Genome-Wide Association Studies
From Polymorphism to Personalized Medicine

Over the last 20 years, Genome-Wide Association Studies (GWAS) have revealed a great deal about the genetic basis of a wide range of complex diseases, and they will undoubtedly continue to have a broad impact as we move to an era of personalized medicine. This authoritative text, written by leaders and innovators from both academia and industry, covers the basic science as well as the clinical, biotechnological, and pharmaceutical potential of these methods.

With special emphasis given to highlighting pharmaco-genomics and population genomics studies using next-generation technology approaches, this is the first book devoted to combining association studies with single nucleotide polymorphisms, copy number variants, haplotypes and expressed quantitative trait loci. A reliable guide for newcomers to the field, as well as for experienced scientists, this is a unique resource for anyone interested in how the revolutionary power of genomics can be applied to solve problems in complex disease.

Krishnarao Appasani is the Founder and Chief Executive Officer of GeneExpression Systems, a global conference-producing organization focusing on biomedical and physical sciences. He is an award-winning scientist and also the editor of *Epigenomics: From Chromatin Biology to Therapeutics* (2012), *MicroRNAs: From Basic Science to Disease Biology* (2007), and *RNA Interference: From Basic Science to Drug Development* (2005), all published by Cambridge University Press.

Genome-Wide Association Studies

FROM POLYMORPHISM TO PERSONALIZED MEDICINE

Edited by

Krishnarao Appasani
GeneExpression Systems, Inc.

Forewords by

Stephen W. Scherer
University of Toronto, Canada

and

Peter M. Visscher
University of Queensland, Australia

CAMBRIDGE UNIVERSITY PRESS

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/9781107042766

First published 2016

Printed in the United States of America by Sheridan Books, Inc.

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

Library of Congress Cataloguing in Publication data
Genome-wide association studies : from polymorphism to personalized medicine / edited by Krishnarao Appasani ; forewords by Stephen W. Scherer and Peter M. Visscher.
p. ; cm.
Includes bibliographical references and index.
ISBN 978-1-107-04276-6 (hardback)
I. Appasani, Krishnarao, 1959–, editor.
[DNLM: 1. Genome – Wide Association Study. 2. Genetic Testing. 3. Genomics. QU 550]
RB155
616′.042–dc23

2015017921

ISBN 978-1-107-04276-6 Hardback

Dedicated to

My friend, ***Arthur Beck Pardee***
Emeritus Professor
Harvard University Medical School, USA

Contents

The colour plates are to be found between pages 222 and 223

Contributors

Krishnarao Appasani
GeneExpression Systems, Inc.
Waltham, Massachusetts, USA

Raghu Appasani
University of Massachusetts Medical School
Worcester, Massachusetts, USA &
MINDS Foundation USA & India
Boston, Massachusetts, USA

Neha S. Bhise
Department of Pharmacotherapy and Translational Research
Center for Pharmacogenomics
College of Pharmacy, University of Florida
Gainesville, FL, USA &
Department of Experimental and Clinical Research
University of Minnesota
St. Paul, Minnesota, USA

Elena Bosch
Institut de Biologia Evolutiva (UPF-CSIC)
Departament de Ciències Experimentals i de la Salut
Universitat Pompeu Fabra
Parc de Recerca Biomèdica de Barcelona
Barcelona, Catalonia, Spain

Erwin Brosens
PhD student Pediatric Surgery & Clinical Genetics
Erasmus MC–Sophia Children's Hospital
GE Rotterdam, The Netherlands

Margit Burmeister
Molecular & Behavioral Neuroscience Institute
University of Michigan
Ann Arbor, Michigan, USA

Ferran Casals
Genomics Core Facility
Departament de Ciències Experimentals i de la Salut
Universitat Pompeu Fabra
Barcelona, Catalonia, Spain

Lata Chauhan
Department of Pharmacotherapy and Translational Research
Center for Pharmacogenomics
College of Pharmacy
University of Florida
Gainesville, Florida, USA

