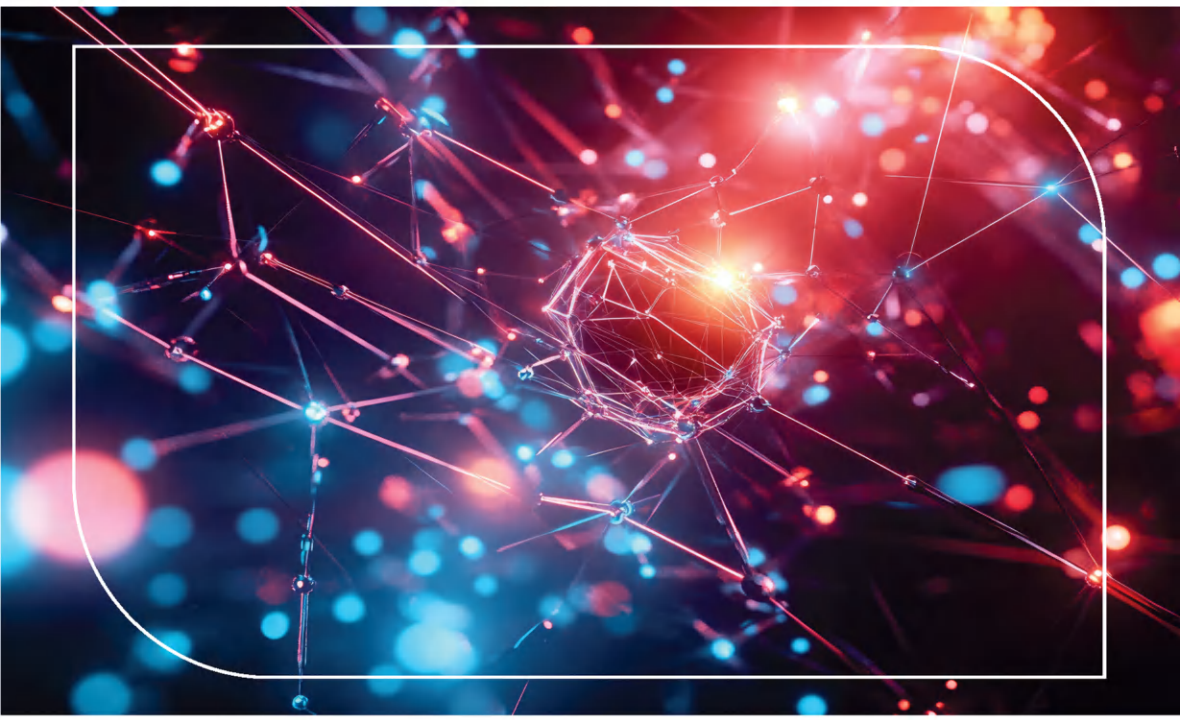
Molecular Genetics, Structures, Mechanisms, and Functions

Volume 1

Principles of Gene Manipulation and Genomics



Tariq Ahmad Bhat
Abdul Rauf Shakoori
Jameel M. Al-Khayri
Editors

CRC Press
Taylor & Francis Group
APPLE ACADEMIC PRESS

MOLECULAR GENETICS, STRUCTURES, MECHANISMS, AND FUNCTIONS

Volume 1

Principles of Gene Manipulation and Genomics

Molecular Genetics, Structures, Mechanisms, and Functions, 2-volume set
ISBN: 978-1-77491-946-0 (hbk)

Molecular Genetics, Structures, Mechanisms, and Functions, Volume 1: Principles of Gene Manipulation and Genomics
ISBN: 978-1-77491-878-4 (hbk)
ISBN: 978-1-00359-881-7 (ebk)

Molecular Genetics, Structures, Mechanisms, and Functions, Volume 2: Applications and Exploring the Nucleus
ISBN: 978-1-77491-948-4 (hbk)
ISBN: 978-1-00360-980-3 (ebk)

