CONTEMPORARY CARDIOLOGY

CHRISTOPHER P. CANNON, MD
SERIES EDITOR

Cardiovascular Disease and the Elderly, edited by Gary Gerstenblith, MD, 2005
Platelet Function: Assessment, Diagnosis, and Treatment, edited by Martin Quinn, MB BCH BAO, PhD, and Desmond Fitzgerald, MD, FRCP, 2005
Diabetes and Cardiovascular Disease, Second Edition, edited by Michael T. Johnstone, MD, CM, FRCP(C), and Aristidis Veves, MD, DSc, 2005
Angiogenesis and Direct Myocardial Revascularization, edited by Roger J. Laham, MD, and Donald S. Baim, MD, 2005
Interventional Cardiology: Percutaneous Noncoronary Intervention, edited by Howard C. Herrmann, MD, 2005
Principles of Molecular Cardiology, edited by Marschall S. Runge, MD, PhD and Cam Patterson, MD, 2005
Heart Disease Diagnosis and Therapy: A Practical Approach, Second Edition, by M. Gabriel Khan, MD, FRCP(C), FRCP (LONDON), FACP, 2005
Cardiovascular Genomics: Gene Mining for Pharmacogenomics and Gene Therapy, edited by Mohan K. Raizada, PhD, Julian F. R. Paton, PhD, Michael J. Katovich, PhD, and Sergey Kasparov, MD, PhD, 2005
Surgical Management of Congestive Heart Failure, edited by James C. Fang, MD and Gregory S. Couper, MD, 2005
Cardiopulmonary Resuscitation, edited by Joseph P. Ornato, MD, FACP, FACC, FACEP and Mary Ann Peberdy, MD, FACC, 2005
CT of the Heart: Principles and Applications, edited by U. Joseph Schoepf, MD, 2005
Cardiac Transplantation: The Columbia University Medical Center/New York-Presbyterian Hospital Manual, edited by Niloo M. Edwards, MD, Jonathan M. Chen, MD, and Pamela A. Mazzeo, 2004
Heart Disease and Erectile Dysfunction, edited by Robert A. Kloner, MD, PhD, 2004
Coronary Disease in Women: Evidence-Based Diagnosis and Treatment, edited by Leslee J. Shaw, PhD and Rita F. Redberg, MD, FACC, 2004
Complementary and Alternate Cardiovascular Medicine, edited by Richard A. Stein, MD and Mehmet C. Oz, MD, 2004
Nuclear Cardiology, The Basics: How to Set Up and Maintain a Laboratory, by Frans J. Th. Wackers, MD, PhD, Wendy Bruni, BS, CNMT, and Barry L. Zaret, MD, 2004
Minimally Invasive Cardiac Surgery, Second Edition, edited by Daniel J. Goldstein, MD, and Mehmet C. Oz, MD 2004
Cardiovascular Health Care Economics, edited by William S. Weintraub, MD, 2003
Heart Failure: A Clinician’s Guide to Ambulatory Diagnosis and Treatment, edited by Mariell L. Jessup, MD and Evan Loh, MD, 2003
Peripheral Arterial Disease: Diagnosis and Treatment, edited by Jay D. Coffman, MD and Robert T. Eberhardt, MD, 2003
Cardiac Repolarization: Bridging Basic and Clinical Science, edited by Ihor Giussak, MD, PhD, Charles Antzelevitch, PhD, Stephen C. Hammill, MD, Win K. Shen, MD, and Preben Bjerregaard, MD, DMS, 2003
PRINCIPLES
OF MOLECULAR CARDIOLOGY

Edited by
Marschall S. Runge, MD, PhD
Department of Medicine
UNC School of Medicine, Chapel Hill, NC

Cam Patterson, MD
Carolina Cardiovascular Biology Center
Chapel Hill, NC

Foreword by
James T. Willerson, MD
University of Texas Medical School
Houston, TX

Humana Press
Totowa, New Jersey
**FOREWORD**

*Principles of Molecular Cardiology* provides a broad and up-to-date treatment of the molecular biology of cardiovascular diseases. This is a timely volume given the impact of cardiovascular diseases in our society and the numerous advances in understanding cardiovascular disease that have followed the molecular and genomic revolutions. This book begins with a broad historical overview of the major themes in molecular cardiology: cardiovascular genetics, the molecular biology of the cardiovascular system, genomics, gene therapy, and stem cells and progenitor cells in cardiovascular disease. These chapters provide concise information about how these areas can be used to understand and develop treatments for cardiovascular diseases.

This overview is followed by a discussion of cardiac function and dysfunction. Subjects covered include the molecular and structural events in heart development; the molecular biology of inherited myocardial diseases with a specific focus on hypertrophic cardiomyopathies and muscular dystrophies that affect the heart; the molecular regulation of inotropic function and the events that occur in cardiomyocytes that contribute to progressive systolic dysfunction; the process of inflammation within the heart with emphasis on how adhesion molecules affect this process; and common cardiac defects and the molecular basis for these defects.

The third section of this book focuses on the important topic of coronary artery disease. This section includes excellent chapters addressing the differentiation of the coronary arteries; the evolution of coronary vascular lesions with emphasis on the effects of platelets, smooth muscle cells, and inflammatory cells in this process; the molecular pathways that activate platelets and the pharmacologic actions of antiplatelet drugs; pathophysiologic events that result in myocardial infarction; the development of arterial disease after cardiac transplantation; the scientific basis for thrombolytic therapy; and the molecular basis for restenosis following percutaneous coronary interventions and recently developed new treatment modalities, including radiation therapy.

The section on cardiac arrhythmias is timely given recent progress in understanding the molecular basis and genetics of arrhythmias and sudden death. Topics covered in this section include the molecular biology of the development of the cardiac conduction system, the electrical events that cause sudden death and the development of arrhythmias, emerging new therapies for arrhythmias based on the molecular understanding, and the genetic basis for arrhythmias.

The section on vascular diseases includes state-of-the-art chapters discussing the molecular events that regulate angiogenesis and the potential for angiogenic therapy, the molecular basis for the metabolism and actions of nitric oxide, the role of inflammation in vascular disease, the pathophysiology of pulmonary hypertension and the genetics that influence its development, the molecular and genetic events that cause malformations of the vascular system, and the molecular events involved in thrombosis.

The book concludes with a section on risk factors for cardiovascular disease. These chapters succinctly summarize the molecular basis for lipid metabolism, the molecular biology of aging and its impact on the cardiovascular system, the role of oxidative species in atherosclerosis, the molecular consequences of diabetes on the cardiovascular system, and the molecular determinants of inflammation and its amelioration with proper risk factor modification.

An understanding of these topics is critically important to anyone who wishes to conduct serious research in any of these areas or teach others about them. Moreover, new therapies that develop in these areas will be based on an understanding of the molecular biology and genetics of these cardiovascular problems.

I believe *Principles of Molecular Cardiology* is an outstanding book. Anyone interested in the development, genetics, and pathologies that affect the cardiovascular system will want to have this book available for ready reference.

*James T. Willerson, MD*

*Texas Heart Institute*

*The University of Texas Health Science Center at Houston*

*Houston, TX*
In recent years, molecular studies have had a major impact on cardiovascular clinical practice and outcomes. The rationale for drug therapies for treating cardiovascular diseases such as heart failure has been based primarily on data derived from basic science investigations. The Human Genome Project has implications for biology and mankind that are unparalleled, in part because having such an abundance of information on the inner workings of humans is unprecedented. Our understanding of genetic links to cardiovascular diseases has increased dramatically, helping redefine the etiology and diagnostic criteria for numerous conditions and leading to new, individualized treatments.

*Principles of Molecular Cardiology* was undertaken to explore the latest developments in molecular cardiology research. In this text, we review the complex process of heart development, explain the molecular bases of cardiovascular diseases, describe the application of research advances in clinical treatment, and provide a historical perspective for important areas within this discipline. This book is intended for researchers, clinicians, students, and healthcare professionals who want to keep abreast of current findings in molecular cardiology research. The authors, all leading specialists, provide a unique perspective of what the future in molecular research holds for their respective fields.

Genetics research and advances in gene therapy are recurring themes throughout this text. Certain genetic mutations are clearly associated with severe cardiovascular disease, and new disease-causing mutations are being identified with increasing frequency. Some researchers estimate that there are probably only 200–300 genes that provide susceptibility for the 20 diseases that account for 80% of all deaths globally. Given the genetic and physical maps of the human genome and the technology of high-throughput nucleotide sequencing, it is conceivable that all human genes that contribute to the genetic risk of major cardiovascular diseases will be known within the next decade.

In the field of vascular biology, the number of genes that have been cloned and linked to vascular wall disease is growing exponentially. Because of their association with cardiovascular function, genes encoding the proteins endothelin-1—a potent vasoactive hormone—and the angiotensin receptor have proven to be attractive sites for pharmacologic intervention, but it is clear that the genes identified today will be the therapeutic targets of tomorrow. In addition, gain-of-function and loss-of-function mice, created through genetic manipulations, have provided enormous insights into such processes as lipid metabolism and the function of cardiac- and vascular wall-specific genes.

Although studies using mouse models have been a major tool to push the field of molecular cardiology forward, advances in human genetics have contributed significantly to the understanding of inherited cardiac diseases such as long QT syndrome and hypertrophic obstructive cardiomyopathy. Although advances have been made in understanding the pathophysiologic and genetic bases of cardiac arrhythmias, current treatment options are still inadequate, prompting a search for genetic strategies to treat these conditions.

Research in the complex area of atherosclerosis continues. Despite the great strides made in recent years, many of the processes involved in atherosclerosis remain poorly characterized. Studies of atherosclerosis in humans are limited by the complexity of the cellular and molecular mechanisms that contribute to the process and the long time course of disease development. There is also significant variability seen in pathogenetic mechanisms. In this text, the authors discuss the latest developments in understanding the pathogenesis of atherosclerosis—its manifestations (coronary artery disease, acute coronary syndromes) and its underlying mechanisms (oxidative stress and inflammation).

Platelets play a central role in the pathogenesis of atherosclerosis. Therefore, platelet inhibition has proven to be a logical therapeutic strategy for acute and chronic treatment of atherosclerosis and its clinical sequelae. The need for efficient inhibition of platelet function is even more evident in the situation of a vascular injury associated with angioplasty. Strategies for inhibiting platelet function are discussed in this text. There are many different potential ways to inhibit platelet activation, and several receptors are considered promising
therapeutic targets, including the thrombin receptor and the TXA₂ receptor.

Restenosis following percutaneous coronary interventions remains a serious problem. Because of the number of molecular targets available for targeting the cell cycle in antirestenosis therapy, gene therapy is a second clinical approach for inhibiting small muscle cell proliferation. Although results from antirestenotic gene therapy trials in humans have not been published, results from animal models are promising. A second anti-restenotic gene therapy that affects the cell cycle makes use of the overexpression of cell cycle inhibitory molecules. Experimental data support the use of gene therapy as a cell cycle inhibitor; however, the application of gene therapy to clinical medicine will depend not only on the ability of cell cycle arrest to block restenosis in clinical settings, but also on the demonstration of acceptable safety profiles.