Mengjie Chen
Department of Biostatistics
School of Public Health
Yale University
New Haven, Connecticut, USA &
Presently at Department of Biostatistics and Genetics
University of North Carolina Chapel Hill, NC, USA

Dongjun Chung
Department of Biostatistics
School of Public Health
Yale University
New Haven, Connecticut, USA

Eske M. Derks
Department of Psychiatry
Amsterdam Medical Center
University of Amsterdam
Amsterdam, The Netherlands

Nevenka Dimitrova
Philips Research
Briarcliff Manor, New York, USA

Martin Eden
Manchester Centre for Health Economics
The University of Manchester
Manchester, UK

H. Eussen
Department of Clinical Genetics
Erasmus Medical Centre
GE Rotterdam, The Netherlands

Joel Gelernter
Department of Genetics
Yale University School of Medicine
New Haven, Connecticut, USA

James Hicks
Cold Spring Harbor Laboratory
Cold Spring Harbor, New York, USA

Takae Hirasawa
Dept. of Epigenetics Medicine
Graduate School of Medicine and Engineering
University of Yamanashi
Yamanashi, Japan

Randall C. Johnson
BSP CCR Genetics Core
Leidos Biomedical Research, Inc
Frederick National Lab
Frederick, Maryland, USA

René S. Kahn
Department of Psychiatry
Rudolf Magnus Institute of Neuroscience
University Medical Center
Utrecht, The Netherlands

Sitharthan Kamalakaran
Philips Research North America
Briarcliff Manor, New York, USA

Tuomas O. Kilpeläinen
The Novo Nordisk Foundation Center for Basic Metabolic Research
Section of Metabolic Genetics
Faculty of Health and Medical Sciences
University of Copenhagen
Copenhagen, Denmark

A. de Klein
Department of Clinical Genetics

Erasmus Medical Centre
GE Rotterdam, The Netherlands

Takeo Kubota
Dept. of Epigenetics Medicine
Graduate School of Medicine and Engineering
University of Yamanashi
Yamanashi, Japan

Chao-Qiang Lai
Jean Mayer USDA Human Nutrition Research Center on Aging
Tufts University
Boston, Massachusetts, USA

Jatinder K. Lamba
Department of Pharmacotherapy and Translational Research
Center for Pharmacogenomics
College of Pharmacy, University of Florida
Gainesville, Florida, USA

Yu-Chi Lee
Jean Mayer USDA Human Nutrition Research Center on Aging
Tufts University
Boston, Massachusetts, USA

Cong Li
Department of Biostatistics
School of Public Health
Yale University
New Haven, Connecticut, USA

Yiyi Ma
Nutrition and Genomics Laboratory
Jean Mayer USDA Human Nutrition Research Center on Aging
Tufts University
Boston, Massachusetts, USA &
Presently at Boston University School of Medicine
Boston, Massachusetts, USA

Urko M. Marigorta
Institut de Biología Evolutiva
Universitat Pompeu Fabra
Biomedical Research Park (PRBB)
Barcelona, Spain &
School of Biology

Georgia Institute of Technology
Atlanta, GA, USA

Sadeep Medhasi
Division of Pharmacogenomics and Personalized Medicine
Mahidol University Faculty of Medicine
Bangkok, Thailand &
Somdech Phra Debaratana Medical Center
Ramathibodi Hospital
Bangkok, Thailand

Barry Merriman
Thermo Fisher Scientific, Inc.
Carlsbad, California, USA &
Presently at Human Longevity, Inc. San Diego, California, USA

Kunio Miyake
Department of Epigenetics Medicine
Graduate School of Medicine and Engineering
University of Yamanashi
Yamanashi, Japan

Arcadi Navarro
Institute of Evolutionary Biology
Universitat Pompeu Fabra, Biomedical Research Park
Barcelona, Spain &
National Institute for Bioinformatics, Barcelona, Catalonia, Spain &
Centre for Genomic Regulation, Barcelona, Catalonia, Spain &
Institució Catalana de Recerca i Estudis Avançats (ICREA)
Barcelona, Catalonia, Spain

