



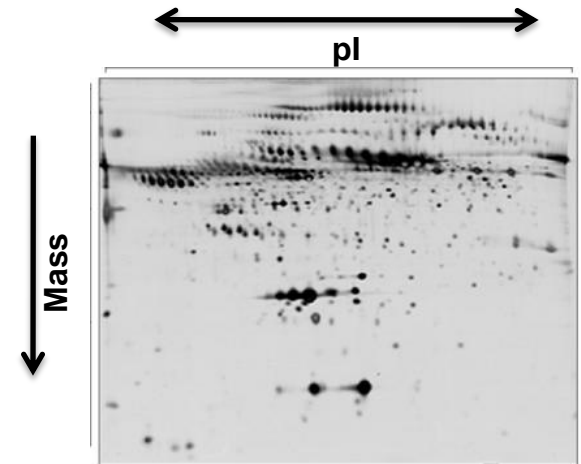
digeR: a graphical user interface R package for analyzing 2D-DIGE (gel) data

Yue Fan

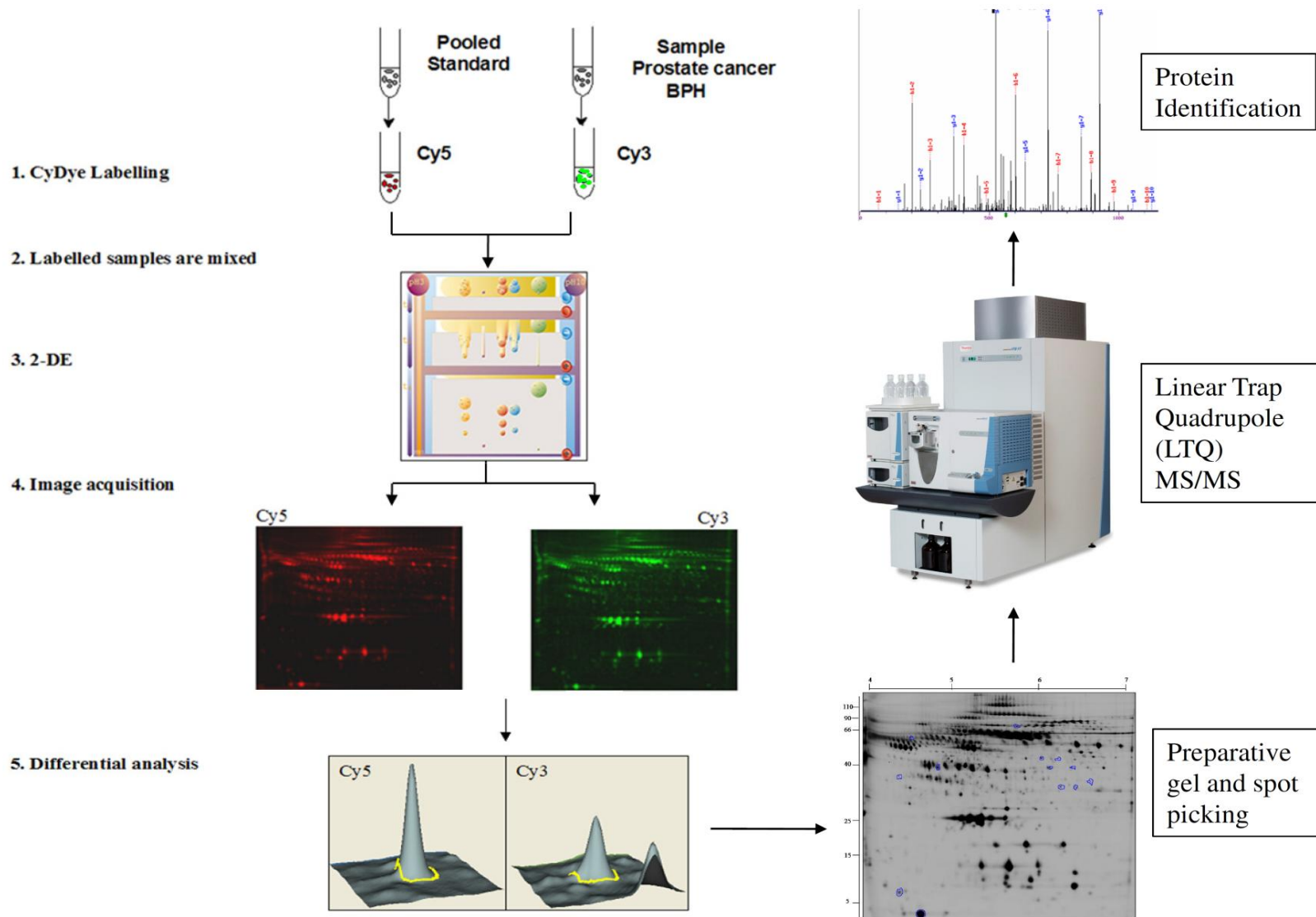
Conway Institute
UCD School of Medicine and Medical Science
University College Dublin, Ireland