The field of developmental biology of the cardiovascular system has also accelerated during the past decade. New developments in the study of blood vessel development and a strong clinical interest in therapeutic angiogenesis have led to greater understanding of the molecular biology of the assembly of cardiovascular structures, and many of these ideas are being translated to clinical practice to treat obstructive vascular disease.

Despite these advances and promising new discoveries, cardiovascular disease remains the leading cause of death in the United States. The aging of the population will undoubtedly be a factor in the increasing incidence of coronary artery disease, heart failure, and stroke. Of the more than 64 million Americans with one or more types of cardiovascular disease, more than 25 million are estimated to be age 65 and older (Heart Disease and Stroke Statistics—2004 Update, American Heart Association). For reasons not entirely clear, there is also an increased prevalence of obesity and type 2 diabetes—the major cardiovascular risk factors—in this country. Related complications—hypertension, hyperlipidemia, and atherosclerotic vascular disease—also have increased.

In the next decade, new tools will be applied to the study of cardiovascular disease and function. These instruments will include DNA microarrays, proteomic approaches, comparative DNA analysis, and markers of human genetic variation. The innovative use of these new and powerful tools hold promise to accelerate the pace of discovery in cardiovascular medicine.

There is an untapped potential for molecular and cellular biology to lead to substantial new discoveries in the near future. These discoveries will only be achieved with intensive and focused research. We hope this text will provide a foundation of knowledge and inspiration for investigators to continue the progress in this crucial field of research. As clinicians and scientists, the advancements in molecular cardiology over the preceding decade have inspired the editors in the laboratory and at the bedside, and we are grateful to our colleagues for moving the field so rapidly during this time.

We would like to thank the many individuals who contributed to the success of this book. We especially commend all our authors for devoting their time, energy, and scholarship to preparing these chapters—we asked for the best from our contributors, and we got it. We also thank the following editors who assisted in preparation of the text. Rebecca Bartow, PhD, was primary manuscript editor, and Jennifer King, PhD, also edited and reviewed many chapters; their contributions to this project can be appreciated whenever consistency and cogency are detected in this book. Angela Rego, BBA, coordinated manuscripts, handled correspondence between physician editors and authors, and served as adjutant general for all aspects of this project. Rebecca Teaff, MA, coordinated the editing process and reviewed manuscripts, for which the editors extend their gratitude. Carolyn Kruse, BS, DC, served as a manuscript editor. Kakky Baugh, BA, Erin Allingham, BA, Elizabeth Schramm, BA, and Kelly Scarlett assisted in manuscript review and formatting, as well as verifying references. Katie O’Brien, MA, and Angela Rego, BBA, assisted in preparation of graphics for the text. We would also like to thank Craig Adams of Humana Press for his enthusiastic support in ushering this book through the publication process.

We also extend our appreciation to our colleagues, collaborators, trainees, and laboratory members, including Nageswara Madamanchi, Yaxu Wu, and Holly McDonough, who inspired our desire to tackle this project. We dedicate this book to the memory of Edgar Haber. Dr. Haber trained many of the contributors to this project and influenced all of them. His family—his wife Carol, his sons Eben, Justin, and Graham, and his sister Ruth—shared him with us, for which we are grateful.

Cam Patterson, MD
Marschall S. Runge, MD, PhD
# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreword</td>
<td>v</td>
</tr>
<tr>
<td>Preface</td>
<td>vii</td>
</tr>
<tr>
<td>Contributors</td>
<td>xii</td>
</tr>
<tr>
<td><strong>PART I. OVERVIEW</strong></td>
<td></td>
</tr>
<tr>
<td>1 Cardiovascular Genetics</td>
<td>3</td>
</tr>
<tr>
<td>\hspace{1em} Christopher Semsarian and Christine Seidman</td>
<td></td>
</tr>
<tr>
<td>2 Molecular Biology Applications in Cardiovascular Medicine</td>
<td>11</td>
</tr>
<tr>
<td>\hspace{1em} Eugene Yang, Scott M. Wasserman, Tatsuro Ishida, Raymond Tabibiazar, and Thomas Quertermous</td>
<td></td>
</tr>
<tr>
<td>3 Genomics and Its Application to Cardiovascular Disease</td>
<td>45</td>
</tr>
<tr>
<td>\hspace{1em} Robert Roberts</td>
<td></td>
</tr>
<tr>
<td>4 Gene Therapy and Cardiovascular Diseases</td>
<td>57</td>
</tr>
<tr>
<td>\hspace{1em} Yi Chu, Neal L. Weintraub, and Donald D. Heistadt</td>
<td></td>
</tr>
<tr>
<td>5 Stem Cells and Progenitor Cells in Cardiovascular Disease</td>
<td>71</td>
</tr>
<tr>
<td>\hspace{1em} Jalees Rehman and Keith L. March</td>
<td></td>
</tr>
<tr>
<td><strong>PART II. MYOCARDIAL DISEASE</strong></td>
<td></td>
</tr>
<tr>
<td>6 Development of the Vertebrate Heart</td>
<td>83</td>
</tr>
<tr>
<td>\hspace{1em} Mabelle Ashe and David Bader</td>
<td></td>
</tr>
<tr>
<td>7 Inherited Myocardial Diseases</td>
<td>105</td>
</tr>
<tr>
<td>\hspace{1em} Elizabeth M. McNally</td>
<td></td>
</tr>
<tr>
<td>8 Receptor-Signaling Pathways in Heart Failure</td>
<td>123</td>
</tr>
<tr>
<td>\hspace{1em} Shayela Suvarna, Liza Barki-Harrington, Miwako Suzuki, Philippe Le Corvoisier, and Howard A. Rockman</td>
<td></td>
</tr>
<tr>
<td>9 Recent Insights into the Molecular Pathophysiology of Viral Myocarditis</td>
<td>145</td>
</tr>
<tr>
<td>\hspace{1em} Tony Tran, Roger D. Rossen, and Douglas L. Mann</td>
<td></td>
</tr>
<tr>
<td>10 Genetic Underpinnings of Cardiogenesis and Congenital Heart Disease</td>
<td>155</td>
</tr>
<tr>
<td>\hspace{1em} Vidu Garg and Deepak Srivastava</td>
<td></td>
</tr>
<tr>
<td><strong>PART III. CORONARY ARTERY DISEASE</strong></td>
<td></td>
</tr>
<tr>
<td>11 Coronary Artery Development</td>
<td>167</td>
</tr>
<tr>
<td>\hspace{1em} Mark W. Majesky</td>
<td></td>
</tr>
<tr>
<td>12 Atherosclerosis</td>
<td>185</td>
</tr>
<tr>
<td>\hspace{1em} George A. Stouffer</td>
<td></td>
</tr>
<tr>
<td>13 Antiplatelet Drugs</td>
<td>203</td>
</tr>
<tr>
<td>\hspace{1em} Karlheinz Peter</td>
<td></td>
</tr>
<tr>
<td>14 Myocardial Infarction</td>
<td>219</td>
</tr>
<tr>
<td>\hspace{1em} D. Douglas Miller and Steven C. Herrmann</td>
<td></td>
</tr>
<tr>
<td>Chapter</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>15</td>
<td>Transplant Arteriopathy: Pathology, Pathogenesis, and Prospects for Treatment</td>
</tr>
<tr>
<td></td>
<td>Joannis Vamvakopoulos, Einari Aavik, Daniel du Toit,</td>
</tr>
<tr>
<td></td>
<td>Pekka Häyrty, and Minnie Sarwal</td>
</tr>
<tr>
<td>16</td>
<td>Thrombolysis: Scientific Basis for Current Therapy</td>
</tr>
<tr>
<td></td>
<td>Guy L. Reed III</td>
</tr>
<tr>
<td>17</td>
<td>Coronary Restenosis</td>
</tr>
<tr>
<td></td>
<td>Julius Aitsebaomo, Martin Moser, Susan Smyth, and Cam Patterson</td>
</tr>
<tr>
<td>PART IV. ARRHYTHMIAS</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Induction and Patterning of Purkinje Fiber Network</td>
</tr>
<tr>
<td></td>
<td>Takashi Mikawa, Robert G. Gourdie, Kimiko Takebayashi-Suzuki, Nobuyuki Kanzawa,</td>
</tr>
<tr>
<td></td>
<td>David J. Pennisi, Clifton P. Poma, and Maxim Shulimovich</td>
</tr>
<tr>
<td>19</td>
<td>Mechanisms of Sudden Cardiac Death</td>
</tr>
<tr>
<td></td>
<td>Wayne E. Cascio</td>
</tr>
<tr>
<td>20</td>
<td>Gene Therapy for Cardiac Arrhythmias</td>
</tr>
<tr>
<td></td>
<td>J. Kevin Donahue and Eduardo Marbán</td>
</tr>
<tr>
<td>21</td>
<td>Genetics and Arrhythmias</td>
</tr>
<tr>
<td></td>
<td>Kui Hong and Ramon Brugada</td>
</tr>
<tr>
<td>PART V. VASCULAR DISEASES</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Therapeutic Angiogenesis for Ischemic Vascular Disease</td>
</tr>
<tr>
<td></td>
<td>Jai Pal Singh and J. Anthony Ware</td>
</tr>
<tr>
<td>23</td>
<td>Nitric Oxide in Cardiovascular Biology and Pathophysiology</td>
</tr>
<tr>
<td></td>
<td>Marshall A. Corson</td>
</tr>
<tr>
<td>24</td>
<td>Vasculitis</td>
</tr>
<tr>
<td></td>
<td>Ronald J. Falk and J. Charles Jennette</td>
</tr>
<tr>
<td>25</td>
<td>Primary Pulmonary Hypertension</td>
</tr>
<tr>
<td></td>
<td>Evangelos D. Michelakis and Stephen L. Archer</td>
</tr>
<tr>
<td>26</td>
<td>The Molecular Basis of Cerebrovascular Malformations</td>
</tr>
<tr>
<td></td>
<td>Douglas A. Marchuk</td>
</tr>
<tr>
<td>27</td>
<td>Thrombotic Vascular Disease</td>
</tr>
<tr>
<td></td>
<td>Stephan Moll and Gilbert C. White II</td>
</tr>
<tr>
<td>PART VI. RISK FACTORS</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Risk Factors</td>
</tr>
<tr>
<td></td>
<td>Ngoc-Anh Le and W. Virgil Brown</td>
</tr>
<tr>
<td>29</td>
<td>Vascular Aging: From Molecular to Clinical Cardiology</td>
</tr>
<tr>
<td></td>
<td>Samer S. Najjar and Edward G. Lakatta</td>
</tr>
<tr>
<td>30</td>
<td>Oxidative Stress</td>
</tr>
<tr>
<td></td>
<td>Nageswara R. Madamanchi and Marschall S. Runge</td>
</tr>
<tr>
<td>31</td>
<td>Diabetes Mellitus: An Important Cardiovascular Risk Factor</td>
</tr>
<tr>
<td></td>
<td>David R. Clemmons</td>
</tr>
<tr>
<td>32</td>
<td>Vascular Inflammation as a Cardiovascular Risk Factor</td>
</tr>
<tr>
<td></td>
<td>Allan R. Brasier, Adrian Recinos III, and Mohsen S. Eledrisi</td>
</tr>
<tr>
<td>Index</td>
<td></td>
</tr>
</tbody>
</table>
CONTRIBUTORS

