

# **Systems Genetics**

Linking Genotypes and Phenotypes

Whereas genetic studies have traditionally focused on explaining heritance of single traits and their phenotypes, recent technological advances have made it possible to comprehensively dissect the genetic architecture of complex traits and to quantify how genes interact to shape phenotypes. This exciting new area has been termed systems genetics and is born out of a synthesis of multiple fields, integrating a range of approaches and exploiting our increased ability to obtain quantitative and detailed measurements on a broad spectrum of phenotypes.

Gathering the contributions of leading scientists, both computational and experimental, this book shows how experimental perturbations can help us to understand the link between genotype and phenotype. A snapshot of current research activity and state-of-the-art approaches to systems genetics is provided, including work from model organisms such as *Saccharomices cerevisiae*, *Drosophila melanogaster*, as well as from human studies.

Researchers and graduate students in genetics, functional genomics, bioinformatics, computational biology, systems biology, and biotechnology will find this a valuable and timely resource.

**Florian Markowetz** is a Group Leader at Cancer Research UK's Cambridge Research Institute. His research is concerned with developing statistical and mathematical models of complex biological systems and analysing large-scale molecular data. His research interests range from the analysis of molecular clinical data to inference of cellular networks from high-throughput gene perturbation screens and integration of heterogeneous data sources using machines learning techniques and probabilistic graphic models.

**Michael Boutros** is a Group Leader at the German Cancer Research Centre (DKFZ) in Heidelberg, where he heads the Division of Signalling and Functional Genomics. He also holds a Professorship at the University of Heidelberg. His research focuses on the systematic dissection signalling pathways in *Drosophila* and mammalian cells, which are important during development and cancer. He attempts to define key components of signalling pathways, discovering interaction between pathways, and characterisation of signalling networks under normal and perturbed conditions.



## **CAMBRIDGE SERIES IN SYSTEMS GENETICS**

**Series Editors:** 

**Jason A. Moore, PhD**University of Pennsylvania

Scott M. Williams, PhD

Geisel School of Medicine at Dartmouth College

Systems Genetics is the study of DNA sequence variation and biological traits as a complex system characterised by spatial and temporal interactions. This includes phenomena such as epigenetics, epistasis, plastic reaction norms, and locus heterogeneity. The Cambridge Series in Systems Genetics covers all areas of genetics approached from a complex systems point of view. This series is of interest to researchers across several areas of the life sciences including bioinformatics, evolution, genomics, human genetics, molecular genetics, precision medicine, and systems biology.



# **Systems Genetics**

# Linking Genotypes and Phenotypes

Edited by

FLORIAN MARKOWETZ

Cancer Research UK Cambridge Institute

MICHAEL BOUTROS

German Cancer Research Center, Heidelberg





# **CAMBRIDGE**UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781107013841

© Cambridge University Press 2015

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2015

Printed in the United Kingdom by TJ International Ltd., Padstow, Cornwall

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data
Systems genetics: linking genotypes and phenotypes / edited by
Florian Markowetz, Cancer Research UK Cambridge Institute,
Michael Boutros, German Cancer Research Center, Heidelberg.
pages cm
Includes index.
ISBN 978-1-107-01384-1
1. Phenotype. 2. Genetics. 3. Functional genomics.

Phenotype.
 Genetics.
 Functional genomics.
 Markowetz, Florian.
 Boutros, Michael.
 QH438.5.S97
 2015

576.5–dc23

2014050295

ISBN 978-1-107-01384-1 Hardback

Additional resources for this publication at www.cambridge.org/SystemsGenetics

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.



# **Contents**

	List	of contributors	page ix
1	An introduction to systems genetics Florian Markowetz and Michael Boutros		1
	1.1	Definition of systems genetics	1
	1.2	History of systems genetics	3
	1.3	Future challenges	7
	1.4	What is covered in the book	8
2	Computational paradigms for analyzing genetic interaction networks Carles Pons, Michael Costanzo, Charles Boone, and Chad L. Myers		12
	2.1	Definition of genetic interaction	12
	2.2	Toward the first reference global genetic interaction network: Synthetic	
		Genetic Array analysis in yeast	15
	2.3	Computational paradigms for genetic interaction networks	17
	2.4	Perspectives	29
3	Mapping genetic interactions across many phenotypes in metazoan cells Christina Laufer, Maximilian Billmann, and Michael Boutros		
	3.1	A short history of genetic interaction analysis	36
	3.2	Perturbation-based genetic interaction studies in yeast	37
	3.3	Genetic interaction analysis in <i>Drosophila</i>	39
	3.4	Expanding genetic interaction mapping towards the genomic scale	42
	3.5	Towards genetic interaction mapping in human cells	45
	3.6	Conclusions	48
4	Genetic interactions and network reliability		
	Edgar Delgado-Eckert and Niko Beerenwinkel		
	4.1	Biological networks	51
	4.2	Epistasis	52