Dorien H. Nieman
Department of Psychiatry
Amsterdam Medical Center
University of Amsterdam
Amsterdam, The Netherlands

Sheikh Nizamuddin
CSIR-Centre for Cellular and Molecular Biology
Hyderabad, India

José M. Ordovás
Department of Epidemiology
Centro Nacional Investigaciones Cardiovasculares &
Instituto Madrileño de Estudios Avanzados en Alimentación
Madrid, Spain

Kouichi Ozaki
Laboratory for Cardiovascular Diseases
RIKEN Center for Integrative Medical Sciences
Yokohoma City, Kanagawa, Japan

Laurence D. Parnell
Jean Mayer USDA Human Nutrition Research Center on Aging
Tufts University
Boston, Massachusetts, USA

Katherine Payne
Manchester Center for Health Economics
The University of Manchester
Manchester, UK

Apichaya Puangpetch
Division of Pharmacogenomics and Personalized Medicine
Mahidol University Faculty of Medicine
Bangkok, Thailand &
Somdech Phra Debaratana Medical Center
Ramathibodi Hospital
Bangkok, Thailand

Juan Antonio Rodriguez
Institut de Biología Evolutiva
Universitat Pompeu Fabra
Biomedical Research Park (PRBB)
Barcelona, Spain &
National Institute for Bioinformatics (INB), Spain

Pål Sætrom
Department of Computer and Information Science
Department of Cancer Research and Molecular Medicine
Norwegian University of Science and Technology
Trondheim, Norway

Stephen W. Scherer
The Centre for Applied Genomics
The Hospital for Sick Children
University of Toronto
Toronto, Ontario, Canada

Caren E. Smith
Jean Mayer USDA Human Nutrition Research Center on Aging
Tufts University
Boston, Massachusetts, USA

K.G. Snoek
Department of Pediatric Surgery
Erasmus Medical Centre–Sophia Children's Hospital
GE Rotterdam, The Netherlands

Sven Stringer
Department of Psychiatry
Amsterdam Medical Center
University of Amsterdam &
Department of Psychiatry
Rudolf Magnus Institute of Neuroscience
University Medical Center
Utrecht, The Netherlands

Chonlaphat Sukasem
Division of Pharmacogenomics and Personalized Medicine
Mahidol University Faculty of Medicine
Bangkok, Thailand &
Somdech Phra Debaratana Medical Center
Ramathibodi Hospital
Bangkok, Thailand

Rakesh Tamang
CSIR-Centre for Cellular and Molecular Biology &
Osmania University, Hyderabad &
University of Calcutta
Kolkata, India

Toshihiro Tanaka
Laboratory of Cardiovascular Diseases
RIKEN, Yokohoma City, Kanagawa, Japan &
Department of Human Genetics and Disease Diversity
Graduate School of Medical and Dental Sciences
Tokyo Medical and Dental University
Tokyo, Japan

Man-Hung Eric Tang
Cold Spring Harbor Laboratory
Cold Spring Harbor, New York, USA &
Department of Oncology
Clinical Sciences
Lund University, Lund, Sweden

Kumarasamy Thangaraj
Evolutionary and Medical Genetics Laboratory

CSIR-Centre for Cellular and Molecular Biology
Hyderabad, India

Laurent F. Thomas
Department of Cancer Research and Molecular Medicine
Norwegian University of Science and Technology
Trondheim, Norway

D. Tibboel
Department of Pediatric Surgery
Erasmus Medical Center–Sophia Children's Hospital
GE Rotterdam, The Netherlands

Elisa M. Trucco
Department of Psychiatry
Molecular & Behavioral Neuroscience Institute
University of Michigan
Ann Arbor, Michigan, USA