MOLECULAR GENETICS, STRUCTURES, MECHANISMS, AND FUNCTIONS

Volume 1

Principles of Gene Manipulation and Genomics

Tariq Ahmad Bhat, PhD
Abdul Rauf Shakoori, PhD
Jameel M. Al-Khayri, PhD
Editors

APP APPLE ACADEMIC PRESS

First edition published 2026

Apple Academic Press Inc.
1265 Goldenrod Circle, NE,
Palm Bay, FL 32905 USA

760 Laurentian Drive, Unit 19,
Burlington, ON L7N 0A4, Canada

CRC Press
2385 NW Executive Center Drive,
Suite 320, Boca Raton FL 33431

4 Park Square, Milton Park,
Abingdon, Oxon, OX14 4RN, UK

© 2026 by Apple Academic Press, Inc.

Apple Academic Press exclusively co-publishes with CRC Press, an imprint of Taylor & Francis Group, LLC

Reasonable efforts have been made to publish reliable data and information, but the authors, editors, and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors, editors, and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged, please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, access www.copyright.com or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. For works that are not available on CCC please contact mpkbookspermissions@tandf.co.uk

Trademark notice: Product or corporate names may be trademarks or registered trademarks and are used only for identification and explanation without intent to infringe.

For Product Safety Concerns and Information please contact our EU representative GPSR@taylorandfrancis.com Taylor & Francis Verlag GmbH, Kaufingerstraße 24, 80331 München, Germany

Library and Archives Canada Cataloguing in Publication

.....

CIP data on file with Canada Library and Archives

.....

Library of Congress Cataloging-in-Publication Data

.....

CIP data on file with US Library of Congress

.....

ISBN: 978-1-77491-878-4 (hbk)
ISBN: 978-1-00359-881-7 (ebk)

Dedication

This book is dedicated to



Abu Dawud Sulayman ibn Hassan Ibn Juljul (944–994 AD; his active period of research was 960–994 AD). He was an Arab physician and pharmacologist. His notable work is Tabqat-ul-atiba Wal-Hukama. His work on pharmacology was frequently quoted during the 10th and 11th centuries.



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

About the Editors



Tariq Ahmad Bhat, PhD
Senior Lecturer on Botany, Department of Education, Govt. of Jammu and Kashmir, India

Tariq Ahmad Bhat, PhD, is serving as a Senior Lecturer Botany in the Department of Education, Govt. of Jammu and Kashmir, India. He is engaged in active research of molecular biology, plant biotechnology, mutation breeding, and genetic improvement of legumes and medicinal plants. He has published more than 15 international books, 100 research papers, and many review articles and book chapters and has participated in 50 national and international conferences, training programs, and workshops. Dr. Bhat has 20 years of active teaching experience. He is serving as the District Coordinator Anantnag of a prestigious project on medicinal plants under financial assistance of the National Medicinal Plants Board (NMPB), New Delhi (Ministry of AYUSH), India. He is also founder member of Hon'ble CM's Super 50 initiative. He is a life member of several scientific societies and a reviewer for several international scientific journals. The Government of India conferred him with a Best Innovative Science Teacher Award 2014 to recognize his meritorious services. He was conferred with a Sir Syed Gold Medal along with certificate of merit (2022) for outstanding contributions in the field of education and research by the Jammu and Kashmir Alig Society.



Abdul Rauf Shakoori, PhD
Distinguished National Professor and Professor Emeritus, School of Biological Sciences, The University of Punjab, Pakistan

Abdul Rauf Shakoori, PhD, is a Distinguished National Professor and Professor Emeritus at the School of Biological Sciences at the University of the Punjab, Pakistan.

Prof. Shakoori has put in more than 60 years of meritorious service as a teacher and research scientist. He is among the pioneers who introduced modern trends in biological sciences and played a key role in introducing cellular and molecular biology as a subject of study and research at the universities of Pakistan. His research publications in the domains of bioremediation, bioinsecticides, and herbal medicine have significant implications on the economy of Pakistan.