EINARI AAVIK, MS • Transplantation Laboratory, Haartman Institute, Rational Drug Design Programme, Biomedicum Helsinki, University of Helsinki, Helsinki, Finland
JULIUS AITSEBAOMO, MD • Carolina Cardiovascular Biology Center, University of North Carolina at Chapel Hill, Chapel Hill, NC
STEPHEN L. ARCHER, MD • Department of Medicine, University of Alberta Hospital, Edmonton, AL, Canada
MABELLE ASHE • Department of Medicine, Vanderbilt University, Nashville, TN
DAVID BADER, PhD • Department of Medicine, Vanderbilt University, Nashville, TN
LIZA BARKI-HARRINGTON, PhD • Department of Medicine, Duke University Medical Center, Durham, NC
ALLAN R. BRASIER, MD • Departments of Internal Medicine and Human Biological Chemistry and Genetics, The University of Texas Medical Branch at Galveston, Galveston, TX
W. VIRGIL BROWN, MD • Department of Medicine, Emory University School of Medicine, Atlanta, GA
RAMON BRUGADA, MD • Masonic Medical Research Laboratory, Utica, NY
WAYNE E. CASCIO, MD • Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC
YI CHU, PhD • Department of Internal Medicine, University of Iowa College of Medicine, Iowa City, IA
DAVID R. CLEMMONS, MD • Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC
MARSHALL A. CORSON, MD • Department of Medicine, University of Washington School of Medicine, Seattle, WA
J. KEVIN DONAHUE, MD • Institute of Molecular Cardiobiology, Johns Hopkins University School of Medicine, Baltimore, MD
DANIEL DU TOIT • Department of Surgery, Stellenbosch University, Cape Town, South Africa
MOHSEN S. ELEDRISE, MD • Department of Internal Medicine, King Abdulaziz National Guard Medical Center, Saudi Arabia
RONALD J. FALK, MD • Departments of Medicine and Pathology and Laboratory Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC
VIDU GARG, MD • Department of Pediatrics, The University of Texas Southwestern Medical Center, Dallas, TX
ROBERT G. GOURDIE, PhD • Department of Cell Biology and Anatomy, Medical University of South Carolina, Charleston, SC
PEKKA HÄYRY, MD, PhD • Transplantation Laboratory, Haartman Institute, Rational Drug Design Programme, Biomedicum Helsinki, University of Helsinki, Helsinki, Finland and Juvantia Pharma Ltd, Turku, Finland
DONALD D. HEISTAD, MD • Department of Internal Medicine, University of Iowa College of Medicine, Iowa City, IA
STEVEN C. HERRMANN, MD, PhD • Department of Internal Medicine, St. Louis University, St. Louis, MO
KUI HONG, MD, PhD • Masonic Medical Research Laboratory, Utica, NY
TATSURO ISHIDA, MD, PhD • Department of Medicine, Stanford University School of Medicine, Stanford, CA
J. CHARLES JENNETTE, MD • Departments of Medicine and Pathology and Laboratory Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC
NOBUYUKI KANZAWA, PhD • Department of Cell Biology, Cornell University Medical College, New York, NY
EDWARD G. LAKATTA, MD • National Institute on Aging, National Institutes of Health, Baltimore, MD
NGOC-ANH LE, PhD • Department of Medicine, Emory University School of Medicine, Atlanta, GA
PHILIPPE LE CORVOISIER, MD • Department of Medicine, Duke University Medical Center, Durham, NC
NAGESWARA R. MADAMANCHI, PhD • Carolina Cardiovascular Biology Center, University of North Carolina at Chapel Hill, Chapel Hill, NC
MARK W. MAJESKY, PhD • Departments of Medicine and Genetics, Carolina Cardiovascular Biology Center, University of North Carolina at Chapel Hill, Chapel Hill, NC
Douglas L. Mann, MD • Winters Center for Heart Failure Research, Baylor College of Medicine, Houston, TX

Eduardo Marbán, MD, PhD • Institute of Molecular Cardiobiology, Johns Hopkins University School of Medicine, Baltimore, MD

Keith L. March, MD, PhD • Departments of Medicine and Cellular and Integrative Physiology, Indiana University School of Medicine, Indianapolis, IN

Douglas A. Marchuk, PhD • Department of Genetics, Duke University Medical Center, Durham, NC

Elizabeth M. McNally, MD, PhD • Departments of Medicine and Human Genetics, The University of Chicago, Chicago, IL

Evangelos D. Michelakis, MD • Department of Medicine, University of Alberta Hospital, Edmonton, AL, Canada

Takashi Mikawa, PhD • Department of Cell Biology, Cornell University Medical College, New York, NY

D. Douglas Miller, MD • Department of Internal Medicine, St. Louis University, St. Louis, MO

Stephan Moll, MD • Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC

Martin Moser, MD • Carolina Cardiovascular Biology Center, University of North Carolina at Chapel Hill, Chapel Hill, NC

Samer S. Najjar, MD • National Institute on Aging, National Institutes of Health, Baltimore, MD

Cam Patterson, MD • Carolina Cardiovascular Biology Center, University of North Carolina at Chapel Hill, Chapel Hill, NC

David J. Pennisi, PhD • Department of Cell Biology, Cornell University Medical College, New York, NY

Karlheinz Peter, MD • Department of Cardiology and Angiology, University of Freiburg, Freiburg, Germany

Clifton P. Poma, BSc • Department of Cell Biology, Cornell University Medical College, New York, NY

Thomas Quertermous, MD • Department of Medicine, Stanford University School of Medicine, Stanford, CA

Adrian Recinos III, PhD • Department of Internal Medicine, The University of Texas Medical Branch at Galveston, Galveston, TX

Guy L. Reed III, MD • Department of Genetics and Complex Diseases, Harvard School of Public Health, Boston, MA

Jalees Rehman, MD • Krannert Institute of Cardiology, Indiana University School of Medicine, Indianapolis, IN

Robert Roberts, MD • Departments of Medicine and Cell Biology, Baylor College of Medicine, Houston, TX

Howard A. Rockman, MD • Departments of Medicine, Cell Biology, and Genetics, Duke University Medical Center, Durham, NC

Roger D. Rosser, MD • Department of Immunology, Baylor College of Medicine, Houston, TX

Marshall S. Runge, MD, PhD • Carolina Cardiovascular Biology Center, and Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC

Minnie Sarwal, MD, PhD • Department of Pediatrics, Stanford University Medical Center, Stanford, CA

Christine Seidman, MD • Departments of Medicine and Genetics, Harvard Medical School, Boston, MA

Christopher Semasarian, MB, BS, PhD • Molecular Cardiology Group, Centenary Institute, Newtown, NSW, Australia

Maxim Shulimovich, BSc • Department of Cell Biology, Cornell University Medical College, New York, NY

Jai Pal Singh, PhD • Cardiovascular Research Division, Eli Lilly and Company, Indianapolis, IN

Susan Smyth, MD, PhD • Carolina Cardiovascular Biology Center, University of North Carolina at Chapel Hill, Chapel Hill, NC

Deepak Srivastava, MD • Departments of Pediatrics and Molecular Biology, The University of Texas Southwestern Medical Center, Dallas, TX

George A. Stouffer, MD • Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC

Shayela Suvarna, PhD • Department of Medicine, Duke University Medical Center, Durham, NC

Miwako Suzuki, MD, PhD • Department of Medicine, Duke University Medical Center, Durham, NC

Raymond Tabibiazar, MD • Department of Medicine, Stanford University School of Medicine, Stanford, CA

Kimiko Takebayashi-Suzuki, PhD • Department of Cell Biology, Cornell University Medical College, New York, NY

Tony Tran, MD • Department of Medicine, Baylor College of Medicine, Houston, TX

Ioannis Vamvakopoulos, MS, PhD • Transplantation Laboratory, Haartman Institute, Rational Drug Design Programme, Biomedicum Helsinki, University of Helsinki, Helsinki, Finland
J. ANTHONY WARE, MD • Cardiovascular Research
Division, Eli Lilly and Company, Indianapolis, IN
SCOTT M. WASSERMAN, MD • Department of Medicine,
Stanford University School of Medicine, Stanford,
CA
NEAL L. WEINTRAUB, MD • Department of Internal
Medicine, University of Iowa College of Medicine,
Iowa City, IA

GILBERT C. WHITE II, MD • Departments of Medicine
and Pharmacology, University of North Carolina
at Chapel Hill, Chapel Hill, NC
JAMES T. WILLERSON, MD • Texas Heart Institute,
The University of Texas Health Science Center at
Houston, Houston, TX
EUGENE YANG, MD • Department of Medicine, Stanford
University School of Medicine, Stanford, CA
I

OVERVIEW
INTRODUCTION

When the double helical structure of DNA was first proposed in 1953 by Watson and Crick in a two-page Nature article (1), no one could have predicted the tremendous impact this discovery would have in establishing the study of human genetic diseases. This discovery was an important landmark in the development of the field of cardiovascular genetics. In the last 30 years, several technological advances have fueled a surge in cardiovascular genetic research. Such advances include the understanding of biochemical components of DNA, the development of cloning techniques and DNA sequencing, the amplification of DNA by polymerase chain reaction (PCR), the identification of restriction enzymes (the molecular biologist’s “scalpel”) for handling small pieces of DNA, and the undertaking of the Human Genome Project (2). Today, cardiovascular genetics is characterized by the integration of high-technology laboratory studies and clinical medicine. Within the last decade, cardiovascular genetics has redefined the etiology and diagnostic criteria for numerous diseases and has led to the development of new, individualized treatment for cardiovascular diseases.

GENETIC BASIS OF CARDIOVASCULAR DISORDERS

A genetic basis has been identified for many cardiovascular disorders (Table 1). Hypertrophic cardiomyopathy, an autosomal dominant disorder, was the first primary cardiomyopathy identified as having a genetic basis and, therefore, has served as a paradigm for the study of genetic cardiovascular disorders. After initial genetic studies in 1989 mapped the gene for familial hypertrophic cardiomyopathy to chromosome 14q1 (3), mutations in the β-myosin heavy chain gene were identified as the cause of hypertrophic cardiomyopathy (Fig. 1A). In the last 10 years, more than 200 mutations in only 10 genes have been identified as causing hypertrophic cardiomyopathy (5,6). Because all 10 genes encode sarcomeric proteins, hypertrophic cardiomyopathy has been redefined as a “disease of the sarcomere.” Over the last 5 years, mutations in several genes have been identified as contributing to other cardiovascular diseases (Table 1), including dilated cardiomyopathy (7–9), cardiomyopathies of the right ventricle such as arrhythmogenic right ventricular dysplasia (10), and mitochondrial myopathies (11). In addition, genetic mutations have been linked to arrhythmogenic disorders such as the autosomal dominant (Romano–Ward syndrome) and recessive (Jervell and Lange-Nielsen syndrome) forms of long QT syndrome, and the Brugada syndrome (12–14). These arrhythmogenic disorders have been called “ion channelopathies,” because the mutations lie in genes encoding sodium or potassium channel proteins.

