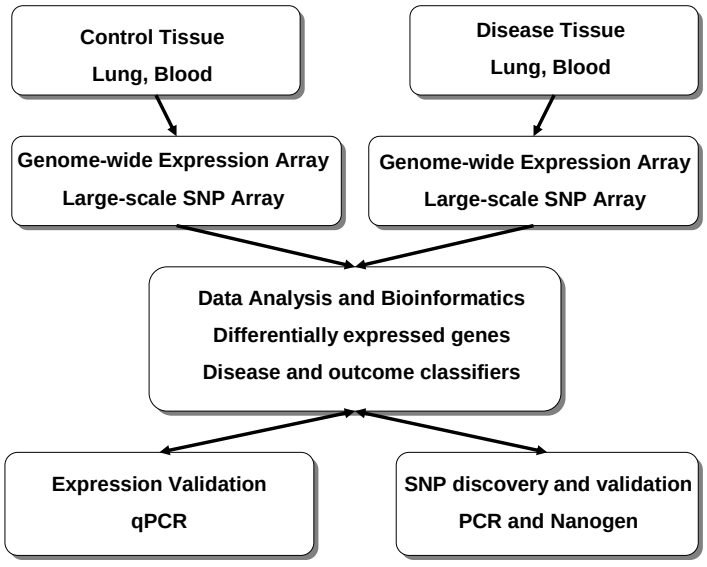
Mark W. Geraci, M.D.: University of Colorado, Denver

Center for Genomics

To provide excellence in genomic analysis for CMREF IPAH researchers.

Rationale: Methods for genomic analysis have proven powerful tools for identification of potential new pathobiology, biomarkers, and drug targets. The field of genomics and its associated technologies is rapidly evolving. Initially, the use of expression arrays became a powerful avenue for hypothesis generation and discovery. More recently, the ability to perform high-throughput SNP analysis has also enabled investigators to explore the genome in disease states. With such density of SNPs on an array, it is now possible to consider large-scale association studies, such as was done in patients with esophageal cancer. Complex traits can also be examined, as has been performed in inbred murine strains. Indeed, SNP analysis has been shown to increase information content compared to traditional microsatellite approaches. With the density and content of these arrays increasing, the ability to perform powerful association studies and examining quantitative trait loci should possible. Affymetrix currently has available a 100k SNP set, with plans to release a 500k set within the year. As in the early days of microarray expression analysis, handling such data has proven to be complex, but newer algorithms, some freely available in the public domain, are becoming established. As a further step to improve statistical power, clusters of SNPs can be evaluated en block, as haplotypes, which reduces the numbers of statistical tests performed and lessens the chance of type I errors due to false positives.



With the "state of the art" being, in fact, a moving target, we propose that establishing a compressive expression and SNP approach may provide the most information for future research. A general schema for the evaluation of tissues and blood by both expression analysis and SNP analysis is shown below in the figure to the left.

By providing a well-characterized control data set (as detailed in Specific Aim 2), annotated with clinical information, expression analysis and large-scale SNP determination, comparison to disease tissue is made much more meaningful and robust.