Prof. Shakoori has published 663 original research articles in the journals of international repute, supervised research work of 77 PhD and 55 MPhil students. He has written 23 books published by national and international publishers. He has also contributed several chapters in a book published by Springer-Verlag, Germany. Prof. Shakoori developed collaborative research programs with scientists at the University of Massachusetts, Worcester, MA, USA, where he also held a position of visiting professor of cell biology for more than 25 years. His research collaboration also extended to the Estee Lauder Companies, Melville, NY, USA; Department of Pharmacy, King's College, University of London; Department of Agriculture and Environmental Science, University of Newcastle upon Tyne, UK; University of Marburg, Germany; University of Vienna, Austria; University of Essex, Colchester, UK; and Shandong Academy of Agricultural Sciences, Jinan, Shandong, China. Prof. Shakoori is elected fellow of The World Academy of Sciences (TWAS) (Italy) and the Pakistan Academy of Sciences and SIGMA Xi (USA). In recognition of his outstanding services for the development of science and his achievements in the field of cell and molecular biology, he was awarded the highest academic award, *Aizaz-i-Kamal,* for the year 1996 by the President of Pakistan; a Civil Award *Tamgha-i-Imtiaz* by the Government of Pakistan in 1999; Distinguished Scientist of the Year Award by Pakistan Academy of Sciences in 2011; and an international prestigious award in the field of Science and Technology by the Economic Cooperation Organization (ECO), which was conferred on him at Baku, Azerbaijan, at the ECO Summit in 2012. In 2019, he was honored and inducted as an academician in the Shandong Academy of Agricultural Sciences, Jinan, Shandong, China. Prof. Shakoori has played a pivotal role in establishing the Center of Advanced Molecular Biology (1983), Department of Microbiology & Molecular Genetics (2002), School of Biological Sciences (2003), and Cancer Research Centre (2019) in the University of the Punjab and the Center for Biotechnology Research, Shandong Academy of Agricultural Sciences, Jinan, Shandong, China, in 2020.

He earned his master's degree in zoology from the University of the Punjab, Lahore, Pakistan, in 1962 and his doctoral degree from University of Hohenheim, Stuttgart, West Germany, in the field of biochemistry/cell biology in 1972.



Jameel M. Al-Khayri, PhD
Professor of Plant Biotechnology, Department of Agricultural Biotechnology, College of Agriculture and Food Sciences, King Faisal University, Saudi Arabia

Professor Jameel M. Al-Khayri, PhD, is currently a Professor of Plant Biotechnology at the Department of Agricultural Biotechnology, College of Agriculture and Food Sciences, King Faisal University, Saudi Arabia, where he had formerly held several administrative posts as the Assistant Director of Date Palm Research Center; Head of Department of Plant Biotechnology; and Vice Dean for Development and Quality Assurance. He has been dedicated to the research of date palm biotechnology for the last three decades. He has published over 60 research articles and reviews in international journals in addition to 30 book chapters. He is the editor of several special issues of international journals on date palm, biotechnology, and sustainable agriculture under abiotic and biotic stress. He is the editor of 15 Springer reference books including *Date Palm Biotechnology, Date Palm Genetic Resources and Utilization* (2 volumes), *Date Palm Biotechnology Protocols* (2 volumes), and *Advances in Plant Breeding Strategies* (9 volumes). He has been appointed as a member of the editorial board and reviewers' panel of several international journals. He has participated in organizing scientific committees of international scientific conferences and has contributed to over 50 research presentations. In addition to teaching, he advises graduate students and conducts funded research projects. He is an active member of the International Society for Horticultural Science and Society for in vitro Biology and is serving as the National Correspondent of the International Association of Plant Tissue Culture and Biotechnology. He has served as a member of the Majlis Ash-Shura (Saudi Arabia Legislative Council) Fifth Session. Currently, he is leading an active research program on date palm, mainly focusing on genetic transformation, secondary metabolites, and *in vitro* mutagenesis to enhance

tolerance to abiotic and biotic stress. He is interested in the role of biotechnology in enhancing food security and the impact of global climate change on agriculture. He earned his BS in Biology in 1984 from the University of Toledo, Ohio, USA; and MSc in Agronomy in 1988 and PhD in Plant Science in 1991 from the University of Arkansas, USA.