Cardiovascular genetics has had an impact on the study of congenital heart diseases and vascular disorders. For example, mutations in the transcription factor TBX5 gene cause Holt–Oram syndrome (15), whereas genetic
<table>
<thead>
<tr>
<th>Type of disorder</th>
<th>Pattern of inheritance</th>
<th>Locus</th>
<th>Gene product</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Arrhythmias</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARVD</td>
<td>Dominant</td>
<td>1q42</td>
<td>Ryanodine receptor</td>
</tr>
<tr>
<td>Brugada syndrome</td>
<td>Dominant</td>
<td>3p21-24</td>
<td>Sodium channel SCN5A</td>
</tr>
<tr>
<td>Long QT syndrome</td>
<td>Dominant</td>
<td>3p21-24</td>
<td>Sodium channel SCN5A</td>
</tr>
<tr>
<td>Naxos (ARVD + palmar plantas keratoderma)</td>
<td>Recessive</td>
<td>6p24</td>
<td>Desmplakin</td>
</tr>
<tr>
<td>Stress-induced ventricular tachycardia</td>
<td>Dominant</td>
<td>1q42</td>
<td>Ryanodine receptor</td>
</tr>
<tr>
<td><strong>Cardiomyopathies</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barth syndrome</td>
<td>X-linked</td>
<td>Xq28</td>
<td>Tafazzin</td>
</tr>
<tr>
<td>Dilated</td>
<td>Dominant</td>
<td>1q3</td>
<td>Cardiac troponin T</td>
</tr>
<tr>
<td>Dilated + conduction disease</td>
<td>Dominant</td>
<td>1p1-q21</td>
<td>Lamin A and C splice variant</td>
</tr>
<tr>
<td>Dilated + muscular dystrophy</td>
<td>Dominant</td>
<td>2q35</td>
<td>Desmin</td>
</tr>
<tr>
<td>Dilated + muscular dystrophy</td>
<td>Dominant</td>
<td>6q23</td>
<td>Unknown</td>
</tr>
<tr>
<td>Hypertrophic</td>
<td>Dominant</td>
<td>1q3</td>
<td>Cardiac troponin T</td>
</tr>
<tr>
<td>Hypertrophic + conduction disease</td>
<td>Dominant</td>
<td>7q3</td>
<td>γ2 regulatory subunit</td>
</tr>
<tr>
<td><strong>Congenital</strong></td>
<td></td>
<td></td>
<td>AMP-activated protein kinase</td>
</tr>
<tr>
<td>Alagille syndrome</td>
<td>Dominant</td>
<td>20q12</td>
<td>Jagged-1 (Notch receptor ligand)</td>
</tr>
<tr>
<td>Anomalous pulmonary venous return</td>
<td>Dominant</td>
<td>4q13-q12</td>
<td>Unknown</td>
</tr>
<tr>
<td>ASD + atrial aneurysm</td>
<td>Dominant</td>
<td>5p</td>
<td>Unknown</td>
</tr>
<tr>
<td>ASD + AV block</td>
<td>Dominant</td>
<td>5q35</td>
<td>Transcription factor NKX2.5</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Type of disorder</th>
<th>Pattern of inheritance</th>
<th>Locus</th>
<th>Gene product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carney complex (myxomas)</td>
<td>Dominant</td>
<td>17q2</td>
<td>Protein kinase A regulatory subunit 1α</td>
</tr>
<tr>
<td>Char syndrome (patent ductus arteriosus)</td>
<td>Dominant</td>
<td>2</td>
<td>Unknown</td>
</tr>
<tr>
<td>Heterotaxy + structural malformations</td>
<td>X-linked</td>
<td>6p12-p21</td>
<td>Neural crest transcription factor TFAP2B</td>
</tr>
<tr>
<td>Holt-Oram syndrome (ASD, VSD)</td>
<td>Dominant</td>
<td>12q2</td>
<td>T-box transcription factor TBX5</td>
</tr>
<tr>
<td>Velocranial facial syndrome</td>
<td>Dominant</td>
<td>22q11.21-q11.23</td>
<td>T-box transcription factor TBX1</td>
</tr>
</tbody>
</table>

**Metabolic**

| Amyloidosis (iron storage)      | Dominant               | 18q11.2-q12.1  | Transthyretin                                        |
| Familial hypercholesterolemia   | Dominant               | 19p13.2        | LDL receptor                                         |
| (↑LDL-cholesterol)              | Recessive              | 1p35           | Adaptor protein                                       |
| Familial hypoapolipoproteinemia | Dominant               | 11q23          | Apolipoprotein A-1                                   |
| (↓HDL-C)                        |                        | 9q22-q31       | ATP-binding cassette transporter-1                   |
| Hemochromatosis                 | Recessive              | 6p21.3         | HFE (HLA-H)                                          |
| Homocystinuria                  | Recessive              | 21q22          | Cystathionine beta-synthase                          |
| Hypobetalipoproteinemia (↓LDL)  | Dominant               | 2p24           | Apolipoprotein (apoB)                                |
| ↑ Lipoprotein (a)               | Dominant               | 6q26-q27       | Lp(a) lipoprotein                                    |
| Pseudoxanthoma elasticum        | Recessive              | 16p13          | ATP-binding cassette transporter-C                   |
| Sitosterolemia + hypercholesterolemia | Recessive              | 2p21           | ATP-binding cassette transporters (ABCG8, ABCG5)     |
| Tangier disease                 | Recessive              | 9q22-q31       | ATP-binding cassette transporter-1                   |

**Vascular**

| Ehlers-Danlos syndrome          | Dominant               | 2q14-q21       | Type III procollagen                                 |
| Familial aortic aneurysm        | Dominant               | 2q14-q21       | Type III procollagen                                 |
| Marfan’s syndrome               | Dominant               | 15q21          | Fibrillin                                             |
| Osler Webb Rendu syndrome       | Dominant               | 9q             | Endoglin                                              |
| Supravascular aortic stenosis   | Dominant               | 7q11.2         | Elastin                                               |

ARVD, arrhythmogenic right ventricular dysplasia; ASD, atrial septal defect; AV, atrioventricular; HDL, high-density lipoprotein; HLA-H, human leukocyte antigen-H; LDL, low-density lipoprotein.

defects in another transcription factor gene, NKX2.5, cause familial atrial septal defects with conduction disease (16). In addition, genetic defects have been identified in vascular disorders characterized by a cardiac phenotype. Marfan's syndrome, an autosomal dominant disorder characterized by multisystem clinical features including cardiac, ocular, and skeletal malformations, is caused by mutations in the gene for fibrillin, a major component of the microfibrils that function in adhesion of connective tissue structures (17). Furthermore, lipid disorders, hypertension, and other vasculopathies have been associated with genetic mutations (18–20). Over the next 10 years,
A

Fig. 1. Molecular genetics of hypertrophic cardiomyopathy. (A) Pathologic specimen showing classic anatomical features of hypertrophic cardiomyopathy caused by a missense mutation in the β myosin heavy chain gene. Note marked biventricular hypertrophy (bar = 25 mm) of the free wall and the interventricular septum and enlargement of the left atrium. Chronic atrial fibrillation may have contributed to the development of mural thrombus (arrowhead) present within each atrium. (B) DNA sequence analysis showing a missense mutation in the β myosin heavy chain gene that causes hypertrophic cardiomyopathy. Genomic DNA extracted from peripheral blood was amplified by PCR, and the β myosin heavy chain gene was sequenced. The sequence trace indicates an ambiguous nucleotide (N) because one β cardiac myosin heavy chain gene contains a mutation. The substitution of one nucleotide alters the amino acid residue encoded at position 719, substituting glutamine for arginine.

B

Fig. 1. (Continued)
Fig. 2. Genotype versus phenotype in familial hypertrophic cardiomyopathy. Autosomal dominant transmission of clinical disease (filled symbols) and a missense mutation in a β cardiac myosin heavy chain gene (+) is evident in generations I and II. Young children in generation III who carry the mutation do not exhibit cardiac hypertrophy (open symbols), a finding that is consistent with age-dependent penetrance of phenotype in hypertrophic cardiomyopathy. Squares, males; circles, females; and slash, deceased.

functional and physiological effects of a disease gene in humans (26). In the Arg403Gln model, mice develop the features of the hypertrophic cardiomyopathy in humans, including cardiac hypertrophy, myofibrillar disarray, fibrosis, diastolic dysfunction, and, in some cases, sudden death (26,31). This model has provided information on how the mutation affects sarcomere function (32) and alters Ca²⁺ handling (33), and how the phenotype can be modified by genetic background (34). Such animal models are invaluable for understanding disease pathogenesis, assessing pharmacological therapies, and identifying new molecular targets for therapy.

**IMPACT OF CARDIOVASCULAR GENETICS ON CLINICAL MEDICINE**

The effects of cardiovascular genetics on clinical medicine are numerous. Identification of disease-causing genetic defects has led to more rapid and accurate diagnoses. For example, in families in which the genetic mutation for hypertrophic cardiomyopathy is known, family members can be screened early for identification of the defective gene (Fig. 2), and the diagnosis can be made before symptoms or clinical features develop. The disease-causing mutation has been identified in many individuals in whom cardiac hypertrophy has not developed; these patients, previously classified as unaffected, are now identified as genotype positive/phenotype negative with this preclinical diagnosis. The most important goal of a preclinical diagnosis in the absence of cardiac hypertrophy is the prevention of cardiac events. Although preclinical diagnosis may have deleterious psychosocial implications, identifying these patients allows for the early initiation of treatment and the possible prevention of clinical features. In addition, genetic diagnosis has led to the recognition that certain gene mutations cause late-onset disease. For example in hypertrophic cardiomyopathy, defects in the gene encoding myosin-binding protein C usually produce hypertrophy in the fourth and fifth decades of life, whereas mutations in the β-myosin heavy chain gene cause hypertrophy by the age of 20 years in more than 90% of individuals with the mutation (4,5,23,35). Thus, the absence of hypertrophy does not exclude an individual in a family with known hypertrophic cardiomyopathy from carrying the genetic defect.

A second, often overlooked, impact of molecular diagnosis in genetic cardiovascular disorders is the "negative result." The identification of the mutation in one family member allows for the screening of any member of the family. Thus, testing children can preempt longitudinal clinical evaluations. If a child does not inherit the mutation, annual clinical screening is not necessary, the child can participate in school sports and extracurricular activities, and parents find relief from worry.

