CS2220 Introduction to Computational Biology
Student Presentations

This presentation contributes 15% to the course grade

You may choose to earn up to 15% of the course grade by picking a paper below and making a presentation on it from 18 October 2012 onwards.

You will be graded according to:

- the quality of your ppt (readability, organization, attractiveness)
- the quality of your presentation (organization, delivery, Q&A)
- the level of understanding of what your are presenting
- inputs from your fellow students

Background

The possibility of using gene expression profiling by microarrays for diagnostic and prognostic purposes has also generated much excitement and research in the last ten years. Nevertheless, a number of issues persist such as how to rectify batch effects (i.e., non-biological variations) [bolstad-2003], how to handle missing values [troyanskaya-2001] and, most importantly, how to identify genes that are meaningful in explaining the difference in disease phenotypes [zhang-2009]. There are three main groups of approaches, that make use of biological pathways (e.g., enzymatic pathways, gene regulatory pathways, and protein interaction networks), for improving gene selection and for transitioning from the selected genes to the understanding of the sequences of causative molecular events. The first group are the overlap analysis methods [doniger-2003, khatri-2005, zeeberg-2003], which test the significance of the intersection of differentially expressed genes with a biological pathway. The second group are the direct group analysis methods [goeman-2004, kim-2005, subramanian-2005], which test whether a biological pathway is differentially expressed as a whole. The third group are the network-based analysis methods [chuang-2007, sivachenko-2007, sohler-2004, soh-2012], which zoom into a subnetwork of a biological pathway and test whether the subnetwork is differentially expressed. All of these approaches have their basis on the fact that every disease phenotype has some underlying biological causes. Therefore, it is reasonable to analyse the gene expression profiles of disease phenotype with respect to the biological contexts provided by biological pathways and protein interaction networks.

Group I: Issues in Microarray Analysis

[1. bolstad-2003] B. M. Bolstad, R. A. Irizarry, M. Astrand, T. P. Speed. **A comparison of normalization methods for high density oligonucleotide array data based on variance and bias**. *Bioinformatics*, 19(2):185-193, 2003.

[2. troyanskaya-2001] O. Troyanskaya, M. Cantor, G. Sherlock, P. Brown, et al. **Missing value estimation methods for DNA microarrays**. *Bioinformatics*, 17(6):520-525, 2001.

[3. zhang-2009] M. Zhang, L. Zhang, J. Zou, C. Yao, et al. **Evaluating reproducibility of differential expression discoveries in microarray studies by considering correlated molecular changes**. *Bioinformatics*, 25(13):1662-1668, 2009.

[4. venet-plos2011] D. Venet, J. E. Dumont, V. Detours. **Most random gene expression signatures are significantly associated with breast cancer outcome**. *PLoS Computational Biology*, 7(10):e1002240, 2011.

Group II: Overlap-Based Approaches

[5. khatri-2005] P. Khatri, S. Draghici. **Ontological analysis of gene expression data: Current tools, limitations, and open problems**. *Bioinformatics*, 21(18):3587-3595, 2005.

[6. zeeberg-2003] B. R. Zeeberg, W. Feng, G. Wang, M. D. Wang, et al. **GoMiner: A resource for biological interpretation of genomic and proteomic data**. *Genome Biology*, 4(4):R28, 2003.

Group III: Direct Group Approaches

[7. goeman-2004, **FCS**] J. J. Goeman, S. A. van de Geer, F. de Kort, H. C. van Houwelingen. **A global test for groups of genes: Testing association with a clinical outcome**. *Bioinformatics*, 20(1):93-99, 2004.

[8. kim-2005, **PAGE**] S. Y. Kim, D. J. Volsky. **PAGE: Parametric analysis of gene set enrichment**. *BMC Bioinformatics*, 8(6):144, 2005.

[9. subramanian-2005,**GSEA**] A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, et al. **Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles**. *Proc. Nat. Acad. Sci. USA*, 102(43):15545-15550, 2005.

Group IV: Network-Based Approaches

[10. sivachenko-2007, **NEA**] A. Y. Sivachenko, A. Yuryev, N. Daraselia, I. Mazo. **Molecular networks in microarray analysis**. *Journal of Bioinformatics and Computational Biology*, 5(2b):429-546, 2007.

[11. soh-2012.pdf, **SNet**] D. Soh, D. Dong, Y. Guo, L. Wong. **Finding consistent disease subnetworks across microarray datasets**. *BMC Genomics*, 12(Suppl 13):S15, 2011.

[12. hanczar-cabios07.pdf] B. Hanczar, J. D. Zucker, C. Henegar, L. Saitta. **Feature construction from synergic pairs to improve microarray-based classification**. *Bioinformatics*, 23(21):2866—2872, 2007.

[13. draghici-gr07.pdf, **Pathway Express**] S. Draghici et al. **A systems biology approach for pathway level analysis**. *Genome Research*, 17:1537-1545, 2007.