۷İ

6

Cambridge University Press 978-1-107-01384-1 - Systems Genetics: Linking Genotypes and Phenotypes Edited by Florian Markowetz and Michael Boutros Frontmatter More information

**Contents** 

screens

Conclusions

Xin Wang, Ke Yuan, and Florian Markowetz

5.4

6.1

	4.3	Network reliability	
	4.4	Epistasis on networks	
	4.5	Inferring function from observed genetic interactions	
	4.6	Conclusions	
5	Synthetic lethality and chemoresistance in cancer Kimberly Maxfield and Angelique Whitehurst		
	5.1	Cancer chemotherapy	
	5.2	Employing small interfering RNA (siRNA) to identify modifiers of chemotherapeutic responsiveness	
	5.3	Mobilizing new therapeutic opportunities with large-scale RNAi	

Joining the dots: network analysis of gene perturbation data

	6.2	Scenario 2: Single gene silenced, multi-level dynamic phenotype	86
	6.3	Scenario 3a: Pathway components perturbed with global	
		transcriptional phenotypes	86
	6.4	Scenario 3b: Capturing rewiring events during network evolution	92
	6.5	Scenario 4: Multi-parametric screen, up to genome-wide	96
	6.6	Conclusions	102
7	High-content screening in infectious diseases: new drugs against bugs		108
	André	P. Mäurer, Peter R. Braun, Kate Holden-Dye, and Thomas F. Meyer	

Scenario 1: Genome-wide screens with single reporters

7.1	The challenge of fighting infectious diseases	108
7.2	Classic strategies for antimicrobial drug development and their	
	limitations	109
7.3	Post-genomic approaches for investigating host-pathogen interactions	115
7.4	Advanced high-content screening in pathogen research	121
7.5	Single-cell population analyses in high-content screening	129
7.6	Future directions	131

8	Inferring genetic architecture from systems genetics studies	139
	Xiaoyun Sun, Stephanie Mohr, Arunachalam Vinayagam, Pengyu Hong, and Norbert Perrimon	

8.2	Identification of network components by RNAi	141
8.3	Identification of network components using proteomics	145
8.4	Integration of RNAi and proteomic data sets	148
8.5	Network modeling: the next step	149
8.6	Applications of network reconstruction	155

8.1

Introduction

139

54575961

65

65

68

74

77

83

83



		Contents	vii	
9	Bayesian inference for model selection: an application to aberrant signalling pathways in chronic myeloid leukaemia			
	Lisa E. M. Hopcroft, Ben Calderhead, Paolo Gallipoli, Tessa L. Holyoake, and Mark A. Girolami			
	9.1	The oncology of chronic myeloid leukaemia	161	
	9.2	Introduction to model comparison	170	
	9.3	Modelling the JAK/STAT pathway in response to TKI and/or JakI	171	
	9.4	The statistical methodology: Riemannian manifold population MCMC	174	
	9.5	A proof-of-concept study with synthetic data	178	
	9.6	Beyond a proof of concept: considering a more biologically		
		realistic dataset	181	
	9.7	Discussion	187	
10	Dynai	mic network models of protein complexes	191	
	Yongjin Park and Joel S. Bader			
	10.1	Dynamic network data	191	
	10.2	Block models of a network	195	
	10.3	Learning algorithms	197	
	10.4	Results	203	
	10.5	Discussion	209	
11	Phenotype state spaces and strategies for exploring them			
	Andreas Hadjiprocopis and Rune Linding			
	11.1	Introduction	214	
	11.2	Phenotype: a constructive generality	215	
	11.3	Cellular noise	216	
	11.4	Genome evolution, protein families, and phenotype	217	
	11.5	Complex networks	222	
	11.6	Random Boolean networks	223	
	11.7	Genomic state spaces	229	
12		nated behavioural fingerprinting of <i>Caenorhabditis elegans</i> mutants	234	
	André E. X. Brown and William R. Schafer			
	12.1	The worm as a model organism	234	
	12.2	High-throughput data collection and information extraction	238	
	12.3	Linking behaviours and genes	243	
	12.4	Outlook	246	
	12.5	Conclusions	251	
	Index		257	

The colour plate section can be found between pages 148 and 149.