Making a genetic diagnosis in a medical disease has numerous advantages. A genetic diagnosis can help resolve ambiguous diagnoses, such as those in patients with a borderline or modest increase in left ventricular wall thickness that can occur in some trained athletes with hypertrophy. Genetic testing can aid in the diagnosis of patients with systemic hypertension and suspected hypertrophic cardiomyopathy and patients with hypertrophy in less common sites such as apical hypertrophic cardiomyopathy, where diagnosis can be difficult. Furthermore, antenatal diagnoses can be made in families in which a mutation is known. Although associated with many ethical issues,
prenatal diagnosis may be warranted in a family with a clearly documented "malignant" phenotype.

Limitations in the availability of laboratory equipment and the labor-intensive nature of genetic techniques preclude the routine use of genetic diagnosis in patients with suspected cardiovascular diseases. Patients with a family history of disease are the most likely ones to undergo genetic testing. Alternative genetic techniques, such as linkage analysis, can be used to identify which gene is involved; therefore, an accurate family history, the cornerstone of management of patients with genetic disease, is of utmost importance in all inherited medical disorders. Of particular importance in obtaining an accurate family history, the ages at death and the causes of death may give an estimation of the risk of death within a family. Patients with a family history of a genetic disorder are at much greater risk of disease or death than the general population, and clinical screening of such patients is justified. However, within the next decade, the development of molecular technology such as DNA chips or mass spectrometry will allow researchers to assess thousands of mutations within hours, thereby enabling more widespread screening of familial and sporadic cases of genetic cardiovascular disorders.

Molecular studies have contributed significantly to clinical medicine by identifying potential sites of intervention and by allowing targeting of treatment. Certain genetic mutations (malignant mutations) are clearly associated with severe cardiovascular disease (4,23,35). Identification of malignant mutations, which put patients at high risk of sudden cardiac death, allows clinicians to consider preventive measures, such as implantation of a cardioverter-defibrillator to prevent sudden death. Future interventions in cardiovascular genetic disorders may involve correcting the underlying genetic defects. Furthermore, with an increased understanding of genetic mechanisms, target therapy may be developed to mitigate the genetic defect or to correct the molecular abnormality, thereby curing the disease.

FUTURE STUDIES

During the last 10 years, our knowledge of genetic cardiovascular diseases has increased remarkably. With the recent completion of the Human Genome Project, the next decade should bring an unprecedented escalation in our identification and understanding of genetic cardiovascular diseases. Furthermore, new disease-causing genes are being identified with increasing frequency. The challenge that lies ahead is determining what these genes do, how they function, how they interact with other genes, how they activate signal pathways to cause disease, and how environmental factors affect the development of disease. More investigation is required to maximize the potential of molecular genetics for improving clinical medicine. We must analyze the effect of genotype on disease expression and therapeutic response to understand fully how genetic studies can be used to enhance health care in the 21st century. Answering such questions will require collaborative efforts between clinicians and basic scientists. An improved understanding of disease pathogenesis and the development of individualized therapeutic regimens based on a patient's genetic profile will have a substantial impact on the diagnosis and treatment of cardiovascular diseases. Proper integration of genetic technologies and clinical medicine will eventually reduce human disease and suffering and prolong human life.

REFERENCES


Molecular Biology Applications in Cardiovascular Medicine

Eugene Yang, MD, Scott M. Wasserman, MD, Tatsuro Ishida, MD, PhD, Raymond Tabibiazar, MD, and Thomas Quertermous, MD

INTRODUCTION

Basic science research has made great contributions to the field of cardiovascular medicine. Scientific studies have had a major impact on clinical practices and outcomes. For example, the principles of cardiac contractile function and unique aspects of hemodynamic loading on the ventricles were defined in animal studies. These findings translated directly into pressure monitoring devices used for patients in the acute care setting. The rationale for drug therapies for treating cardiovascular diseases was based primarily on data derived from basic science investigations. For example, the treatment of heart failure and cardiac arrhythmias evolved from elegant pharmacologic and physiologic studies. A clear path has emerged from the basic science laboratory to the bedside.

During the past decade, scientists have supported the application of cellular and molecular biology to the study of cardiovascular disease and function. Endothelin-1, a potent vasoactive hormone, and angiotensin receptor were among the most heralded new genes to be cloned because of their association with cardiovascular function. The field of vascular biology has expanded rapidly; many endothelial cell genes have been cloned and linked to vascular wall disease. Gain-of-function and loss-of-function mice, created through genetic manipulations, have provided great insight into lipid metabolism and the function of cardiac- and vascular wall–specific genes. In addition, the field of developmental biology of the cardiovascular system has developed during this decade. Increased study of blood vessel development and a strong clinical interest in therapeutic angiogenesis led to great advances in the understanding of the molecular biology of the assembly of cardiovascular structures. Finally, human genetics studies have contributed significantly to the understanding of inherited cardiac diseases such as long QT syndrome and hypertrophic obstructive cardiomyopathy.

Unlike earlier scientific discoveries that successfully translated to clinical practice, the application of recent studies in cellular and molecular cardiovascular science has been slow. Although the identification of genes and their encoded proteins has generated optimism about
finding new therapeutic targets, the path from target to drug development is a complex, arduous process; therefore, success has been limited in this era of cloning and molecular biology with the exception of plasminogen activator molecules. Furthermore, gene therapy has been viewed as a way to apply directly the findings of basic genetic research to patient treatment, but again this therapeutic approach has been slow to materialize. During this decade, therefore, the major advances in the treatment of cardiovascular disease have evolved not from basic science applications, but from the development of new technologies such as the angioplasty catheter and the coronary artery stent.

In the next decade, new tools will be applied to the study of cardiovascular disease and function. These instruments will include DNA microarrays, protein chips, and other proteomics methodologies, vast amounts DNA sequences from humans and other species, and markers of human genetic variation. The innovative use of these new and powerful tools may accelerate the pace of discovery. For example, DNA microarrays can provide quantitative information about differences in gene expression between two conditions for 20,000 genes simultaneously and can relate these changes to virtually all other genes in the transcriptome. Genes that have coordinate regulation across many manipulations would be expected to have similar functions or to be involved in the same, or related, signaling pathways.

The greatest promise of DNA microarrays and other new tools is that cardiovascular researchers will be able to evaluate the behavior of every gene involved in a specific disease process or biological pathway. We can now identify all genes that are differentially regulated in the blood vessel wall of patients with vascular disease by comparing their gene expression profile with those generated from the normal vessel wall of individuals with known risk factors (such as diabetes) and those without risk factors. Human genetic epidemiology studies can be used to evaluate the association of disease and the vascular wall genes whose expression is temporally linked to vascular disease and risk factor status. Identification of the pathways that underlie cardiovascular disease will accelerate the development of new therapeutic strategies.

We believe that the great, untapped potential of molecular and cellular biology and the emergence of new genomics and human genetics initiatives will lead to substantial new discoveries in cardiovascular disease. However, these discoveries will not be achieved without intensively focused research. In this chapter, we provide a brief overview of basic molecular biology techniques and applications useful to a cardiovascular researcher.

OVERVIEW OF RECOMBINANT DNA TECHNOLOGY

The study of the human genome, which encompasses approximately 3 billion nucleotides, is a daunting task that has been made possible by recombinant DNA technology. In its basic form, recombinant DNA technology is the process of combining DNA from two or more sources. This technology is based on several common techniques of molecular biology, including the use of restriction endonucleases to cleave DNA into manageable segments, sequence analysis of purified DNA to confirm the identity of a particular fragment, DNA cloning strategies to produce large amounts of identical DNA sequence(s), and hybridization methodologies to identify particular nucleic acid sequences. Improvements in reagents and techniques have allowed for the recent, large-scale sequencing of the human genome.

In general terms, cloning is the process of isolating and amplifying a particular DNA sequence. A clone, which is defined as a large number of exact copies of a unique DNA sequence, can exist as a double-stranded DNA fragment in solution or in a vector (e.g., plasmid, bacteriophage, or plasmid–bacteriophage hybrid). Polymerase chain reaction (PCR)-based or cell-based techniques can be used for cloning. In PCR-based cloning, a target sequence is exponentially amplified by adding oligonucleotide primers designed for the sequence of interest and a DNA template to a PCR reaction mixture. This amplicon can then be used for various purposes, such as creating a probe for in situ hybridization or for ligating into a vector for cell-based cloning. Although fast, this PCR-based technique has several problems associated with in vitro enzymatic reaction systems, including failed PCR reactions, decreased efficiency of cloning lengthy sequences (>2 kb DNA segments), limited production of and random mutations in the amplicon, a need for a priori knowledge of sequence information for primer design, and high costs. Innovations in the field, such as the use of enzymes such as Pfu, with its proofreading feature, have resolved some of these issues. Cell-based cloning, however, remains the most versatile and widely used method for cloning.

Cell-based cloning involves inserting a foreign DNA sequence into a vector and introducing this vector into a host cell that reproduces the foreign DNA in large quantities. The genetic material of interest can be obtained directly from cells or tissues by a PCR-based approach.
that uses oligonucleotide primers to generate an ampli-
con. In addition, DNA can be obtained from a comple-
mentary DNA (cDNA) library, which is a collection of
distinct DNA sequences generated by converting cellular
mRNA into cDNA, or by fragmentation of genomic
DNA. The isolated DNA sequence is packaged into a
vector, which is then introduced into host cells (i.e., bac-
teria or yeast) in a process known as transformation.
Once inside the host, the vector does not incorporate into
the host genome and therefore reproducibly itself inde-
pendent of the host cell. This process generates large
amounts of the unique DNA sequence. After the host cells
have been transformed, the clones are plated onto agar at
a density that permits the isolation of individual clones.
These clones multiply and become colonies (for plasmid
cloning vectors) or plaques (phage cloning vectors).
Each colony or plaque represents one clone comprising a
group of host cells with a unique, genetically identical
recombinant DNA sequence. The colony or plaque with
the DNA element of interest is picked through a screen-
ing process and further expanded in liquid media. The
target DNA is isolated by using standard biochemical
protocols for further analyses such as sequencing or
screening a library.

Cloned DNA fragments isolated from the vector
sequence can be used for several purposes, including
Northern blotting and in situ hybridization studies to
evaluate mRNA expression levels of a specific gene.
When the cloned DNA fragment encodes a cDNA, pro-
tein expression can be used for a more detailed analysis.
In expression cloning, the expressed protein rather than
the DNA sequence can be identified after transfection of
the target sequence into a cell or tissue. Expression and
purification of proteins derived from cloned cDNAs can
be used in protein function studies, such as receptor–ligand
binding assays and other functional assays (e.g., ion flux
and phosphorylation). Finally, expressed nucleic acid
sequences can be used for in situ hybridization or for
functional analysis with antisense strategies.

