13.4 Conclusion

We have reviewed Bayesian approaches for variable selection for linear settings and for mixture models and have described novel extensions that aim at addressing important problems in the analysis of genomic data. Bayesian variable selection techniques can cope with a large number of regressors and can handle a number of covariates larger than the sample size. These methods allow the evaluation of the joint effect of sets of variables and the use of stochastic search techniques to explore the high-dimensional variable space. The approaches we have presented offer a coherent framework in which variable selection and prediction, in the linear setting, and variable selection and sample classification/clustering, in the mixture models, are performed simultaneously. In addition, the flexible prior models allow the incorporation of additional information in quite a natural way. We have seen in particular how gene–gene networks can be captured via MRF priors for the analysis of microarray data. We have also investigated graphical models, integrating gene expression data with microRNA expression data.

Acknowledgments

Vannucci's research is partially supported by National Institutes of Health/ National Heart, Lung, and Blood Institute grant no. P01-HL082798 and National Science Foundation/Division of Mathematical Sciences grant no. 1007871.

References

Brown, P. J. 1993. *Measurement, Regression and Calibration*. Clarendon.

Brown, P. J., Vannucci, M., and Fearn, T. 1998. Multivariate Bayesian variable selection and prediction. *J. R. Stat. Soc. B*, **60**, 627–641.

Brown, P. J., Vannucci, M., and Fearn, T. 2002. Bayes model averaging with selection of regressors. *J. R. Stat. Soc. B*, **64**, 519–536.

Chipman, H., George, E., and McCulloch, R. E. 2001. The practical implementation of Bayesian model selection. In: *Model Selection*. IMS, pp. 67–116.

Clyde, M., and George, E. I. 2004. Model Uncertainty. *Stat. Sci.*, **19**, 81–94.

Clyde, M. A., DeSimone, H., and Parmigiani, G. 1996. Prediction via orthogonalized model mixing. *J. Am. Stat. Assoc.*, **91**, 1197–1208.

Dawid, A. P. 1981. Some matrix-variate distribution theory: notational considerations and a Bayesian application. *Biometrika*, **68**, 265–274.

Denkert, C., Winzer, K. J., and Hauptmann, S. 2004. Prognostic impact of cyclooxygenase-2 in breast cancer. *Clin. Breast Cancer*, **4**, 428–33.

Dudoit, S., Fridlyand, J., and Speed, T. P. 2002. Comparison of discrimination methods for the classification of tumors using gene expression data. *J. Am. Stat. Assoc.*, **97**, 77–87.