Vinay Varadan
Case Comprehensive Cancer Center
Case Western Reserve University
Cleveland, USA

D. Veenma
Department of Pediatric Surgery & Clinical Genetics
Erasmus Medical Centre–Sophia Children's Hospital
GE Rotterdam, The Netherlands

Sandra Villafuerte
Department of Psychiatry
Molecular & Behavioral Neuroscience Institute
University of Michigan
Ann Arbor, Michigan, USA

Peter M. Visscher
Queensland Brain Institute
University of Queensland
St. Lucia, Brisbane, Queensland, Australia

Edwin Wang
Bioinformatics and Systems Biology
National Research Council Canada &
McGill University Center for Bioinformatics
Montreal, Canada

Cheryl A. Winkler
Senior Investigator
SAIC-National Cancer Institute, NIH
Frederick, Maryland, USA

Can Yang
Department of Mathematics
Hong Kong Baptist University
Kowloon Tong, Hong Kong

Meredith Yeager
Core Genotyping Facility
SAIC-National Cancer Institute, NIH
Gaithersburg, Maryland, USA

Naif Zaman
Bioinformatics and Systems Biology
National Research Council Canada &
McGill University Center for Bioinformatics
Montreal, Canada

Michael Q. Zhang
Cold Spring Harbor Laboratory
Cold Spring Harbor, New York, USA &
The University of Texas at Dallas
Richardson, Texas, USA &
Tsinghua University
Beijing, China

Hongyu Zhao
Department of Biostatistics
School of Public Health
Yale University
New Haven, Connecticut, USA

Foreword

Stephen W. Scherer

I am grateful to Krishnarao Appasani for making the Herculean effort to prepare this volume on the rapidly expanding fields of Genome-Wide Association Studies (GWAS) and personalized medicine, and for inviting me to offer a few of my own introductory statements.

The complexity of the human disease state has always been an area of human curiosity. Over the last decade, GWASs have enabled us to expand our understanding of complex diseases using genetic-based approaches. We now see GWAS as a technology platform that promises to help move us into the era of personalized medicine.

Genome-Wide Association Studies: From Polymorphism to Personalized Medicine edited by Dr. Appasani has assembled the contributing chapters into five main areas encompassing: an introduction to GWAS in medicine, GWAS in pharmacogenomics, different classes of genetic variants for GWAS, new technologies including next-generation sequencing, and population genetics. The component chapters will be highly valuable, not only to those who are experimentally active in these aspects of research, but also to those interested in potential drug discovery applications. The historical perspectives offered also bring forward a unique vantage point into ongoing and future research in this field.

I believe that this book will become a reliable guide for anyone attempting to understand the successes with GWAS, planning new experiments, as well as its potential for the advancement of medicine. We approach a time where advances in genome sequencing technologies will deliver the long-awaited $1,000 genome, which promises to enable capture of all classes of genetic variants in a single experiment empowering new and innovative future studies. As such, studying the content of this book will allow us to pause and reflect, of where the field has come from, and where it needs to now go.

August 05, 2015

Professor Stephen W. Scherer, PhD. FRSC
Director of The Centre for Applied Genomics and
Director, McLaughlin Centre
University of Toronto
Toronto, Ontario, Canada

Foreword

Peter M. Visscher

Almost 10 years of GWAS: a feast of discoveries

Genome-wide association studies (GWASs) are less than 10 years old but have revolutionized discoveries in human genetics. GWAS is an experimental design based upon association because it exploits the fact that genetic variants that are close together tend to be statistically correlated. Driven by advances in array technologies these genome surveys of genetic variation have led to the discovery of thousands of DNA variants that are associated with complex traits, including many diseases. They have also led to new insights in human evolution and population differences.