**GENETIC TOOLS**

**Restriction Enzymes**

Restriction endonucleases are bacterial enzymes that
protect the host bacteria by degrading foreign double-
stranded DNA. The endonucleases act as molecular sci-
sors by cutting double-stranded DNA at unique nucleotid-
ne recognition sequences. Bacteria can protect these particu-
lar sites from enzymatic degradation by methylating
certain nucleotide residues (i.e., adenine and cytosine).

Because of their ability to recognize and cut specific
sequence motifs, restriction endonucleases can be used by
molecular biologists to cut large pieces of DNA into more
manageable fragments. These smaller fragments can be
cloned, sequenced, and even used to decipher the primary
structure of the larger sequence. Several hundred enzymes
have been isolated, characterized, and named based on the
bacteria from which they were isolated. For example,
EcoR1 originates from Escheria coli, Smal from Serratia
marcescens, and PstI from Providencia stuartii (1).

Each restriction enzyme recognizes and cuts a particular
4–8 basepair sequence. Enzymes that recognize the same
sequence are called isoschizimers. Based on the probability
of finding a 4–8 nucleotide recognition site composed of
any combination of bases, an 8-base cutter, such as PucI,
should cut human genomic DNA every $1 \times 10^6$ basepairs,
whereas a 4-base cutter, such as HaeII, should cut every
250 nucleotides (2). This calculation assumes that each
nucleotide is equally represented in the genome, which is
not completely true. Furthermore, endonuclease restriction
sites are not evenly and randomly placed in the genome.
Enzymes that have CpG in their recognition sequence more
frequently in regions of the genome that are rich in C and
G nucleotides. Recognition sites are often palindromic, and
most have a twofold axis of symmetry with the nucleic acid
sequence for each strand being the same when read in the 5'
to 3' direction.

After binding to its recognition site, the restriction
endonuclease hydrolyzes a phosphodiester bond at the
same point in each 5' to 3' sequence. This cleavage can
occur symmetrically or asymmetrically. Restriction frag-
ments produced by restriction endonucleases that cut sym-
metrically have blunt-ends, whereas enzymes that cut
asymmetrically yield fragments with 5' or 3' overhangs.
These asymmetric overhangs are often called cohesive ter-
mini, or sticky ends, because the overhanging ends are
complementary and can associate with themselves by
hydrogen bonds to yield a circular molecule or with each
other to form linear or circular concatamers. The fragments
with either blunt or sticky ends can be ligated into a plasmid
cloning vector that has been linearized (i.e., cut with a
restriction enzyme). These intramolecular and intermolecular
associations of restriction fragment ends depend on the
DNA concentration and whether the fragment and cloning
vector ends are blunt or have compatible overhangs.

High DNA concentrations favor concatamerization,
whereas low DNA concentrations favor intramolecular
cyclization (3). When ligating or joining by enzymatic
reaction the DNA fragment of interest with a linearized
vector, the concentrations of the DNA insert and vector
are optimized to ensure ligation of one DNA insert copy per vector. Several techniques can be used to improve the efficiency of cloning the fragment into a vector, such as the use of two restriction enzymes that do not produce complementary ends or the use of vector dephosphorylation.

Restriction enzyme digestion of DNA, either genomic or cDNA, produces fragments of different lengths. A digest reaction can be complete or partial. In a complete restriction digest, all of the recognition sites for that restriction endonuclease are cleaved. Partial digests, which occur when the amount of enzyme is limited or when the time for digestion is decreased, result in a random cleavage of only a fraction of the available restriction sites. Complete digestion of the target DNA is usually desired. Restriction fragments can be separated by gel electrophoresis, and their sizes can be calculated by comparing the fragments to DNA standards of known molecular weights. This DNA fingerprint can be compared to fingerprints generated from digests with other restriction enzymes to yield a restriction map. These restriction maps, which provide useful information for other molecular biological tools, are linear or circular maps that document the relative positions of various restriction sites for a specific segment of DNA.

**Cloning Vectors**

A cloning vector is a vehicle for the delivery of foreign DNA into a host cell and the replication of that DNA independent of the host cell cycle. Foreign DNA is first inserted into the vector in vitro. This hybrid molecule is transferred to a host cell where it uses the host’s cellular machinery to generate multiple copies of itself without incorporating into the host’s genome. The most commonly used cloning vectors are plasmids and bacteriophages. Both of these naturally occurring species have been genetically engineered to replicate in a foreign host, to transform the host efficiently, and to allow the recovery of vector with its target DNA. Several new hybrid vectors, such as cosmids, yeast artificial chromosomes (YACs), and bacterial artificial chromosomes (BACs), have been constructed to clone larger pieces of DNA. In addition, cloning vectors can be engineered to produce nucleic acid sequences or proteins (4). These vectors, called expression vectors, have special sequences that direct the transcription of nucleic acids and the translation of amino acids independent of the host cell cycle. The generated nucleic acid and protein sequences can be used in several assays such as *in situ* hybridization, and the amino acid sequences can be assessed for function or antibody binding.

**Plasmids**

Plasmids are naturally occurring, circular molecules of extra-chromosomal, double-stranded DNA that can self-replicate. Plasmids carry genes that can confer unique properties on their host, such as drug resistance, sexual fertility, and the ability to synthesize a rare amino acid. Found in the cytoplasm of many prokaryotes and eukaryotes, such as bacteria, yeast, and mammalian cells, plasmids can be transferred to neighboring cells through bacterial conjugation or to daughter cells through host cell division. Because of their innate ability to carry genes and their compatibility with various cell types, these natural structures have been genetically engineered for use as cloning vectors.

Plasmid vectors have various unique properties based on their intended use; however, they all have three basic characteristics. First, plasmid vectors must have the means to self-replicate; thus, plasmids contain an origin of replication site and other genes required for replication. Second, plasmid vectors need a polynlinker, a 50-150 nucleotide sequence comprising multiple, distinct recognition sites for restriction endonucleases, that allows the simple insertion and removal of foreign DNA sequence in and out of the vector. The third feature is the need for selectable markers that will differentiate host cells with no vector, empty vector, or vector with foreign DNA. These markers usually bestow properties on the transformed cell that it normally does not possess. For example, bacteria transformed with a plasmid that encodes an antibiotic resistance gene will no longer die in the presence of that antibiotic.

Several genes have been used as selectable markers, including genes that encode resistance to ampicillin (β-lactamase), chloramphenicol (chloramphenicol acetyltransferase), kanamycin (kanamycin phosphotransferase), and tetracycline (3). Bacteria transformed with plasmids encoding these genes will survive when grown on agar implanted with the appropriate antibiotic.

In addition, markers can be used to determine whether a plasmid vector contains foreign DNA sequence. To achieve this goal, β-galactosidase can be inserted into the polynlinker site. When foreign DNA is cloned into the polynlinker, the marker gene is interrupted, resulting in insertional inactivation. Thus, cells transformed with empty vector turn blue when grown on X-gal impregnated agar, whereas those cells transformed with plasmid containing the insert will be colorless. Furthermore, this technique of insertional inactivation can be used with antibiotic resistance genes. In this case, two antibiotic resistance genes are used; one is placed within the polynlinker and the other is outside of the polynlinker. Transformed bacteria are grown in duplicate on agar plates containing both antibiotics. Colonies that are sensitive to the antibiotic resistance gene in the
polylinker, but resistant to the antibiotic gene outside of the polylinker, contain the plasmid with the foreign DNA. Colonies that are resistant to both antibiotics contain empty vector.

Foreign DNA can be ligated into plasmid vectors with DNA ligase in vitro. Plasmid vectors are ideal for the cloning of cDNAs because most mRNA transcripts are smaller than 5 kb. Bacteria are transformed with the plasmid and grown on an agar plate impregnated with the appropriate selectable marker(s). Colonies with the insert are expanded in large volumes of media. Several plasmid purification kits are available for isolating pure plasmid DNA from the bacterial host. Because of their ease of use and widespread availability, plasmids are a fundamental tool in molecular cloning. The major limitations with plasmids are that the foreign DNA insert must be smaller than 5–10 kb and the inefficiency of bacterial transformation (1). Plasmid vectors are commercially available from several vendors. These products range from the straightforward plasmids used for cloning restriction fragments to more advanced vectors designed to produce functional proteins.

**BACTERIOPHAGE**

Bacteriophage λ, a bacterial virus engineered for cloning, can transform bacteria more efficiently than plasmids and can accept inserts up to 23 kb in size. The wild-type bacteriophage λ virion comprises a protein coat that holds nearly 50 kb of linear double-stranded DNA. A temperate virus, bacteriophage λ can exist in a lytic or lysogenic growth state once it has infected a host. In the lytic cycle, the virus makes multiple copies of its genome and coat proteins. The viral genome is then packaged into new phage particles that are released when the host cell lyses. The neighboring cells are infected by the phage and the cycle of replication and release continues. In the lysogenic cycle, the viral genome is integrated into the host chromosome and replicates along with the host chromosome.

The bacteriophage and its lytic growth cycle, when used in recombinant DNA technology, is an effective and efficient cloning vector. The middle portion of the bacteriophage λ genome encodes proteins that are not vital for lytic growth and can be excised so that foreign DNA can be ligated between the two “arms” of essential genetic sequence. This recombinant DNA, which must be 37–52 kb, is then packaged in vitro by coat proteins into infectious phage particles (1). The infectious virions are grown with bacteria. Bacteria infected by the bacteriophage λ undergo lysis, which leaves a “hole” known as a plaque on the bacterial lawn. Each plaque represents a single clone with a unique insert. Clones are picked and expanded in a broth medium. Then, the phage are purified and DNA is isolated. Because they are easier to handle than cosmids, YACs, and BACs, bacteriophage are ideal for cloning midsized 10–20 kb segments (5).

**COSMIDS**

Cosmids are hybrids of plasmids and bacteriophage λ designed to hold 30–44 kb DNA fragments that efficiently transform bacteria (1). Cosmids contain a plasmid backbone composed of selection markers, an origin of replication, a polylinker, and a cos site from the bacteriophage λ. Cos sites are cohesive termini found at the 5' ends of linear phage DNA that are ligated by the host cell, which produces a circular molecule capable of replication. For cloning, the cosmid vector is linearized and mixed with the DNA fragments to be cloned. DNA ligase joins cut vector and insert fragments into concatemeric molecules. These molecules are mixed with packaging extract containing proteins necessary to package naked phage DNA into phage heads. The infectious phage inject their DNA, which contains the insert into the host bacteria. The linear cosmid DNA recircularizes in the host and replicates as a plasmid. Bacterial colonies are chosen based on their selection markers. Although these hybrid molecules clone larger fragments than plasmids or phage, and transform bacteria as efficiently as phage, cosmids tend to rearrange and/or delete DNA insert segments.