2D gel (2D-DIGE)

- 2D gel
 - Isoelectric point (pI) and mass
 - Intact protein level quantitation
 - Post translational modification (PTMs)
- 2D-DIGE:
 - Fluorescent labeling: higher sensitivity
 - Better normalization and reduced variation



2D-DIGE work flow



Data analysis in 2D gel study

- Control FDR
- Dimension reduction
- Classification use one or multiple biomarkers
- Prioritization of identified biomarkers
- Potential PTMs and different protein isoforms
- Number of gels required to differentiate groups

R package for biologist

- Graphical user interface
- Easy parameter setting
 - Flexibility and default value
- Generate publishable plots
- Identify best predictive protein spots
- Power analysis helps experiment design

digeR functions

- Spot correlation analysis
 - Identify potential PTMs
- Score plot
 - Dimension reduction
- Classification
 - Single or multiple biomarkers predictive ability
- Feature selection
 - Prioritization of identified biomarkers
- Power analysis
 - Number of gels for the study



Fan, Y., Murphy, T.B. and Watson, R.W.G. (2009) digeR: a graphical user interface R package for analyzing 2D-DIGE data, *Bioinformatics*, 25, 3033-3034.

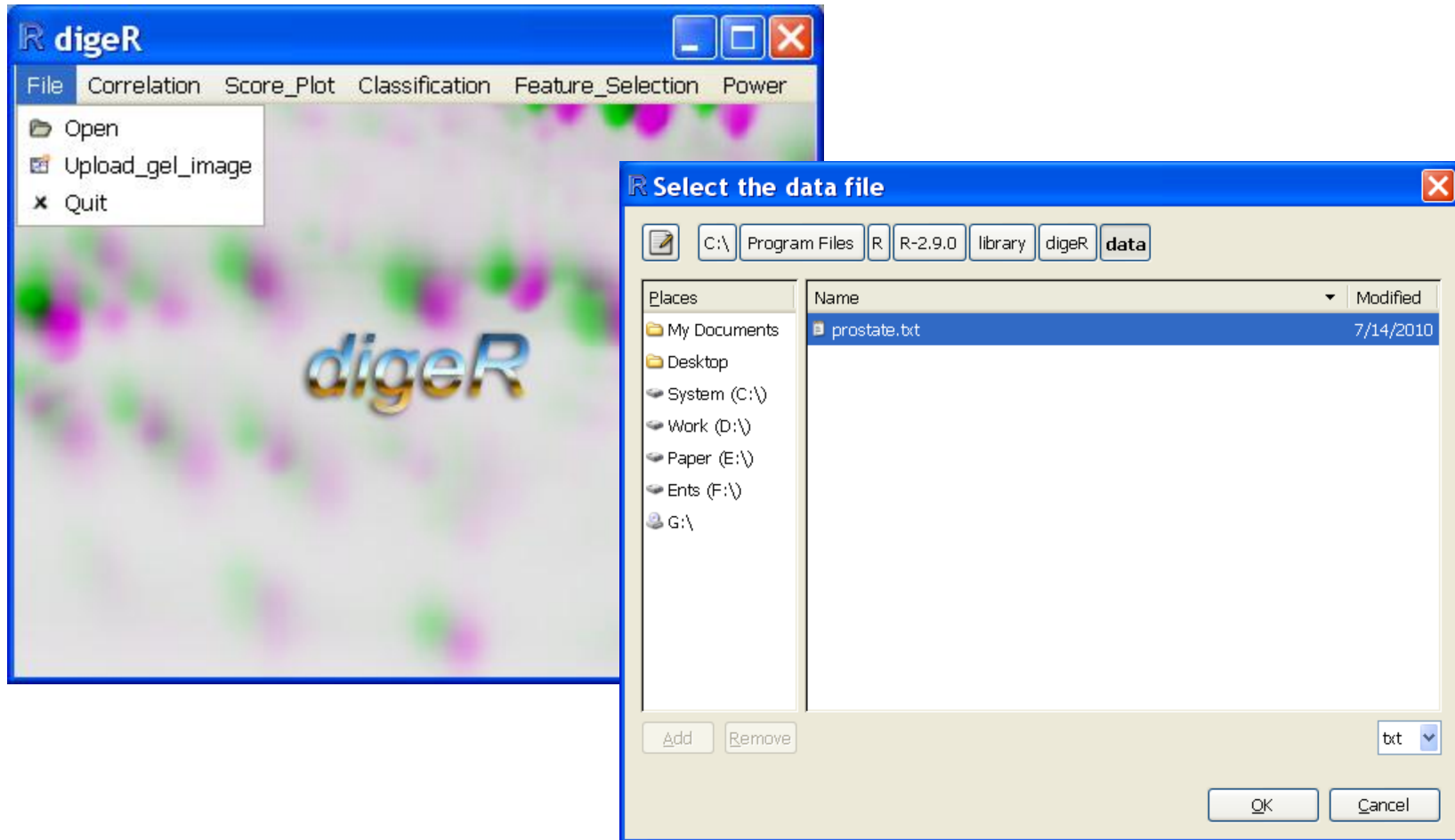
Package dependency

- Graphical user interface
 - Gtk, gWidgets and gWidgetsRGtk2
- Score plot
 - caTools, pls and elipse
- Classification
 - MASS, pls, e1071 and ROCR
- Feature selection
 - MASS, pls, randomForest and adabag

