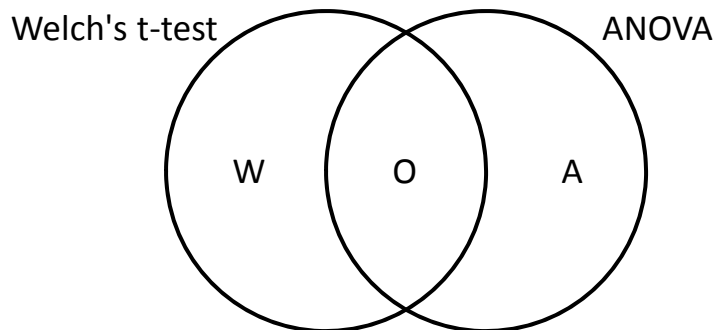
Sample Group ID	Sample Composition A% + B%	Replicates
1	Sample B	6
2	0.5% + 99.5%	3
3	1% + 99%	3
4	5% + 95%	3
5	10% + 90%	3
6	25% + 75%	3
7	50% + 50%	3
8	75% + 25%	3
9	90% + 10%	3
10	95% + 5%	3
11	99% + 1%	3
12	99.5% + 0.5%	3
13	Sample A	6



The number of DEGs can be used as a measurement of treatment effect size.

When Treatment is Strong and Noise is Weak

Compare	O	W	A	concordance	Compare	O	W	A	concordance
A_vs_B Lab1	21053	41	2	99.9	B_vs_C Lab1	18044	40	1	99.89
A_vs_B Lab2	19951	43	0	99.89	B_vs_C Lab2	16802	48	0	99.86
A_vs_B Lab3	20758	42	0	99.9	B_vs_C Lab3	17680	32	1	99.91
A_vs_B Lab4	21524	6	129	99.69	B_vs_C Lab4	18509	10	95	99.72
A_vs_B Lab5	20787	28	11	99.91	B_vs_C Lab5	17837	50	0	99.86
A_vs_B Lab6	19482	29	0	99.93	B_vs_C Lab6	16805	23	1	99.93
A_vs_C Lab1	3676	63	3	99.11	B_vs_D Lab1	5460	34	2	99.67
A_vs_C Lab2	2724	28	2	99.45	B_vs_D Lab2	4335	37	0	99.58
A_vs_C Lab3	2912	7	0	99.88	B_vs_D Lab3	5055	33	0	99.67
A_vs_C Lab4	3628	33	19	99.29	B_vs_D Lab4	5034	3	148	98.52
A_vs_C Lab5	2889	23	7	99.48	B_vs_D Lab5	4616	42	0	99.55
A_vs_C Lab6	2647	7	0	99.87	B_vs_D Lab6	4491	24	1	99.72
A_vs_D Lab1	16303	56	3	99.82	C_vs_D Lab1	10132	39	0	99.81
A_vs_D Lab2	15236	42	0	99.86	C_vs_D Lab2	9984	42	1	99.79
A_vs_D Lab3	16070	45	1	99.86	C_vs_D Lab3	10658	52	4	99.74
A_vs_D Lab4	16752	1	184	99.45	C_vs_D Lab4	10765	7	133	99.35
A_vs_D Lab5	15729	28	16	99.86	C_vs_D Lab5	10264	57	0	99.72
A_vs_D Lab6	14524	34	1	99.88	C_vs_D Lab6	9612	30	0	99.84



