Our first meeting

• Major charge of this task force is to identify and prioritize IT needs within BUMC in domains of research, education, clinical, administrative etc.....

• Strategic plan needed by June 2009, but our goal is to define needs in advance and submit to several committees
But why me?

- Our lab’s experience and interaction with IT over last 6 years represents microcosm of issues facing many within BUMC

- Solutions (and mistakes) we have adopted are worth exploring

- Links to CTSI Bioinformatics Core
Why IT is important to our lab

• Study molecular changes in epithelial cells that line the respiratory tract in smokes

• Develop early diagnostic biomarkers for lung cancer
The problem........

Use molecular tools that allow us to measure the expression of all genes in human genome in a single experiment

22,500 genes/patient $\times$ 22 spots/gene $\times$ $\sim$400 patients = $\sim$200 million data points

Preprocessing, analysis, storage and dissemination

Only the tip of the iceberg: all-exon array measures $\sim$5 million datapoints per sample!
Our solutions....

- Bought server and stored it on E6 with Microarray Core
- Hired part-time “consultant”/administrative assistant to manage server
- Graduate students building individual databases for each project
- High-school student (Newton North) built website and wiki over summer
- LinGA server for computation- Andi
The good......

• Papers, grants, 15 minutes of fame

Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer


Lung cancer is the leading cause of death from cancer in the U.S. and the world. The high mortality rate (80-85% within 5 years) results, in part, from a lack of effective tools to diagnose the disease at an early stage. Given that cigarette smoke elicits a field of injury throughout the airway, we sought to determine if gene expression in histologic normal large airway epithelial cells obtained at bronchoscopy from smokers with suspicion of lung cancer could be used as a lung cancer biomarker. Using a training set (n = 37) and gene expression profiles from Affymetrix HU-133A microarrays, we identified 80 gene biomarkers that could be used in the classification of smokers with and without lung cancer. We used the biomarker on an independent test set (n = 56), with an accuracy of 83% (80% sensitive, 86% specific), and on an additional validation set independently obtained from five medical centers (n = 35). Our data suggest that for stage 1 cancer across all smoking groups. Combining cytologie of lower airway cells obtained by bronchoscopy with the biomarker yielded 95% sensitivity and a 95% negative predictive value. Those findings indicate that gene expression in cytologically normal large-airway epithelial cells can serve as a lung cancer biomarker, potentially owing to a cancer-specific airway-wide response to cigarette smoke.
The bad....

Administrative assistant running website business off of our server

Constantly running out of space.....

Intermittent server shut down due to server room flooding, temperature fluctuations etc....
The truly ugly.....

Sold more Viagra in 2007 than did Pfizer
Why our group is invested in this process

Gene Expression Profiles

Proteomic Profiles
Metabolic Profiles

Genomic Data
SNPs
Epigenetic

Clinical Data
Treatments
Family history
Demographics
Environmental

Integrated Models

Predictions:
Risk
Individualized Prognosis & Diagnosis
Drug Response
Environment (eg Diet) Response