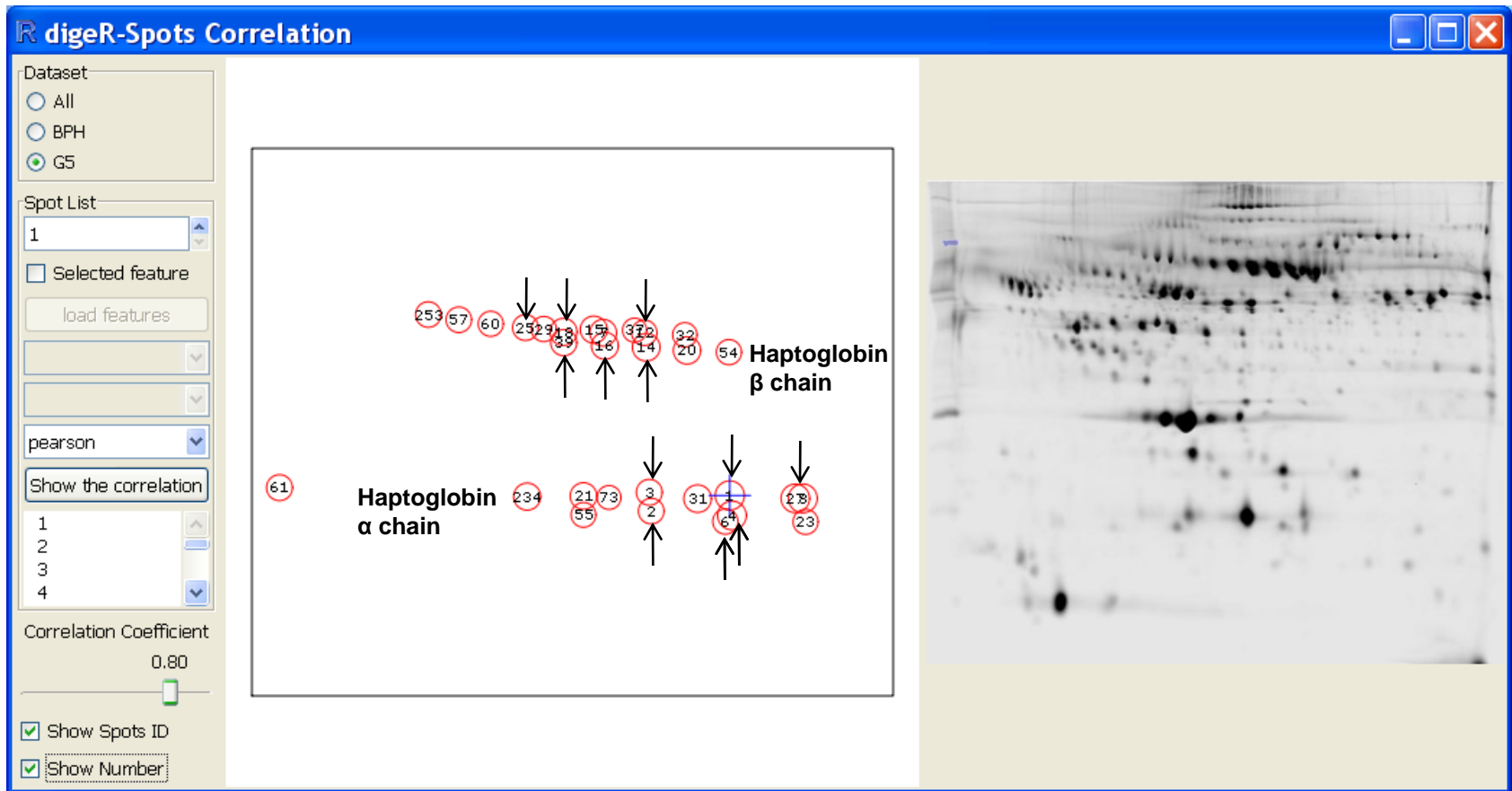
Example: Prostate Cancer

- PSA is not specific in prostate cancer diagnosis, especially BPH patient
- 2D-DIGE study with 50 serum samples from BPH and prostate cancer patient.
- 32 samples (14 BPH and 18 prostate cancer), 747 spots