Despite its undeniable success in finding many replicable associations between genetic variants and complex traits, GWAS has been criticized for a number of supposed flaws and with ever-shifting goalposts. Initial criticism was that it wouldn't work at all. This was followed by critique that not enough variants were detected per trait. When tens to hundreds of variants were detected for a disease or trait, the criticism became that not enough of the genetic variation was explained by those variants. Then, when a significant chunk of variation was accounted for, the perceived problem was that there was no biological insight. And finally, when that critique was shown by empirical data to be unjustified, the supposed fault with GWAS has become that its findings have not yet been translated into the clinic. And that for an experimental design less than 10 years old! I have never fully understood the criticism and attribute it to a mixture of ignorance about what association studies are designed for and prejudice with respect to the nature of genetic variation. As emphasized in this book, GWAS has shown that there are many genes with small effects that contribute to risk of common disease. That is not a shortcoming of the design, but revealing the true state of nature.

The facts speak for themselves, as highlighted and summarized nicely in this book. For many diseases and quantitative traits, tens to hundreds of robust associations have been discovered, including for diseases such as schizophrenia for which no genes were identified before the advent of GWAS. But GWAS has gone well beyond the discovery of robust associations. It has informed us about

the nature of trait variation – genetic variation underlying disease and quantitative traits is highly polygenic and a surprisingly large proportion of genetic variation is tagged by common SNPs. It had led to the discoveries of new biological pathways involved in disease and has led to new biological knowledge about how DNA variation can affect gene regulation. Although it is early days, there are also signs that GWAS findings will lead to new drug development and may contribute to the adoption of early intervention strategies by stratifying people according to their genetic risk.

This book contains chapters from many of the players who have contributed to GWAS discovery in the last decade. It provides a helpful overview of the many facets of genome-wide data, their analysis and the interpretation of discoveries. We are living in an era of high-throughput hypothesis generating science, and, as demonstrated in this book, GWAS is a testament to the variety of discoveries that can be made when taking a whole genome approach to genetics, combined with large experimental sample sizes.

July 20, 2015

Peter M. Visscher, PhD.
Professor and Chair of Quantitative Genetics
Queensland Brain Institute
University of Queensland
St. Lucia, Brisbane, Queensland, Australia

Preface

DNA neither cares nor knows. DNA just is. And we dance to its music
Richard Dawkins, English evolutionary biologist and science writer (1941–)

In 1994, more than two decades ago, Francis Collins (the present Director of National Institutes of Health) expressed his view that "finding genes is like trying to find a needle in a haystack." In the pre-genomic era (prior to the Human Genome Project), geneticists were primarily focused on family-based linkage studies that examined simple Mendelian disorders, hereditary diseases caused by the malfunction of a single gene. In those days, it was difficult to study complex diseases like cancer or diabetes, which are influenced by multiple genes and multiple environmental factors. After completion of the sequencing of the "human genome" in early 2001, the big buzz in biology switched to understanding the complex diseases. From that time onwards, scientific efforts became increasingly globalized and increasingly intimate collaboration between academia and industry developed. Advances in genomic technologies have contributed to the development and use of new tools such as Genome-Wide Association Studies. Genome-Wide Association Studies (GWAS) were designed to survey the role of common genetic variations in complex human diseases. It was thought that GWAS would have the advantage of not relying on prior knowledge of biological pathways compared with "candidate gene-based studies," and it was also expected that GWAS would have higher power and finer resolution to identify genetic variants in the genome.

Another global collaborative effort was initiated in 2003, with the "HapMap Project" being established to characterize the haplotype patterns in the human genome and subsequently to identify single nucleotide polymorphisms. In the past decade GWAS has been adopted to study the genomics of complex traits of several diseases ranging from leukemia, obesity, diabetes, and coronary heart diseases to neurological diseases. GWAS research has aided in the discovery of hundreds of common variants whose allele frequencies are statistically correlated

with various illnesses and traits. However, these studies did not establish any significant biological or clinical relevance in terms of prognosis and/or treatment.

