

"An excellent book for beginners and occasional practitioners"
Reviewer, *Journal of the American Medical Association*

Bioinformatics

FOR

DUMMIES®

2nd Edition

**A Reference
for the
Rest of Us!®**

FREE eTips at dummies.com®

Jean-Michel Claverie, PhD

Research Director, France's Centre National
de la Recherche Scientifique (CNRS)

Cedric Notredame, PhD

Professor of Bioinformatics, Switzerland's
Lausanne University and the CNRS

Updated to cover
multiple new
genomes and
databases



Bioinformatics

FOR

DUMMIES[®]

2ND EDITION

**by Jean-Michel Claverie, PhD
and Cedric Notredame, PhD**

Bioinformatics

FOR

DUMMIES®

2ND EDITION

Bioinformatics

FOR

DUMMIES®

2ND EDITION

**by Jean-Michel Claverie, PhD
and Cedric Notredame, PhD**

For Dummies®, 2nd Edition

ng, Inc.

et

7030-5774

m

7 by Wiley Publishing, Inc., Indianapolis, Indiana

Wiley Publishing, Inc., Indianapolis, Indiana

Simultaneously in Canada

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted in writing by Permissions 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 646-8600. For all other use, permission should be addressed to the Legal Department, Wiley Publishing, 111 River Street, Hoboken, NJ 07030, (201) 748-6000, fax (201) 748-6050, or online at www.wiley.com/go/permissions.

This book contains the Wiley Publishing logo, For Dummies, the Dummies Man logo, A Reference for the Dummies Way, Dummies Daily, The Fun and Easy Way, Dummies.com, and related trade names or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates in the United States and other countries, and may not be used without written permission. All other trademarks are the property of their respective owners. Wiley Publishing, Inc., is not associated with any product or vendor mentioned in this book.

LIMITATION OF WARRANTY: THE PUBLISHER AND THE AUTHOR MAKE NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE ACCURACY OR COMPLETENESS OF THE CONTENTS OF THIS WORK AND SPECIFICALLY DISCLAIM ALL WARRANTIES, INCLUDING WITHOUT LIMITATION ANY IMPLIED WARRANTIES OF FITNESS FOR A PARTICULAR PURPOSE. NO WARRANTY MAY BE CREATED BY SALES OR PROMOTIONAL MATERIALS. THE ADVICE AND STRATEGIES CONTAINED HEREIN MAY NOT BE SUITABLE FOR EVERY SITUATION. THIS WORK IS SOLD WITH THE UNDERSTANDING THAT THE PUBLISHER IS NOT ENGAGED IN RENDERING LEGAL, ACCOUNTING, OR OTHER PROFESSIONAL SERVICES. IF PROFESSIONAL ASSISTANCE IS REQUIRED, THE SERVICES OF A QUALIFIED PROFESSIONAL PERSON SHOULD BE SOUGHT. NEITHER THE PUBLISHER NOR THE AUTHOR SHALL BE LIABLE FOR DAMAGES ARISING HEREFROM. THE FACT THAT AN ORGANIZATION OR PERSON IS REFERRED TO IN THIS WORK AS A CITATION AND/OR A POTENTIAL SOURCE OF INFORMATION DOES NOT MEAN THAT THE AUTHOR OR THE PUBLISHER ENDORSES THE

the Authors

Jean-Michel Claverie is Professor of Medical Bioinformatics at the School of Medicine of the Université de la Méditerranée, and a consultant in genomics and bioinformatics. He is the founder and current head of the Structural & Genomic Information Laboratory, located in Marseilles, a sunny city on the Mediterranean coast of France. Using science as a pretext to travel, Jean-Michel has held positions in Paris (France), Sherbrooke (PQ, Canada), the Salk Institute (La Jolla, CA), the Pasteur Institute (Paris), Incyte pharmaceutical (Palo Alto, CA); and the National Center for Biotechnology Information (Bethesda, MD). He has used computers in biology since the early days — his Ph.D. work involved modeling biochemical reactions by programming an 8K Honeywell 516 computer right from the console switches! Although he has no clear recollection of it, he has been credited with introducing the French word “bioinformatique” in the late eighties, before involuntarily coining the catchy “bioinformatics” by mistranslating it while giving a talk in English!

Jean-Michel’s current research interests are in microbial and structural genomics, and in the development of bioinformatic methods for the prediction of gene function. He is the author or coauthor of more than 150 scientific publications, and a member of numerous international review panels and scientific councils. In his spare time, he enjoys the relaxed pace of life in Marseilles, with his wife Chantal and their two sons, Nicholas and Raphael.

ation

his is for my parents Monique and Jack, for keeping me in school, and for
hantal, for keeping me happy — in and out of the lab. It's also for my daugh-
er Vanessa, and my sons Nicholas and Raphael, for reminding me that not
everything in life is scientific.

- J-MC

his is for my wife Marita, my daughter Lina, my mother Marie and in
emory of my grandparents, Simone and Louis.

- CN

rs' Acknowledgments

he entire Wiley staff did a great job pulling together to publish this book on
ght deadlines. We'd especially like to thank our tireless project editor, Paul
evesque, and Barry Childs-Helton, who did a great job copyediting a text full
f obscure biochemical words.

We'd also like to thank Amey Godse, our technical editor. Amey nailed down
major and minor inaccuracies alike. His many suggestions did much to
improve the book