Contents

Contributors ..... *xiii*

Acknowledgment ..... *xv*

Preface ..... *xvii*

1. Future of Genetic Engineering **1**
Aqsa Tariq, Ambreen Ahmed, and Iqra Mustafa

2. Biotechnology and Its Applications **29**
Tooba Momin and Malabika Bhowmik

3. Recombinant DNA Technology **41**
Tanveer Hussain and Muhammad Ishtiaq

4. Construction and Screening of DNA Libraries and the Molecular Analysis of DNA, RNA, and Protein **51**
Maria Butt, Mahnoor Imran, Tanzeela Rehman, and Uzma Qaisar

5. Genetic Engineering Tools: Restriction Enzymes and Vectors **75**
Roopashree M. S. and Venkatesh K. H.

6. Restriction Enzymes and Vectors **101**
Mohammad Amin Lone

7. Gene and Cell Division **119**
Aaradhya Bhatnagar, Sachchit Patel, and Arpit Shukla

8. Linkage, Crossing Over, and Recombination **143**
Mahnoor Imran, Tanzeela Rehman, Maria Butt, Muhammad Naveed Shahid, and Uzma Qaisar

9. Molecular Techniques and Methods for Mutation Detection and Screening in Plants **161**
Maham Akram and Ali Razzaq

10. Mutations and Mechanisms **185**
Muhammad Tariq Zahid, Habiba Ramzan, Ghulam Mustafa, Farah Rauf Shakoori, and Tariq Ahmad Bhat

11. Population Genetics 199
Shagufta Perveen, Faiza Akhtar, Nadia Yasmeen, Khadija Hafeez, Abida Parveen, and Sara Zafar

12. Sexuality in Bacteria 233
Mahnoor Imran, Maria Butt, Tanzeela Rehman, Bushra Tabassum, and Uzma Qaisar

Index ***255***

Contributors

Ambreen Ahmed
Institute of Botany, University of the Punjab, Quaid-e-Azam Campus, Lahore, Pakistan

Faiza Akhtar
Government College University, Faisalabad, Pakistan

Maham Akram
Centre of Excellence in Molecular Biology (CEMB), University of the Punjab, Lahore, Pakistan

Tariq Ahmad Bhat
Department of Education, Government of Jammu and Kashmir, India

Aaradhya Bhatnagar
Indrashil University, Rajpur, Gujarat, India

Malabika Bhowmik
Department of Biology, University of Dayton, Dayton, OH, USA

Maria Butt
School of Biological Sciences, University of the Punjab, Lahore, Pakistan

Khadija Hafeez
Government College University, Faisalabad, Pakistan

Tanveer Hussain
Department of Botany, Mirpur University of Science and Technology (MUST), Mirpur (AJK), Pakistan

Mahnoor Imran
School of Biological Sciences, University of the Punjab, Lahore, Pakistan

Muhammad Ishtiaq
Department of Botany, Mirpur University of Science and Technology (MUST), Mirpur (AJK), Pakistan

Venkatesh K. H.
Department of Life Science, Bangalore University, Jnanabharathi Campus, Bengaluru, Karnataka, India

Mohammad Amin Lone
Assistant Professor, Department of Zoology, Government Degree College Baramulla (Autonomous), Jammu and Kashmir, India

Tooba Momin
Department of Biology, University of Dayton, Dayton, OH, USA

Roopashree M. S.
Department of Life Science, Bangalore University, Jnanabharathi Campus, Bengaluru, Karnataka, India