Genome-wide Association Studies: From Polymorphism to Personalized Medicine is intended for those in the biotechnology, genetics, genomics, pharmaco-genomics, and molecular medicine fields. There are a few books already available covering genomic structural variants or copy number variations (Kehrer-Sawatzki, H. and Cooper, D.N. (2009) *Copy Number Variation and Disease*, Basel, Switzerland; Karger A.G., Zeggini, E. and Morris, A. (2010) *Analysis of Complex Disease Association Studies*. New York, USA, Academic Press; Feuk, L. (2012) *Genomic Structural Variants*, New Jersey, USA, Springer Press; Gondro, C., van der Werf, J. and Hayes, B. (2013) *Genome-Wide Association Studies and Genomic Prediction*, New Jersey, USA, Springer). For example, Kehrer-Sawatzki and Cooper (2009) highlighted the copy number variations, whereas Zeggini and Morris (2010) nicely covered the importance of disease association studies. The book by Feuk (2012) emphasized the protocols involved with identifying the structural variants, whereas a recent book by Gondro *et al.* (2013) on GWAS focused mainly on statistical approaches. This present book differs, in that it is the first text completely devoted to combining association studies with single nucleotide polymorphisms (SNPs), copy number variations (CNVs), haplotypes, and expressed quantitative trait loci (eQTL). Special emphasis is placed on pharmaco-genomic and population genomics studies using next-generation technology approaches. This book also focuses on the use of association studies in the context of disease biology and personalized medicine. The goal is for this book to serve as a reference for graduate students, post-doctoral researchers, and teachers and as an explanatory analysis for executives and scientists in biotechnology and pharmaceutical companies. Our hope is that this volume will serve as a prologue to the field for both newcomers and those already active in the field. We have carefully chosen the chapters, written by experts in the field from both academia and industry, and have divided the chapters into appropriate sections to support the theme expressed in the subtitle of this book: *From Polymorphism to Personalized Medicine*.

Many people have contributed to making our involvement in this project possible. We thank our teachers for their excellent teaching, guidance, and mentorship, which has helped us to bring about this educational enterprise. We are extremely grateful to all of the contributors to this book, without whose commitment this book would not have been possible. Many people have had a hand in the preparation of this book. Each chapter has been passed back and forth between the authors for criticism and revision; hence, each chapter represents a joint contribution. We thank our readers, who have made the hours spent putting together this volume worthwhile. We are indebted to the staff of Cambridge University Press, and in particular Katrina Halliday for her generosity and efficiency throughout the editing of this book; she truly understands the urgency and need of this volume. We also extend our appreciations to Megan Waddington and Victoria Parrin for their excellent cooperation during the development of this

volume. We want to thank Professor Stephen Scherer, a Canadian Geneticist, and one of the pioneers in the field of copy number variations, for his kindness in writing the Foreword to this book. We also want to thank Professor Peter Visscher, an Australian Quantitative Geneticist, and one of the pioneers in the field of genome association studies, for his thoughtfulness in writing the Foreword to this book. Last, but not least, we thank Shyamala Appasani for her understanding and support during the development of this interesting volume.

This book is the third joint project of father and son. A portion of the royalties will be contributed to the Dr. Appasani Foundation (a non-profit organization devoted to bringing social change through the education of youth in developing nations) and The MINDS Foundation (Mental Illness and Neurological Diseases), which is committed to taking a grassroots approach to providing high-quality mental healthcare in rural India.

