

Postdoctoral position in bioinformatics/statistics for interrogation of omics data to find causes of cancer

Our laboratory seeks to discover causes of cancer by comparing metabolomic and adductomic profiles between disease cases and controls to find discriminating features. Analyses employ two dedicated platforms for liquid chromatography-high resolution mass spectrometry, which can measure thousands of small molecules (metabolomics) and hundreds of protein modifications at particular loci (adductomics) in blood specimens. Blood samples have been obtained from case-control studies nested within prospective cohorts, which archived blood years prior to cancer diagnosis. Cancers being investigated (with NIH funding) include childhood leukemia, breast cancer, prostate cancer, colorectal cancer, lung cancer, and lymphoid cancers. Each experimental run generates gigabytes of data that must be acquired, stored, and processed, using in-house software written in R. Statistical aspects of the project include: experimental design, exploratory data analysis (EDA), quality control (QC), pre-processing (e.g., QC-based sample and feature filtering, imputation, normalization), feature selection, and performance assessment/validation of identified discriminating features.

We seek a candidate, having a recent **PhD in statistics, biostatistics, or bioinformatics/computational biology** (or related discipline), with the ability to work in an interdisciplinary team of chemists, biochemists, epidemiologists, statisticians, and health scientists. Demonstrated facility with the R programming language (<https://www.r-project.org>) is essential and experience with R software packages for omics data (Bioconductor Project; <http://www.bioconductor.org>) and systems for computationally reproducible research is highly desirable (e.g., knitr, Jupyter). Candidates are expected to have expertise in statistical methods for high-dimensional inference, including: exploratory data analysis, visualization, regularized regression, smoothing, resampling, and ensemble prediction and variable selection methods.

For further information, contact Professor Stephen Rappaport, at the UC Berkeley School of Public Health (srappaport@berkeley.edu), or Professor Sandrine Dudoit, at the UC Berkeley Division of Biostatistics and Department of Statistics (<https://www.stat.berkeley.edu/~sandrine/>; sandrine@stat.berkeley.edu).

Start date: As soon as possible.

Funding: One year, with option to renew a second year.

Application materials: Curriculum vitae, statement of research interests, names and e-mail addresses of 3 references.