Ghulam Mustafa
Department of Zoology, Government College University Lahore, Pakistan

Iqra Mustafa
Institute of Botany, University of the Punjab, Quaid-e-Azam Campus, Lahore, Pakistan

Abida Parveen
Government College University, Faisalabad, Pakistan

Sachchit Patel
Indrashil University, Rajpur, Gujarat, India

Shagufta Perveen
Government College University, Faisalabad, Pakistan

Uzma Qaisar
School of Biological Sciences, University of the Punjab, Lahore, Pakistan

Habiba Ramzan
Department of Zoology, Government College University Lahore, Pakistan

Ali Razzaq
Agronomy Department, University of Florida, Gainesville, FL, USA

Tanzeela Rehman
School of Biological Sciences, University of the Punjab, Lahore, Pakistan

Muhammad Naveed Shahid
Department of Botany, Division of Science and Technology, University of Education, Lahore, Pakistan

Farah Rauf Shakoori
Institute of Zoology, University of the Punjab, Lahore, Pakistan

Arpit Shukla
Indrashil University, Rajpur, Gujarat, India

Bushra Tabassum
School of Biological Sciences, University of the Punjab, Lahore, Pakistan

Aqsa Tariq
Institute of Botany, University of the Punjab, Quaid-e-Azam Campus, Lahore, Pakistan

Nadia Yasmeen
Government College University, Faisalabad, Pakistan

Sara Zafar
Government College University, Faisalabad, Pakistan

Muhammad Tariq Zahid
Department of Zoology, Government College University Lahore, Pakistan

Acknowledgment

The editors extend their appreciation to the support of the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia (Project No. GRANT5153).



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Preface

This book, *Molecular Genetics, Structures, Mechanisms, and Functions*, comprises of two volumes. Volume 1 is titled *Principles of Gene Manipulation and Genomics*. The main motivation for writing this book was to put together all classical and advanced aspects of molecular genetics, and gene manipulation in one consolidated place for beginners, experts, and those venturing into the fascinating science of molecular biology.

This book has 12 chapters that deal with the concept of genes, their relationship with chromosomes, and their functional manifestation to the benefit of organisms at large and humans in particular. Some chapters particularly deal with the tools of recombinant technology, which have brought about a revolution in our understanding of various complex biological phenomena. These tools, such as restriction enzymes, vectors, etc., have paved the way for creating organisms of choice and have new horizons in the field of medicine, agriculture, and industry for the enhancement of human welfare. This technology has made a profound effect on society, including better health through improved disease diagnosis, much better understanding of human gene variation, improved drug and pharmaceutical production, vastly more sensitive and specific crime scene forensics, and production of variety of improved livestock and crops.

The book has chapters that cover mutagenesis, construction, and sequencing of DNA libraries with repercussions in improving health, preventing genetic diseases, enhancing food resources, environmental bioremediation, etc. Special chapters have been included on the applications and future prospects of recombinant technology with a focus on national economy. A couple of chapters on linkage, crossing over and recombination, sexuality in bacteria, and population genetics have also been included to help understand different extrapolation of molecular genetics.

Volume 2: Applications and Exploring the Nucleus continues with the coverage of generic engineering, dealing with the concept of genes, their relationship with chromosomes, and their functional manifestation to the benefit of organisms at large and for human in particular. Topics include Mendel's Laws of Inheritance, which explains the inheritance of traits visible through generations; genome diversity and evolution genetic protein synthesis, recombination and evolution of DNA; transposable elements in

genetics; chromosomal aberrations; and more. The volume also addresses genetic engineering in agricultural science to increase crop yields, reduce costs for food or drug production, reduce the need for pesticides, enhance crop quality, etc.

All the chapters included in these two volumes have been contributed by brilliant young and upcoming biologists and highly experienced senior molecular biologists from different universities and research institutes of Pakistan, India, and USA.