Krishnarao Appasani
Raghu K. Appasani

Abbreviations

ABPA	approximate Bayesian polygenic analysis
ACCE	**a**nalytic validity, **c**linical validity, **c**linical utility and associated **e**thical, legal and social implications
ACTG	Adult AIDS Clinical Trials Group
AD	alcohol dependence
ADHD	attention deficit/hyperactivity disorder
ADME	absorption, distribution, metabolism, and elimination
ADR	adverse drug reaction
AFP	alphafetoprotein
AI	allelic imbalance
ALD	admixture linkage disequilibrium
ALPS	autoimmune lymphoproliferative syndrome
AMD	age-related macular degeneration
AML	acute myeloid leukemia
ANI	ancestral North Indian
ANTAC	asymptotically normal estimation with thresholding after adjusting covariates
APA	alternative polyadenylation
AR	allelic ratios
ART	assisted reproductive technologies
ARV	antiretroviral
ASD	autism spectrum disorder
ASI	ancestral South Indian
AUD	alcohol use disorder
BCH	Boston Children's Hospital
BDNF	brain-derived neurotrophic factor
BEDMR	breakpoint-enriched differentially methylated region
BER	breakpoint-enriched region
BGA	biogeographical ancestry
BGI	Beijing Genome Institute

BLUP	best linear unbiased prediction
BPD	bipolar disorder
CAD	coronary artery diseases
cART	combined antiretroviral therapy
CAST	Cohort Allelic Sums Test
CBA	cost–benefit analysis
CBS	circular binary segmentation
CBT	cognitive behavioral therapy
CCFA	Crohns and Colitis Foundation of America
CCMB	Centre for Cellular and Molecular Biology
CCR5	chemokine co-receptor
CD/CV	Common Disease/Common Variant
CDA	cytidine deaminase
CDRI	Central Drug Research Institute
CEA	cost-effectiveness analysis
CFH	complement factor H
cGGM	conditional Gaussian graphical models
CGH	comparative genomic hybridization
CGS	CpG-related SNP
CGS-C	CpG-related SNP with the allele to create CpG dinucleotides
CGS-Cp	proxy CGS-C
CGS-D	CpG-related SNP with the allele to disrupt CpG dinucleotides
CGS-Di	index CGS-D
cM	centimorgan
CMA	cost-minimization analysis
CMC	Combined Multivariate and Collapsing Method
CNP	copy number polymorphism
CNS	central nervous systems
CNV	copy number variations
COGS	cost of goods and services
CS	cleavage site
CSIR	Council of Scientific and Industrial Research
CTPS	cytidine-5′-triphosphate synthetase
CUA	cost–utility analysis
DCDA	dichorionic diamnionic
DCK	deoxycytidine kinase
DCTD	deoxycytidylate deaminase
dCTP	deoxycytidine triphosphate
DHS	DNase I hypersensitivity site
DMP	DNA methylation position
DMR	differentially methylated region
DNMT	DNA methyltransferase
DOHaD	Developmental Origin of Health and Diseases
DSM	Diagnostic and Statistical Manual of Mental Disorders
DZ	dizygotic

EEG	electroencephalography
eGFR	estimated glomerular filtration rate
EH	expression heterogeneity
EM	expectation–maximization
EMMPAT	Evolutionary Mixed Model for Pooled Associated Testing
ENCODE	Encyclopedia of DNA Elements
eQED	eQTL electrical diagrams
eQTL	expressed quantitative trait loci
ES	exome sequencing
ETV-RAM	etravirine-resistance-associated mutation
EWAS	epigenome-wide association studies
FAB	French–American–British
FDR	false discover rate
FISH	fluorescence *in situ* hybridization
FMRP	Fragile X mental retardation protein
FPKM	fragments per kilobase of exon per million fragments mapped
GABA	gamma-aminobutyric acid
GCTA	genome-wide complex trait analysis
GIANT	Genomic Investigation of ANthropometric Traits
GO	gene ontology
GST	glutathione-*S*-transferases
GWA	genome-wide association
GWAS	Genome-Wide Association Studies
HAART	highly active antiretroviral therapy
HCAEC	human coronary artery endothelial cells
HCASMC	human coronary artery smooth muscle cell
HCG	human chorionic gonadotropin
HDL	high-density lipoprotein
HDL-C	high-density lipoprotein cholesterol
HIV	human immunodeficiency virus
HWE	Hardy–Weinberg Equilibrium
IBD	identical-by-descent
IBS	identity-by-state
ICE	inter-sample correlation emended
ICF	immunodeficiency, centromere instability and facial anomalies (syndrome)
ICGC	International Cancer Genome Consortium
IGIB	Institute of Genomics and Integrative Biology
IGV	Indian Genome Variation
IICB	Indian Institute of Chemical Biology
IMTECH	Institute of Microbial Technology
InDel	Insertion/Deletions
InI	integrase inhibitor
ISC	International Schizophrenia Consortium
ITD	internal tandem duplication