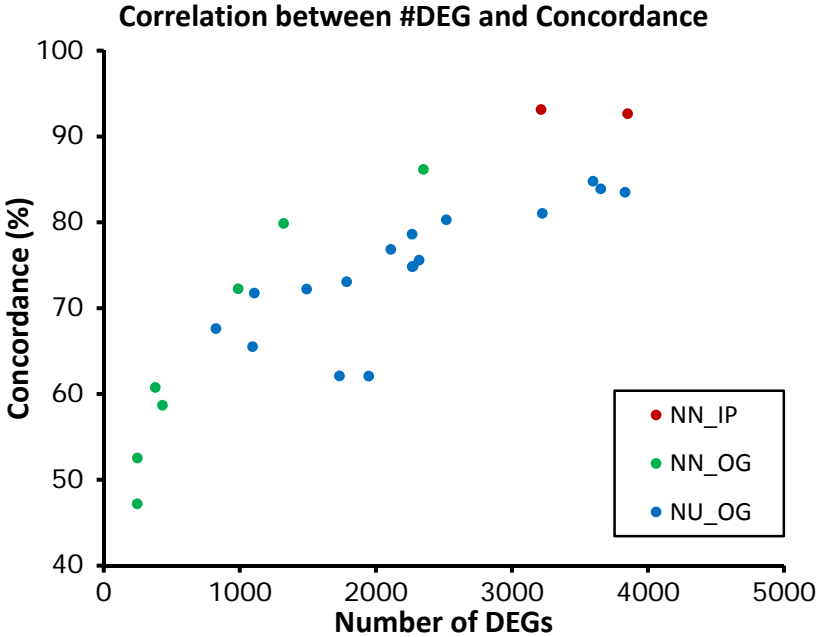
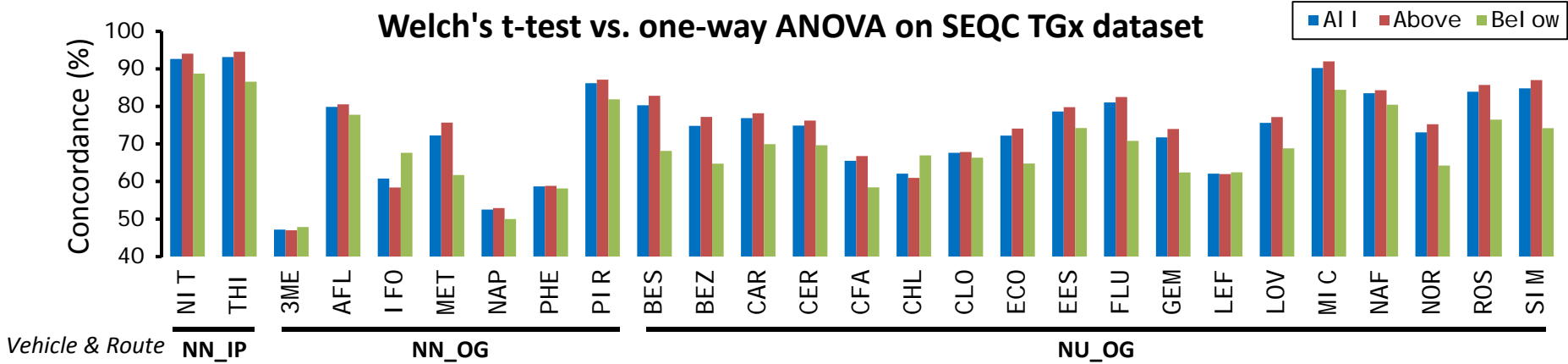
$$\text{concordance} = \frac{2 \times TT}{2 \times TT + TF + FT}$$

- MAQC Data Set – Main study (technical samples)
- The concordance between Welch's t-test and ANOVA is pretty high (>99%).