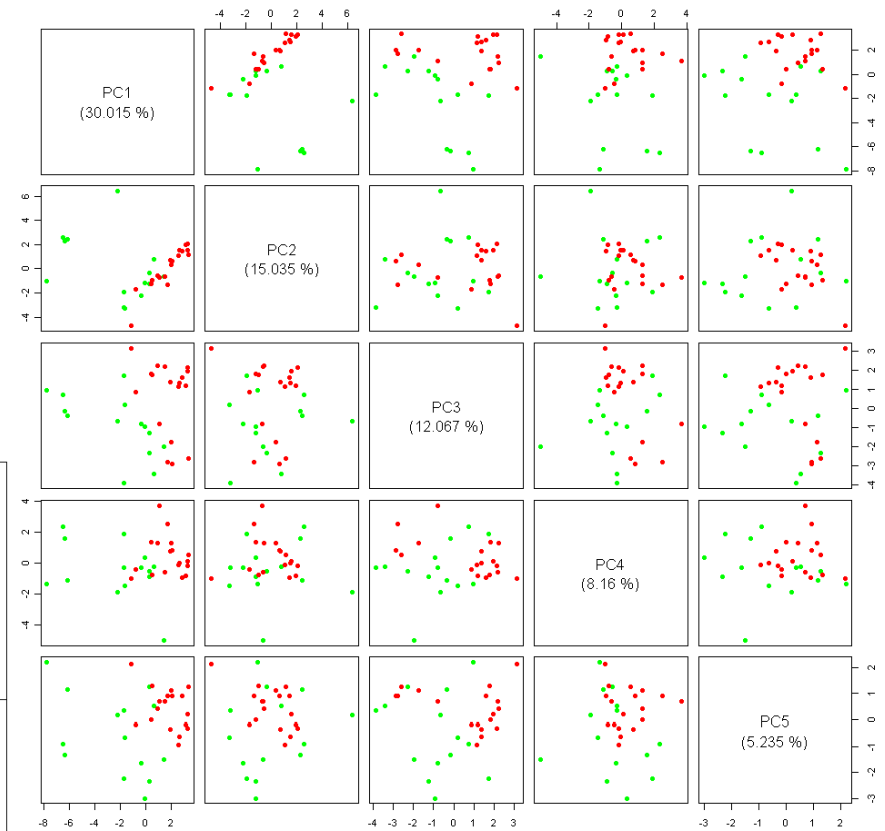
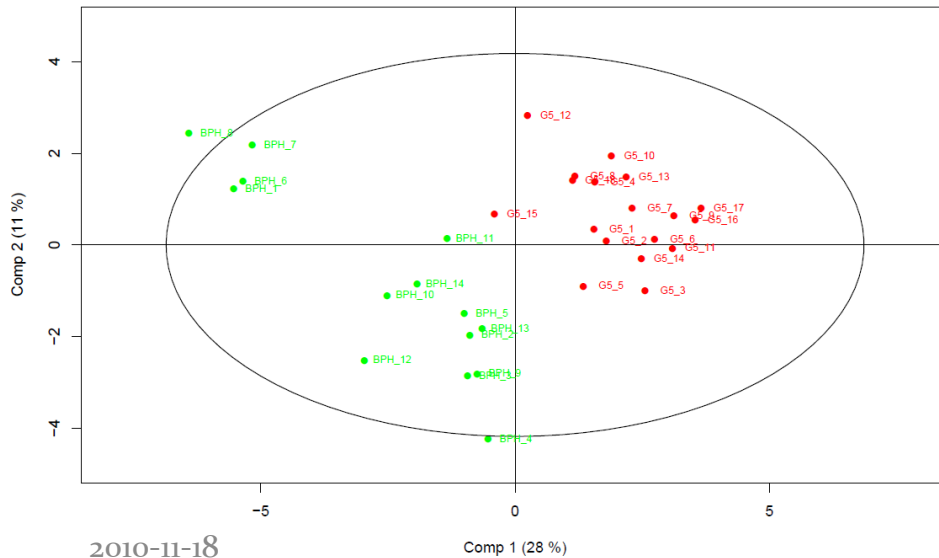
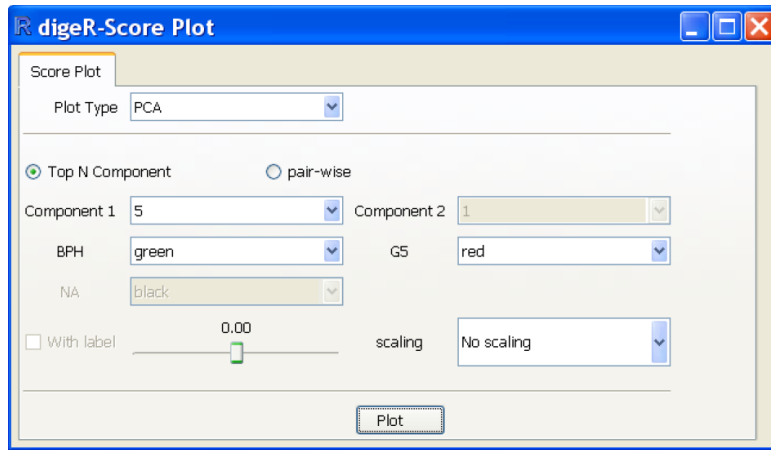
Data import



