The kinship2 R Package for Pedigree Data

Jason P. Sinnwell Terry M. Therneau Daniel J. Schaid

Division of Biomedical Statistics and Informatics, Department of Health Sciences Research, Mayo Clinic,
Rochester, Minn., USA

Key Words

Pedigrees · Genetic linkage analysis · Kinship · Graphics

Abstract

Background: The kinship2 package is restructured from the previous kinship package. Existing features are now enhanced and new features added for handling pedigree objects. **Methods:** Pedigree plotting features have been updated to display features on complex pedigrees while adhering to pedigree plotting standards. Kinship matrices can now be calculated for the X chromosome. Other methods have been added to subset and trim pedigrees while maintaining the pedigree structure. **Conclusion:** We make the kinship2 package available for R on the Contributed R Archives Network (CRAN), where data management is built-in and other packages can use the pedigree object.

© 2014 S. Karger AG, Basel

Introduction

Pedigrees have long been used in genetic linkage and association studies, and pedigree kinship matrices are widely employed in mixed-effects models for survival and regression analyses. The original kinship package, developed by Terry Therneau and ported to R by Jing

Hua Zhao [1], was developed to accompany the coxme package [2] which extends the Cox model to include kinship matrices for related subjects. The kinship2 package has major updates to the core functions to calculate kinship matrices, plot pedigrees, and trim pedigree objects. Family-based studies for linkage, association, and mixed effects need the capability to keep track of pedigree relationships, phenotypes, covariates, and subjects that are informative for the analyses. We utilize the object-oriented framework within R to manage pedigree objects, where the objects contain data for pedigree members.

Methods

Pedigree Plots

Pedigree plots are widely used to check the accuracy of pedigree data. Our pedigree plotting routines within R have compared favorably to other stand-alone and R plotting routines [1, 3–5]. Now in kinship2, we have made the plots more robust and adhere to more of the pedigree plotting standards discussed in Bennett et al. [6]. We provide an example of some of the key additions in figure 1; additional features are demonstrated in the online supplementary material (www.karger.com/doi/10.1159/000363105).

Pedigree Trimming

Pedigree analyses are often based on a subset of pedigree members that are informative for a specific analysis. For example, subjects who are not genotyped and do not have offspring are not

References

- 1 Zhao JH: Pedigree-drawing with R and graphviz. *Bioinformatics* 2006;22:1013–1014.
- 2 Therneau TM, Grambsch PM: *Modeling Survival Data: Extending the Cox Model*. New York, Springer, 2000.
- 3 Fuchsberger C, Falchi M, Forer L, Pramstaller PP: PedVizApi: a Java API for the interactive, visual analysis of extended pedigrees. *Bioinformatics* 2008;24:279–281.
- 4 Trager EH, Khanna R, Marrs A, Siden L, Branham KE, Swaroop A, Richards JE: Madeline 2.0 PDE: a new program for local and web-based pedigree drawing. *Bioinformatics* 2007;23:1854–1856.
- 5 Mäkinen VP, Parkkonen M, Wessman M, Groop PH, Kanninen T, Kashi K: High-throughput pedigree drawing. *Eur J Hum Genet* 2005;13:987–989.
- 6 Bennett RL, French KS, Resta RG, Doyle DL: Standardized human pedigree nomenclature: update and assessment of the recommendations of the National Society of Genetic Counselors. *J Genet Couns* 2008;17:424–433.
- 7 Abecasis GR, Cherny SS, Cookson WO, Cardon LR: Merlin – rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002;30:97–101.
- 8 Lange K: *Mathematical and Statistical Methods for Genetics Analysis*. New York, Springer, 1997.
- 9 Kent JW Jr, Dyer TD, Blangero J: Estimating the additive genetic effect of the X chromosome. *Genet Epidemiol* 2005;29:377–388.
- 10 Schifano ED, et al: SNP set association analysis for familial data. *Genet Epidemiol* 2012, Epub ahead of print.
- 11 Schaid DJ, McDonnell SK, Sinnwell JP, Thibodeau SN: Multiple genetic variant association testing by collapsing and kernel methods with pedigree or population structured data. *Genet Epidemiol* 2013;37:409–418.