**YEAST ARTIFICIAL CHROMOSOMES**

Yeast artificial chromosomes (YACs) are recombinant vectors generated to clone large fragments of genomic DNA by taking advantage of chromosomal replication during cell division (2,6,7). Recombinant DNA technology is used to place four sequences required for chromosomal replication in a plasmid backbone: an autonomous replication sequence, a centromere sequence (central region of chromosome), and two telomere sequences (chromosome ends). This plasmid backbone contains cloning sites and selectable markers. A YAC plasmid vector is linearized and then cut into two fragments. Foreign DNA 0.2–2.0 million bases (Mb) in length is ligated between these two arms, creating an artificial chromosome with two telomeres capping the chromosome ends and a central centromere. The cell walls of the yeast Saccharomyces cerevisiae are removed (thereby producing spheroplasts) and embedded in agar for support and stability. The yeast are transformed with the artificial chromosomes, which have a low transformation efficiency. More starting DNA is required to ensure complete representation of the genome when this strategy is used to generate a recombinant genomic library. The yeast grow in their selective environment and regenerate their
walls. Selection markers allow only YAC-containing yeast to be propagated. The foreign DNA is replicated during cell division, resulting in a single copy per cell. The low yields, creation of inserts composed of two or more noncontiguous genomic fragments, and a complicated DNA isolation process are drawbacks to this technique. Furthermore, generation of recombinant YACs is time consuming and technically difficult. Only a few universities and national laboratories support YAC cloning. Nonetheless, YACs are important tools for generating physical genome maps over multiple megabases and have permitted the cloning of regions with repetitive sequences common to eukaryotes.

**Bacteriophage P1 Vectors and Artificial Chromosomes**

Cloning of eukaryotic genomes has been complicated by the large size of the genome and by the presence of structurally unstable repetitive sequence elements. YACs have been engineered to alleviate these problems but are technically challenging; therefore, several alternative vectors have been created. Bacteriophage P1 is a plasmid–bacteriophage hybrid vector engineered to accommodate 70–100 kb of foreign DNA (8,9). Target DNA sandwiched between two plasmid arms is packaged in P1 phage in vitro. These infectious virions inject their DNA into a host where it circularizes into a plasmid and replicates. P1 artificial chromosomes are a combination of the *Escherichia coli* fertility factor (F-factor) plasmid and bacteriophage P1 (10). BACs based on the F factor plasmid were created as an alternative to YAC cloning (11). The F factor plasmid is a low copy number plasmid (one or two copies per cell), which reduces the potential for genetic recombination events, and the F factor plasmid can accept DNA fragments up to 300 kb. *E. coli* are transformed by electroporation, which is 10–100 times more efficient than yeast spheroplast transformation. Therefore, less starting material is required to create a comprehensive genomic library. Conventional colony lift and hybridization techniques are used in BAC screening. DNA isolation is easier and chimeric DNA inserts are formed less frequently with BACs than with YACs.

**Expression Vectors and Systems**

Expression vectors are cloning vectors designed to express or produce substantial quantities of a gene or its protein. For example, an mRNA product of an expression vector can be used as a probe for in situ hybridization or as an antisense RNA for functional–therapeutic assessment. The expressed gene products or recombinant proteins can be used as molecular reagents such as restriction endonucleases, as therapeutic proteins such as erythropoietin, or as reagents for structural analyses, functional assays, antibody production, or drug screening. These expression vectors have many features of cloning vectors, such as the polylinker cloning sites and selectable markers. In addition, expression vectors have specific sequences that target the cloned cDNA for transcription and translation in the host cell. Some vectors have molecular switches called inducible promoters that can turn the transcription of a gene on and off through the addition of an inert reagent. The inducible promoter protects the host from potentially toxic effects of the recombinant protein or its production. The most commonly used expression vectors are plasmids with bacteriophage sequence elements and viral vectors. Host systems range from bacteria to mammalian cell lines. The vector choice is usually based on the application of the expressed gene or recombinant protein.

**E. coli Recombinant Protein Expression System**

*E. coli* transformed with plasmid/bacteriophage-based expression vectors generate recombinant proteins for antibody screening, functional studies, and structural determination. This system is best suited for generating small, intracellular proteins because large polypeptides cannot be properly folded, and post-translational modifications are minimal or nonexistent in the *E. coli* system. Eukaryotic proteins are usually not very stable in bacteria. Producing a fusion protein by inserting the foreign cDNA downstream of a sequence that encodes a highly expressed host cell gene can increase protein stability and expression. For example, a cDNA of interest can be inserted into the *E. coli* lacZ gene, which results in a fusion protein consisting of the cloned gene product at the carboxy-terminus and β-galactosidase at the amino-terminus. Amino-terminal tags, glutathione-S-transferase and polyhistidine, can be added to the cloned gene product to facilitate protein purification (2). To ensure that expression of the target protein is not harmful to the host, expression vectors often have an inducible promoter that turns on the production of the cloned gene product in the presence of a reagent, such as isopropyl-β-d-thiogalactopyranoside. The *E. coli* recombinant protein system is inexpensive and simple, and most molecular biologists are familiar with the techniques and reagents. However, recombinant protein stability and solubility are problems in this system. Despite efforts to create a “bacterial” fusion gene product, these proteins still degrade easily. The use of *E. coli* strains deficient in proteases has met with limited success. High levels of protein expression can result in the formation of inclusion bodies, which are dense aggregates of insoluble recombinant protein (3). Although they may improve protein purification, the aggregates can cause the protein to be incorrectly folded and inactive. The
level of expression, temperature, and the bacterial strain can be varied to mitigate this problem (3,12).

**Baculovirus-Insect Recombinant Protein Expression System**

Large quantities of recombinant protein can be produced with the baculovirus expression system in insect cells. Baculoviruses are large, enveloped arthropod viruses containing double-stranded DNA. Like other bacteriophage vectors, the sizable 130 kb baculovirus genome contains significant amounts of genetic material that can be discarded without hindering its ability to replicate and infect a host. Therefore, this viral vector system can support large segments of foreign DNA.

The baculovirus vector is prepared through homologous recombination. Target DNA is ligated into a transfer plasmid flanked by the polyhedrin promoter and viral-specific sequences. The polyhedrin protein is a protective viral coat protein that is produced in substantial quantities but is not critical for viral replication or infection. Insect cells lacking polyhedrin cannot form occlusion bodies and thus have unique plaque morphologies. A small transfer plasmid with the foreign DNA fragment and wild-type baculovirus DNA must be applied together and transform *Spodoptera frugiperda* (Sf9)-cultured insect cells. The cloned gene is incorporated into the baculovirus DNA by replacing the polyhedrin gene inside the cell through an inefficient process called homologous recombination. Proper recombination is confirmed by PCR or nucleic acid hybridization. The viral mixture composed of wild-type and recombinant baculovirus is plated onto Sf9 cells, and plaques with recombinant virus are chosen and expanded for protein purification. A substantial quantity of soluble recombinant protein is generated with relative ease. However, this process requires a high level of expertise and is time consuming. Insect cells can be infected with more than one recombinant virus, which permits the expression of more than one protein or protein subunit per cell. Because the baculovirus system is eukaryotic, the recombinant protein often assumes the proper cellular location and can undergo some post-translational modification. However, excessive recombinant protein production can distort the cellular compartmentalization. This recombinant protein strategy is confined to insect cells because baculovirus does not infect vertebrates, and its promoters do not function in mammalian cells.

**Recombinant Protein Expression in Mammalian Cells**

Mammalian cell expression of recombinant proteins is essential for the proper synthesis, processing, and folding of complex polypeptides. Recombinant proteins have successfully been expressed in mammalian cells with multiple techniques, including transient transfection, stable transfection, viral transduction of cultured cells, transgenic animals, and cell lines derived from transgenic animals. In transient transfection, plasmid-based vectors drive the expression of recombinant proteins for a brief period of time (i.e., days to weeks). Several methods, such as DEAE-dextran, electroporation, and liposome formulations, can be used to transfect the host cell with the plasmid vector carrying the foreign DNA. Although it varies with the technique and cell type, the transfection efficiency is low, with approximately 5–50% of the cell population expressing the recombinant protein. This technique is useful for getting quick information about subcellular localization or cellular function, but it is not useful for producing protein for biochemical characterization.

Strategies have been developed to increase the production of recombinant protein by transiently transfecting eukaryotic cells. These methods use COS cells, a simian kidney cell line with a stably integrated, replication origin-defective SV40 genome (13). COS cells transfected with a plasmid vector containing foreign DNA and a SV40 origin of replication produce high levels of target protein expression. This target protein is processed and appropriately secreted or targeted to its subcellular location. Thus, the COS cell expression system is a good transient transfection strategy to determine receptor–ligand interactions and to reliably produce sufficient amounts of recombinant protein in mammalian cells.

The use of viruses is another method for expressing recombinant protein in mammalian cells both in vitro and in vivo. Vaccinia viruses, adenoviruses, and retroviruses have been used to insert foreign DNA into mammalian cells. The gene of interest is usually expressed under the control of viral expression elements. After infecting mammalian cells, the virus uses the host cell translational machinery to produce large amounts of the protein of interest, which can be purified for further analyses. The in vivo application of viral-mediated gene delivery, known as gene therapy, is beyond the scope of this chapter, but will be discussed elsewhere in this book.

Stable integration of a foreign gene into a host cell chromosome through transfection or generation of transgenic animals results in the reliable production of large amounts of recombinant protein. An expression plasmid can be integrated into the cellular genome of cultured cells by selection for stable expression of a drug resistance gene such as the neomycin resistance gene. With these methods, identical stable cell lines can be developed and selected for high level expression of recombinant protein.
The use of engineered Chinese hamster ovary cells in combination with plasmid vectors carrying specific drug resistance genes allows for the production of cell lines with high levels of protein expression. Transgenic animals are produced to accept an expression construct in the germ line through manipulation of fertilized eggs or embryonic cells in culture. In one clever application of this technique, transcriptional regulatory sequences in the expression construct direct expression of the protein in milk-producing cells so that the recombinant protein can be extracted from milk. Alternatively, stable cell lines can be adapted from the transgenic animal for the in vitro production of recombinant protein.

RECOMBINANT LIBRARIES

A DNA library is a compendium of DNA clones isolated from a particular type of cell, tissue, or organism. The goal of developing such a library is to obtain the genetic information that determines a particular species of animal or plant, to examine the expressed genetic information that characterizes a specific cell or tissue, and to observe how a cell or tissue responds to certain environmental stimuli. Cloning and sequence characterization of numerous genomes together with identifying expressed sequence tags (EST) from large numbers of cells and tissues has provided many physical clones that can be obtained directly. Access to these clones has decreased the need for generating and screening libraries. All genes will eventually exist as a catalogued collection of well-characterized clones that can be rapidly obtained for a small fee.

Genomic libraries are constructed from the high-molecular-weight DNA that is contained in the nuclei of eukaryotic cells and transmitted during reproduction. This DNA, which contains all the organism’s genetic information, directs the development, differentiation, and function of cells and specialized tissues. All cells of an organism contain the same genetic information, with the notable exceptions of T and B cells and some tumor cells. Of course, some degree of dissimilarity is seen between the genomes of same species members; these differences define our intraspecies variability. The obvious advantage of this type of library is that it contains all of the inherited genetic material for the species. In additional, the DNA sequences that regulate gene expression can be examined only by studying DNA obtained from genomic libraries. The major disadvantage of this approach, however, is the presence within a structural gene of large regions of untranslated sequence. Genomic libraries are usually screened by hybridization with DNA probes derived from cDNA clones. Alternatively, high-throughput PCR can be used to screen arrayed genomic libraries where DNA is available for each clone. Screening of genomic libraries should decrease with the completion of the entire human genome sequence and the availability of cloned DNA fragments.