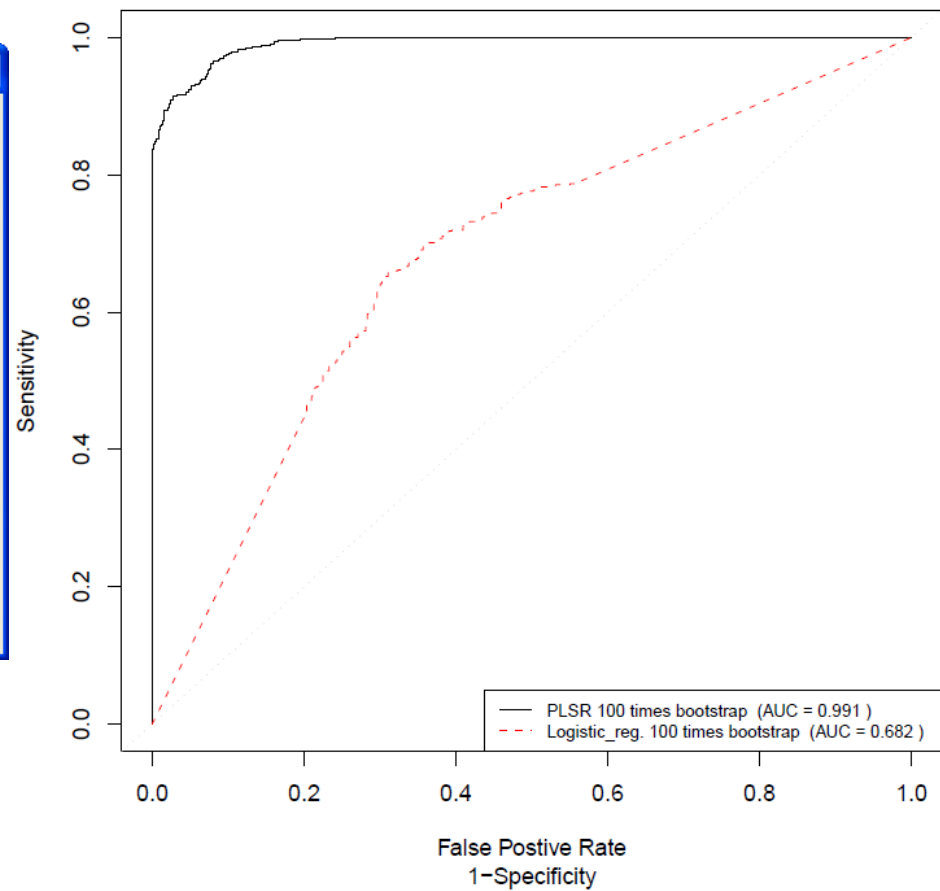
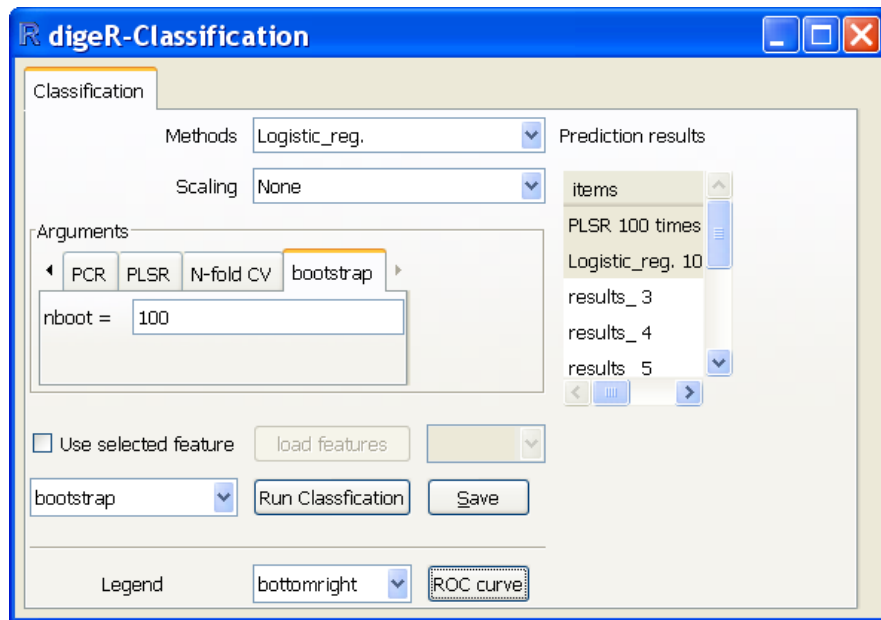
Spot correlation