ITRC	Industrial Toxicological Research Centre
KSR	kinase suppressor of ras
LAPS	Lifetime Alcohol Problems Score
LCL	lymphoblast cell line
LCR	low-copy number region
LD	linkage disequilibrium
LINE	long interspersed elements
LMM	linear mixed model
LoF	loss of function
LORS	low-rank representation and sparse regression
LR	likelihood ratio
MAF	minor allele frequencies
MALD	mapping disease genes by admixture linkage disequilibrium
MALDI/TOF	matrix-assisted laser desorption/ionization–time of flight
MAP	mitogen-activated protein
MAPK	mitogen-activated protein kinase
MBD	methyl-CpG-binding domain (protein)
MCDA	monochorionic diamnionic
MCL	gene ontology-guided Markov cluster (algorithm)
MET	motivational enhancement therapy
MFE	minimum free energy
MGS	Molecular Genetics of Schizophrenia
MHC	major histocompatibility complex
MI	myocardial infarction
miRSNP	SNPs in microRNAs target site
MLS	Michigan Longitudinal Study
MOMA	methylation oligonucleotide microarray analysis
MRD	minimal residual disease
MSA	multiple system atrophy
MSS	maternal serum screening
VP	Million Veteran Program
MZ	monozygotic
NAHR	non-allelic homologous recombination
NAPHA	National Access to Antiretroviral Programs for People who have AIDS
NCATS	National Center for Advancing Translational Science
NDK	nucleoside diphosphate kinase
NFκB	nuclear factor κB
NGS	next-generation sequencing
NHGRI	National Human Genome Research Institute
NIPT	non-invasive prenatal testing
NNIBP	NNRTI binding pocket
NNRTS	non-nucleoside reverse transcriptase inhibitor
NRTI	nucleoside (or nucleotide) reverse transcriptase inhibitor
OATP	organic anion transporting polypeptide

OMIM	Online Mendelian Inheritance in Man
OR	odds ratio
ORF	opening reading frame
$P(x>0)$	percentage of CpG sites with methylation level greater than 0
PA	protease inhibitors (PI),
PAGE	Population Architecture using Genomics and Epidemiology
PAS	polyadenylation signal
PC	principal components
PCA	principal components analysis
PCR	polymerase chain reaction
PD	pharmacodynamics
PGC	Psychiatric Genomics Consortium
PGx	pharmacogenomics
PK	pharmacokinetics
POMC	pro-opiomelanocortin
PROCARDIS	Precocious Coronary Artery Disease
QALY	Quality Adjusted Life Year
QC	quality control
QQ	quantile–quantile
QTL	quantitative trait loci
RA	rheumatoid arthritis
RefSeq	Reference Sequence
REML	restricted maximum likelihood
RISC	RNA-induced silencing complex
ROMA	representational oligonucleotide microarray analysis
RPKM	reads per kilobase per million reads mapped
RR	relative risk
RR	ribonucleotide reductase
RRBS	reduced representative bisulfite sequencing
RVM	relevance vector machine
SAGE	Study of Addiction: Genetics and Environment
SCID	Structured Clinical Interview for DSM-IV
SCZ	schizophrenia
SFS	site frequency spectrum
SINE	short interspersed elements
SMC	smooth muscle cell
SNP	single nucleotide polymorphisms
STR	short tandem repeat
SV	structural variation
SVA	surrogate variable analysis
SVM	support vector machines
T1D	type 1 diabetes
T2D	type 2 diabetes
TCGA	The Cancer Genome Atlas
TDT	transmission disequilibrium test