A cDNA library is constructed from mRNA that is expressed by a cell or tissue at a specific point in time. This library reflects not only the type of cell or tissue used in the study, but also the cell’s response to various hormones, growth factors, and biophysical forces. cDNA does not exist in nature but is synthesized from mRNA, which specifies and directs assembly of protein sequences. Because mRNA cannot be cloned directly, the synthesis of cDNA allows for easy characterization of the coding portion of the structural gene and provides for rapid bioinformatics analysis of the conceptual protein. cDNA libraries offer several unique advantages. First, the coding region of a gene is usually much smaller than the genomic structural sequence that contains both intronic and regulatory sequences. Second, the cDNA can be used to direct synthesis of an mRNA that will encode the same protein as the original mRNA, allowing for rapid development of expression constructs for generating recombinant protein. This feature of cDNAs is used in the screening of expression libraries, where clones are evaluated on the basis of the proteins that they encode. Third, the numbers of genes that are expressed by a single cell or tissue are much smaller than the total number of genes in the genome, and thus the number of clones that need to be examined to identify a cDNA for a specific gene may be smaller. cDNA clones isolated from many cells and tissues have been identified through sequencing and made available to researchers. Availability of these EST clones and the development of PCR cloning is decreasing the need to screen cDNA libraries to obtain clones. However, cloning methodology will continue because of the need to identify genes that encode specific functions through expression cloning, or those that characterize a specific cellular process through subtraction cloning. cDNA libraries are usually screened with labeled DNA probes, but expression libraries can be screened with antibodies and other reagents that depend on protein function.

Obtaining a population of clones that represents the complete genome or all mRNA transcripts can be difficult. Mathematical models are used to ensure that the library is large enough to encompass the entire sequence repertoire (14). The number of independent clones that are formed by ligating the cloned DNA into a cloning vector is called the base of the library. A base of approximately 1 million clones is usually necessary for λ phage genomic libraries and cDNA libraries. The larger the base
of the library, the more likely it is that the gene of interest will be found. To examine all of the clones of a library, a number of clones representing at least three times the base of the library should be screened. Although the availability of genomic DNA for the construction of genomic libraries is rarely a problem, cDNA must be biochemically synthesized and is often limiting. The use of high-efficiency bacteria for the generation of the clones in the primary plating of a library is important to capture as many recombinant molecules as possible. The quality, purity, and size of the DNA are crucial to the generation of a library. High-quality DNA ensures an accurate representation of the DNA because exogenous, contaminating genetic material can cause erroneous results. The creation of a DNA library is not a trivial undertaking. Well-characterized genomic and cDNA libraries are available from academic and commercial sources. Libraries that have been amplified several times by plating on bacteria should be used with care because the library quality declines with overamplification, which can result in loss of rare transcripts.

**Genomic DNA Library**

All nucleated cells of an organism have the same general genomic information. Genomic DNA comprise coding regions or exons and non-coding regions or introns. Within the non-coding regions are sequences that control the patterns of individual gene expression (i.e., transcriptional regulatory elements). Most of the genome of higher eukaryotes is non-coding DNA. A complete genomic DNA library contains at least one copy of each region of the genome being evaluated. A sub-genomic library can be generated that contains only a defined portion of the species genome. Such libraries can be based on fractionated chromosomes, a particular chromosome band, or genomic DNA fragments of specific molecular weight. Genomic DNA libraries have been used to determine structural information such as the location of introns, exons, and transcriptional regulatory elements. In addition, genomic libraries have been used to clone genes with homology to other known genes and genes with little information regarding cellular expression pattern.

Genomic DNA can be prepared from any type of cell or tissue. Accessible cells, such as semen or blood cells, are an excellent source of genomic DNA. In the human genome project, a genomic library for nucleotide sequencing of the human genome was created from the blood of five donors (15). To establish a genomic library, high-molecular-weight DNA is carefully isolated and then randomly digested by partial restriction enzyme digestion or sheared by physical forces. Either process yields overlapping DNA fragments of optimum size to clone into a vector. The fragments of genomic DNA are usually inserted into vectors with a large carrying capacity, such as bacteriophage, cosmids, YACs, or BACs. When the starting sequence material is represented by fewer clones, screening of the cloned sequences is more rapid and thorough. Finally, generating sufficient numbers of clones to represent the entire genome is important in establishing genomic DNA libraries. A genome equivalent, which is an estimate of the number of independent clones required to adequately represent the entire genome (2), can be calculated by dividing the genome size by the average insert size. For example, a genome equivalent for a human genomic DNA library (3 × 10⁹ bases) composed of inserts of 30 kb would be 100,000 clones, creating a onefold library. A library of several genomic equivalents is usually needed to ensure representation of the entire genome. Vectors with high transformation efficiency and the ability to assimilate large DNA inserts minimize the complexity of the library (i.e., the number of independent genomic clones).

**cDNA Library**

mRNA is the genetic blueprint or coding sequence that dictates the primary amino acid sequence of the encoded protein. mRNA represents the combined sequences of the various exons that make up the structural gene, with intronic regions removed from an initial large RNA molecule by RNA splicing. The mRNA is only a small fraction of the total cellular RNA, which also includes ribosomal RNA (rRNA) and transfer RNA (tRNA). mRNA is a fragile, single-stranded molecule sensitive to RNase, a ubiquitous enzyme that degrades RNA. Single-stranded mRNA can be converted into cDNA in vitro by reverse transcriptase. The cDNA molecule can be used as a template to synthesize a second DNA strand, to produce double-stranded cDNA. cDNA is composed of sense and antisense strands; the sense strand has the same sequence as the mRNA. A cDNA library is a collection of different double-stranded molecules that represent the different mRNAs found in the cell or tissue being studied. The frequency of a single clone usually represents the relative abundance of the mRNA for that gene in the starting mRNA sample.

Isolation of mRNA is facilitated by its characteristic poly-adenylated 3' tail. These poly-A tails bind to oligo(dT) or poly(U) columns while the remaining RNA elutes off the column. Application of high-ionic-strength buffers to the column disrupts the non-covalent hydrogen bonds between nucleotides and permits the elution of purified mRNA. Total cellular RNA or mRNA is converted to double-stranded cDNA by reverse transcriptase in vitro. cDNA synthesis requires primers that can be either random,
such as oligo(dT), or gene-specific to begin the reaction. The type of primer used in cDNA synthesis is based on the starting RNA material and the intended use of the cDNA. Double-stranded cDNA is first ligated into a cloning vector and then transformed in a host cell to generate colonies for plating and selection. Because most genes are less than 5 kb, plasmids and bacteriophage are the most commonly used vectors for cDNA library construction.

Like genomic libraries, a critical issue for cDNA libraries is the adequate representation of all mRNAs expressed by a cell or tissue. The abundance of mRNA transcripts varies and is a function of the number of transcripts generated and the rate at which they are degraded. Rare mRNAs may be difficult to detect, whereas other transcripts may not be expressed in a particular cell or tissue. For cDNA cloning, therefore, determining which tissue source will be most enriched for the genes of interest is critical.

**RECOMBINANT LIBRARY SCREENING**

**Nucleic Acid Hybridization**

Nucleic acid hybridization involves the non-covalent association of nucleic acids from two complementary single strands into a double-stranded molecule. In a typical hybridization reaction, DNA/RNA from a library or gel is transferred to a nitrocellulose or nylon membrane. The membrane-bound, denatured double-stranded DNA or single-stranded RNA is exposed to a labeled or tagged nucleic acid probe. The probe binds to complementary RNA/DNA on the membrane, which is washed at various stringencies to remove nonspecific probe. Binding of the nucleic acid probe to the membrane or blot is assessed by autoradiography or phosphoimaging. DNA or RNA that hybridizes to the probe contains either the exact or a highly homologous DNA/RNA sequence. Northern blotting is the technique of RNA nucleic acid hybridization, and Southern blotting is DNA nucleic acid hybridization.

The typical nucleic acid probe is 50–1000 bases long and can be generated by several techniques, including oligonucleotide synthesis, PCR, and expression cloning. Oligonucleotide probes are single-stranded, synthetic molecules 15–50 bases long and are usually custom designed and purchased from a commercial supplier. Because of their short length, oligonucleotide probes can be designed to detect specific protein motifs like zinc-finger or other conserved functional domains. PCR with sequence-specific primers, cell-based cloning, and expression cloning can be used to generate nucleic acid probes. Probes are usually labeled by incorporating a nucleotide with a radioactively labeled phosphate (i.e., $[^{32}P]dCTP$) into the probe synthesis reaction. Less common labeling methodologies include digoxigenin-labeled probes, which can be detected with an enzyme-linked assay, and biotinylated probes, which can be detected with avidin-based systems. Colonies, plaques, or DNA/RNA fragments that bind the labeled probe appear as distinct spots or bands on the membrane. This hybridization information is then correlated with the original plate or gel to isolate the sequence of interest.

**Differential Screening, Subtraction Hybridization, and Subtraction Cloning**

These methods have been developed to clone cDNAs that represent genes that are differentially expressed between different cell types or the same cell type that is undergoing stimulation. Subtraction cloning is a powerful technique that generates a cDNA library or subtraction library, enriched for genes that are different between two conditions. The construction of subtraction libraries is based on the assumption that cells with or without a stimulus will express the same genes except for those genes affected by the stimulus. mRNA is isolated from two cell populations to create a subtraction library. For example, mRNA can be isolated from endothelial cells grown in the presence or absence of the inflammatory cytokine, tumor necrosis factor-$\alpha$ (TNF-$\alpha$). Single-stranded cDNA is synthesized from mRNA of one condition (e.g., endothelial cells exposed to TNF-$\alpha$) and hybridized to an excess of mRNAs from the second condition (unstimulated endothelial cells). DNA–RNA hybrids form between genes that are common to both conditions, whereas genes expressed in only one condition remain single-stranded. These single-stranded cDNA are then converted to double-stranded cDNA and ligated into a cloning vector to produce a library of the differentially expressed genes. This classic approach, however, is tedious and technically challenging. More simple PCR-based methods have recently been developed to generate subtracted libraries enriched for low transcript messages. This technique, called suppression subtraction hybridization, has been used extensively in our laboratory to clone genes that are expressed by endothelial cells undergoing angiogenesis in vitro (16,17).

Differential screening is similar to subtraction hybridization but does not require the creation of a new library. A cDNA library is chosen that is expected to contain the genes that are expressed in the cell under stimulation. Labeled cDNA is made from mRNA isolated from cells or tissues to be compared. The labeled probes from each condition are hybridized separately to replicate