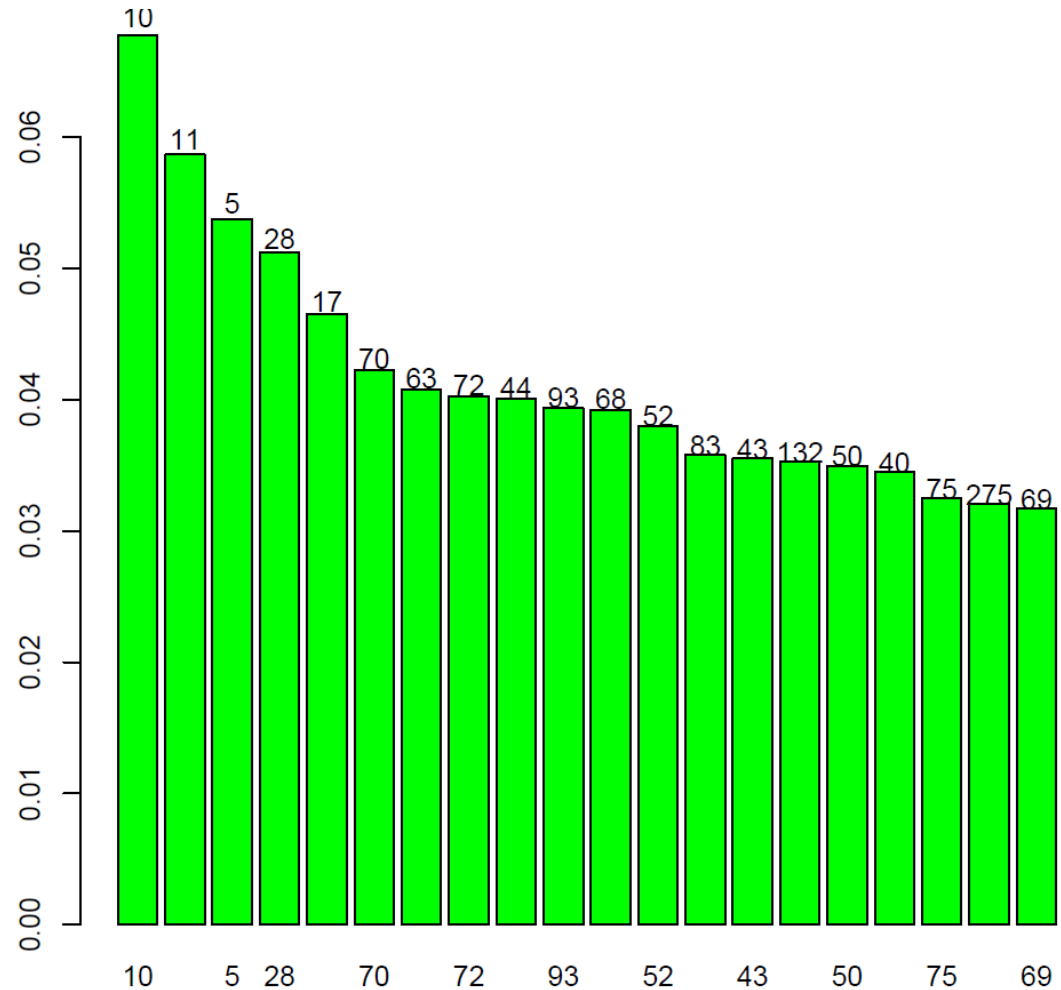
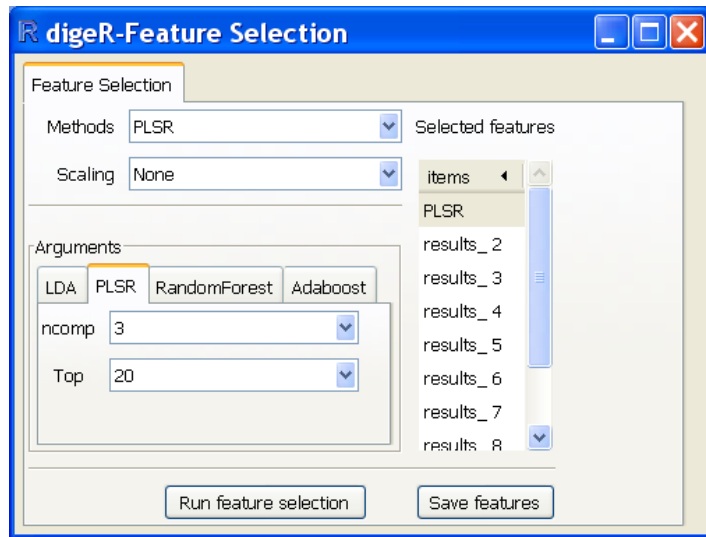
Score plot



Classification



Feature selection



Power analysis

digeR-Sample size

Power Analysis

Single spots Spot number

Gel

Arguments

Effect size Significant level

Power Sample size Per group

Effect size is estimated from data

Hwang, D. et al. (2002) *Determination of minimum sample size and discriminatory expression patterns in microarray data.*

Bioinformatics, **18**, 1184–1193

Next: digeR v1.3

- Interactive plot
- Allow randomly split samples into training and testing
- Apply feature selection with cross validation/bootstrapping
- Test prediction performance through label permutation

Acknowledgements

Prof. Bill Watson

Prof. Brendan Murphy

Dr. Jennifer Byrne

Prof. Des Higgins

Dr. Kenneth Bryan

Irish Prostate Cancer Research Consortium