Fold Change cut-off > 1.5

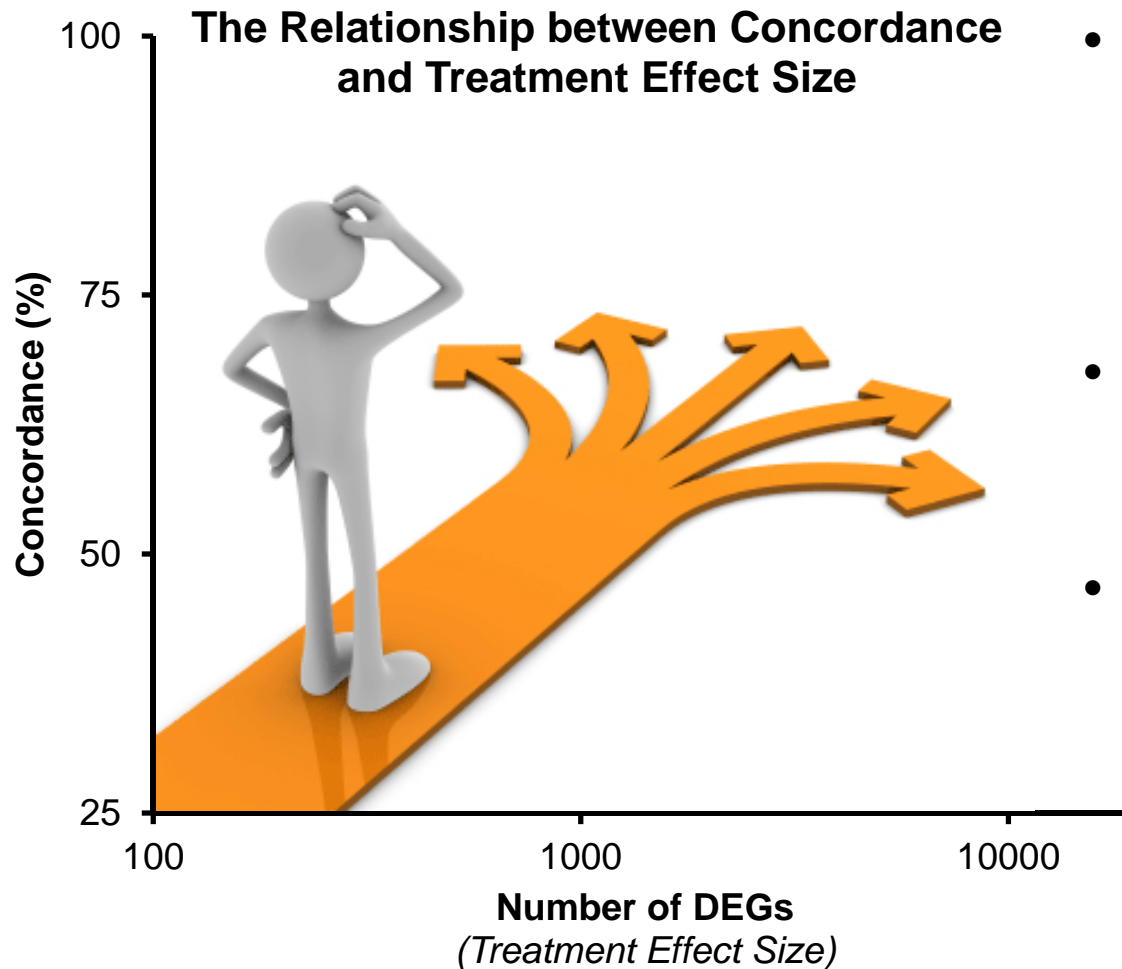
P Value cut-off < 0.05

Treatment Effect Size varies in Real-world application



- The concordance varies from 47% to 93% (74% in average).
- The concordance is generally higher for the highly expressed genes (24 of 27 chemicals), but not much.
- The concordance is linearly correlated with the treatment effect size (#DEG).

Explore on Larger Datasets



- **TG-GATE**
 - ~ 150 chemicals
 - human and rat
 - *in vivo* and *in vitro*
 - liver and kidney
- **MAQC Neuroblastoma dataset**
 - ~ 500 patients
 - 5 stages
- **SEQC Rat BodyMap dataset**
 - 11 organs
 - 4 development stages
 - male and female

ACKNOWLEDGEMENT

- **NCTR/FDA**
- **Fudan University**
- **SAS Institute Inc.**
- **ECETOC** (European Centre for Ecotoxicology and Toxicology of Chemicals)
- **MAQC/SEQC Society**
- **Supervisor: Dr. Weida Tong**
- **Dr. Joshua Xu**
- **Dr. Hong Fang**
- **Dr. James Chen**
- **Dr. Wenjun Bao**
- **Dr. Tzu-Ming Chu**
- **Dr. Leming Shi**
- **Mingrui Li**