# **Contributors**

#### Joel S. Bader

Johns Hopkins University, Baltimore, MD 21218, USA

#### **Niko Beerenwinkel**

Department of Biosystems Science and Engineering, ETH Zurich, CH-4058 Basel, Switzerland

#### **Maximilian Billmann**

German Cancer Research Center (DKFZ) and Heidelberg University, Division of Signaling and Functional Genomics, D-69120 Heidelberg, Germany

#### **Charles Boone**

Banting and Best Department of Medical Research and Department of Molecular Genetics, Terrence Donnelly Center for Cellular and Biomolecular Research, University of Toronto, Toronto, ON M5S 3E1, Canada

#### **Michael Boutros**

German Cancer Research Center (DKFZ) and Heidelberg University, Division of Signaling and Functional Genomics, D-69120 Heidelberg, Germany

#### Peter R. Braun

Steinbeis-Innovationszentrum, Center for Systems Biomedicine, D-14612, and Max Planck-Institüt für Infektionsbiologie, D-10117 Berlin

# André E. X. Brown

MRC Clinical Sciences Centre, Faculty of Medicine, Imperial College, Du Cane Road, London, W12 0NN, UK

#### **Ben Calderhead**

Centre for Mathematics and Physics in the Life Sciences and Experimental Biology (CoMPLEX), University College London, London, WC1E 6BT, UK

# **Michael Costanzo**

Banting and Best Department of Medical Research and Department of Molecular Genetics, Terrence Donnelly Center for Cellular and Biomolecular Research, University of Toronto, Toronto, ON M5S 3E1, Canada



#### x Contributors

#### **Edgar Delgado-Eckert**

Department of Biosystems Science and Engineering, ETH Zurich, CH-4058 Basel, Switzerland

## Paolo Gallipoli

Paul O'Gorman Leukaemia Research Centre, College of Medical, Veterinary and Life Sciences, Institute of Cancer Sciences, University of Glasgow, Glasgow, G12 0ZD, UK

#### Mark A. Girolami

Department of Statistical Science, University College London, London, WC1E 6BT, LIK

#### **Andreas Hadjiprocopis**

The Institute of Cancer Research (ICR), London, SW3 6JB, UK

#### Kate Holden-Dve

Steinbeis-Innovationszentrum, Center for Systems Biomedicine, D-14612 Berlin, Germany

#### Tessa L. Holyoake

Paul O'Gorman Leukaemia Research Centre, College of Medical, Veterinary and Life Sciences, Institute of Cancer Sciences, University of Glasgow, Glasgow, G12 0ZD, UK

#### Penavu Hona

Department of Computer Science, Volen Center for Complex Systems, Brandeis University, Waltham, MA 02454, USA

#### Lisa E. M. Hopcroft

School of Computing Science, College of Science and Engineering, University of Glasgow, Glasgow, G12 8QQ, and Stem Cell and Leukaemia Proteomics Laboratory, Institute of Cancer Sciences, University of Manchester, Manchester, M20 3LJ, UK

#### **Christina Laufer**

German Cancer Research Center (DKFZ) and Heidelberg University, Division of Signaling and Functional Genomics, D-69120 Heidelberg, Germany

#### **Rune Linding**

Cellular Signal Integration Group (C-SIG), Center for Biological Sequence Analysis (CBS), Department of Systems Biology, Technical University of Denmark (DTU), DK-2800 Lyngby, Denmark

# Florian Markowetz

Cancer Research UK Cambridge Institute, Cambridge, CB2 0RE, UK

#### André P. Mäurer

Steinbeis-Innovationszentrum, Center for Systems Biomedicine, Haydnallee 21, D-14612, and Max Planck-Institüt für Infektionsbiologie, Charitéplatz 1, D-10117 Berlin, Germany



**Contributors** 

Χİ

#### **Kimberly Maxfield**

University of North Carolina – Chapel Hill School of Medicine, Chapel Hill, NC 27599, USA

# Thomas F. Meyer

Max Planck-Institüt für Infektionsbiologie, Charitéplatz 1, D-10117 Berlin, Germany

#### **Stephanie Mohr**

Department of Genetics, Harvard Medical School, Boston, MA 02115, USA

#### Chad L. Myers

Department of Computer Science and Engineering, University of Minnesota-Twin Cities, Minneapolis, MN 55455, USA

## Yongjin Park

Johns Hopkins University, Baltimore, MD 21218, USA

#### **Norbert Perrimon**

Department of Genetics, Harvard Medical School, Boston, MA, USA, and Howard Hughes Medical Institute, Harvard Medical School, Boston, MA 02115, USA

#### **Carles Pons**

Department of Computer Science and Engineering, University of Minnesota-Twin Cities, Minneapolis, MN 55455, USA

# William R. Schafer

Medical Research Council Laboratory of Molecular Biology, Cambridge, CB2 0QH, UK

# Xiaoyun Sun

Department of Computer Science, Volen Center for Complex Systems, Brandeis University, Waltham, MA 02454, USA

#### **Arunachalam Vinayagam**

Department of Genetics, Harvard Medical School, Boston, MA 02115, USA

#### Xin Wang

Cancer Research UK Cambridge Institute, Cambridge, CB2 0RE, UK

# **Angelique Whitehurst**

University of North Caroline – Chapel Hill School of Medicine, Chapel Hill, NC 27599, USA

#### Ke Yuan

Cancer Research UK Cambridge Institute, Cambridge, CB2 0RE, UK