Group V: Model-Based Approaches for the Adventurous ☺

[14. geistlinger-cabios11] Geistlinger et al. **From sets to graphs: Towards a realistic enrichment analysis of transcriptomic systems**. *Bioinformatics*, 27(13):i366—i373, 2011.

[15. zampieri-cabios11] Zampieri et al. **A system-level approach for deciphering the transcriptional response to prion infection**. *Bioinformatics*, 27(24): 3407--3414, 2011.

[16. chindelevitch-cabios12] Chindelevitch et al. **Causal reasoning on biological networks: Interpreting transcriptional changes**. *Bioinformatics*, 28(8):1114-1121, 2012.

Grading Scheme

You will be graded according to:

- the quality of your ppt (readability, organization, attractiveness)
- the quality of your presentation (organization, delivery, Q&A)
- the level of understanding of what your are presenting

Your marks for the presentation will be the average of the inputs from your classmates and myself using the distribution scheme below:

	poor	ok	super	remarks
quality of ppt	10	20	30	
quality of presentation	10	20	30	
level of understanding	10	20	30	
gone beyond the paper assigned	0	0	10	

A computational biologist often has to communicate with biologists or computer scientists who either do not have sufficient background in computing, mathematics, or biology. The inputs from your classmates are especially important for assessing whether your presentation is sufficiently clear and easily understood by such non-experts.

Presenter #1: _____

Grader:_____

	poor	ok	super	remarks
quality of ppt	10	20	30	
quality of presentation	10	20	30	
level of understanding	10	20	30	
gone beyond the paper assigned	0	0	10	

Presenter #2: _____

Grader: _____

	poor	ok	super	remarks
quality of ppt	10	20	30	
quality of presentation	10	20	30	
level of understanding	10	20	30	
gone beyond the paper assigned	0	0	10	

Presenter #3:_____

Grader:_____

	poor	ok	super	remarks
quality of ppt	10	20	30	
quality of presentation	10	20	30	
level of understanding	10	20	30	
gone beyond the paper assigned	0	0	10	

Presenter #4:_____

Grader:_____

	poor	ok	super	remarks
quality of ppt	10	20	30	
quality of presentation	10	20	30	
level of understanding	10	20	30	
gone beyond the paper assigned	0	0	10	

Presenter #5: _____

Grader:_____

	poor	ok	super	remarks
quality of ppt	10	20	30	
quality of presentation	10	20	30	
level of understanding	10	20	30	
gone beyond the paper assigned	0	0	10	

Presenter #6: _____

Grader: _____

	poor	ok	super	remarks
quality of ppt	10	20	30	
quality of presentation	10	20	30	
level of understanding	10	20	30	
gone beyond the paper assigned	0	0	10	

Presenter #7:_____

Grader:_____

	poor	ok	super	remarks
quality of ppt	10	20	30	
quality of presentation	10	20	30	
level of understanding	10	20	30	
gone beyond the paper assigned	0	0	10	

Presenter #8:_____

Grader:_____

	poor	ok	super	remarks
quality of ppt	10	20	30	
quality of presentation	10	20	30	
level of understanding	10	20	30	
gone beyond the paper assigned	0	0	10	

Presenter #9: _____

Grader:_____

	poor	ok	super	remarks
quality of ppt	10	20	30	
quality of presentation	10	20	30	
level of understanding	10	20	30	
gone beyond the paper assigned	0	0	10	

Presenter #10: _____

Grader: _____

	poor	ok	super	remarks
quality of ppt	10	20	30	
quality of presentation	10	20	30	
level of understanding	10	20	30	
gone beyond the paper assigned	0	0	10	

Presenter #11:_____

Grader:_____

	poor	ok	super	remarks
quality of ppt	10	20	30	
quality of presentation	10	20	30	
level of understanding	10	20	30	
gone beyond the paper assigned	0	0	10	

Presenter #12:_____

Grader:_____

	poor	ok	super	remarks
quality of ppt	10	20	30	
quality of presentation	10	20	30	
level of understanding	10	20	30	
gone beyond the paper assigned	0	0	10	

Presenter #13: _____

Grader:_____

	poor	ok	super	remarks
quality of ppt	10	20	30	
quality of presentation	10	20	30	
level of understanding	10	20	30	
gone beyond the paper assigned	0	0	10	

Presenter #14: _____

Grader: _____

	poor	ok	super	remarks
quality of ppt	10	20	30	
quality of presentation	10	20	30	
level of understanding	10	20	30	
gone beyond the paper assigned	0	0	10	

Presenter #15:_____

Grader:_____

	poor	ok	super	remarks
quality of ppt	10	20	30	
quality of presentation	10	20	30	
level of understanding	10	20	30	
gone beyond the paper assigned	0	0	10	

Presenter:_____

Grader:_____

	poor	ok	super	remarks
quality of ppt	10	20	30	
quality of presentation	10	20	30	
level of understanding	10	20	30	
gone beyond the paper assigned	0	0	10	