

## **Comparison of SAGE, cDNA microarray, and Affymetrix gene expression platforms for consistency and biological relevance of large-scale global co-expression analyses**

**Obi Griffith**, Erin Pleasance, Debra Fulton, Misha Bilenky, Sheldon McKay, Mehrdad Oveisi, Peter Ruzanov, Kim Wong, Scott Zuyderduyn, Asim Siddiqui and Steven Jones  
Genome Sciences Centre, British Columbia Cancer Agency, Vancouver, British Columbia, Canada, V5Z 4E6

We have conducted a comprehensive comparison of three major expression platforms: cDNA microarray, oligonucleotide microarray and serial analysis of gene expression (SAGE) using large sets of available data for *Homo sapiens*. Several studies have compared two of the three platforms to evaluate the consistency of expression profiles for a single tissue or sample set but none have determined if these translate into reliable co-expression patterns for global analyses across many conditions. To this end we analyzed a recently published data set of 1202 cDNA microarray experiments (Stuart *et al.* 2003), 214 SAGE libraries from CGAP and internal sources, and 525 Affymetrix (HG-U133A) oligonucleotide microarray experiments from GEO. All expression data were assigned to LocusLink Ids resulting in an overlap set of 4169 unambiguously mapped genes represented in all three platforms. Using standard co-expression analysis methods, we have assessed each platform for internal consistency and performed all pairwise platform comparisons. Internal consistency was determined by randomly dividing the datasets in half and comparing the Pearson distances for each subset: Affymetrix gave an  $r = 0.96$ , SAGE an  $r = 0.92$ , and microarray an  $r = 0.58$  ( $p < 0.001$ ). Despite these levels of internal consistency, all pairwise comparisons found poor correlation between platforms ( $r < 0.1$ ,  $p < 0.001$ ). Comparison against the Gene Ontology (GO) showed that all three platforms identify more co-expressed gene pairs with common GO biological process annotations than random data. However, SAGE and Affymetrix performed equally best with microarray performing only slightly better than random.