

1. Record Nr.	UNINA9910299769803321
Autore	Gondro Cedric
Titolo	Primer to Analysis of Genomic Data Using R // by Cedric Gondro
Pubbl/distr/stampa	Cham : , : Springer International Publishing : , : Imprint : Springer, , 2015
ISBN	3-319-14475-8
Edizione	[1st ed. 2015.]
Descrizione fisica	1 online resource (283 p.)
Collana	Use R!, , 2197-5736
Disciplina	572.838
Soggetti	Statistics Gene expression Microarrays R (Computer program language) Statistics for Life Sciences, Medicine, Health Sciences Statistics and Computing/Statistics Programs Gene Expression
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Description based upon print version of record.
Nota di bibliografia	Includes bibliographical references.
Nota di contenuto	R basics -- Simple marker association tests -- Genome wide association studies -- Population and genetic architecture -- Gene expression analysis -- Databases and functional information -- Extending R -- Final comments -- Index -- References.
Sommario/riassunto	Through this book, researchers and students will learn to use R for analysis of large-scale genomic data and how to create routines to automate analytical steps. The philosophy behind the book is to start with real world raw datasets and perform all the analytical steps needed to reach final results. Though theory plays an important role, this is a practical book for advanced undergraduate and graduate classes in bioinformatics, genomics and statistical genetics or for use in lab sessions. This book is also designed to be used by students in computer science and statistics who want to learn the practical aspects of genomic analysis without delving into algorithmic details. The datasets used throughout the book may be downloaded from the publisher's website. Chapters show how to handle and manage high-throughput genomic data, create automated workflows and speed up

analyses in R. A wide range of R packages useful for working with genomic data are illustrated with practical examples. In recent years R has become the de facto tool for analysis of gene expression data, in addition to its prominent role in the analysis of genomic data. Benefits to using R include the integrated development environment for analysis, flexibility and control of the analytic workflow. At a time when genomic data is decidedly big, the skills from this book are critical. The key topics covered are association studies, genomic prediction, estimation of population genetic parameters and diversity, gene expression analysis, functional annotation of results using publically available databases and how to work efficiently in R with large genomic datasets. Important principles are demonstrated and illustrated through engaging examples which invite the reader to work with the provided datasets. Some methods that are discussed in this volume include: signatures of selection; population parameters (LD, FST, FIS, etc); use of a genomic relationship matrix for population diversity studies; use of SNP data for parentage testing; snpBLUP and gBLUP for genomic prediction. Step-by-step, all the R code required for a genome-wide association study is shown: starting from raw SNP data, how to build databases to handle and manage the data, quality control and filtering measures, association testing and evaluation of results, through to identification and functional annotation of candidate genes. Similarly, gene expression analyses are shown using microarray and RNAseq data. .