CHAPTER 1

Future of Genetic Engineering

AQSA TARIQ, AMBREEN AHMED, and IQRA MUSTAFA

Institute of Botany, University of the Punjab, Lahore, Pakistan

ABSTRACT

Genetic engineering has led scientists to explore the DNA. Molecular tools allowed them to play with genome sequences and manipulate them for the benefit of mankind. Various advance techniques such as cloning, and CRISPR system make genetic modifications easier and help mankind in presenting best possible solution of various incurable diseases. Through these advance techniques mankind is facilitated in various fields, but illegal use of these technologies is hazardous to humanity. CRISPR-Cas9 is highly revolutionized genome-editing tool benefiting mankind which might be used to produce disease-free super species. Scientists have come way to far from genome editing of plants to animals and now humans. The cloning has shown bright future, but the experimentations on human germline to produce genetically modified embryo have become great concern due to their unexpected and uncontrollable consequences. Scientific communities do not agree on any clinical trials using human embryo. There must be some regulatory protocols that should be followed for the implications of these genetic modifications.

1.1 INTRODUCTION

Discovery of DNA as a hereditary unit led scientists to investigate more about its functionality and genetic implications. Molecular technologies such as DNA sequencing, restriction endonucleases, cloning, and genome

Molecular Genetics, Structures, Mechanisms, and Functions, Volume 1: Principles of Gene Manipulation and Genomics. Tariq Ahmad Bhat, Abdul Rauf Shakoori, and Jameel M. Al-Khayri (Eds.)
© 2026 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

editing made it easier to manipulate genomic DNA at organismic level. Initial investigations regarding gene manipulation involved the infusion of viral nucleic acid with cultivated cell lines and led to the adoption of markers of interest within plasmids. This has paved the way for microinjection and electroporation of cultured cells for random DNA integration. Furthermore, viral methods of transferring DNA to grown cells have developed at the same rate as physical methods (Lanigan et al., 2020). The efficient manipulation of nucleic acids of organism to modify genes and their control has been the ambition of research for long time. Scientists can investigate function of genes as well as can correct defects by genetic modification, but this method is frequently impractical due to the technical hurdle called site-specific nucleic acid targeting. Induced double-standard DNA (dsDNA) breakage in eukaryotic chromosomes was used to enable targeted gene editing, although difficult technologies based on the engineering of direct protein-DNA recognition were required (Knott and Doudna, 2018). Climatic change is a great concern for ecosystem biodiversity, productivity, ecophysiology, and even it alters plant growth. Its impact on agricultural system has proven to be the most complicated research predicaments that scientists have to cope with. Plants face these climatic ambiguities using various intricate and enigmatic approaches that vary among crops depending on region (Sood et al., 2019). Genetic engineering has the potential to revolutionize agriculture by increasing crop production, producing novel species and cloning to enhance the quality of life (Joy, 2020). Recombinant DNA manipulation began with cloning of DNA fragments using microbes that developed into a gigantic field where entire genome can be cloned and transferred from cell to cell using various strategies. Society might effectively overcome maladies by using genome sequences or embryos to dodge wellbeing issues, but it may allow the production of super species that are designed for certain specific characteristics leading to extraordinary separation in the fraternity (Sohn et al., 2020). This chapter delves into the advance techniques used for genetic manipulation through cloning and gene editing with reference to their applications in various fields, especially in medicine and agriculture.

1.2 GENETIC ENGINEERING: AN ADVANCE BIOTECHNOLOGICAL TOOL

Advancement in various techniques such as genetic engineering, cloning, and recombinant DNA technology has enabled analysts to manipulate DNA

sequences. The advancement in the field of biotechnology has enabled scientists to exploit nature to benefit mankind. Biotechnological tools have evolved tremendously over past few decades. Some major achievements in this field are presented in Figure 1.1. There are two popular advance genetic techniques used in the current era: (1) Cloning; (2) CRISPR Cas system.

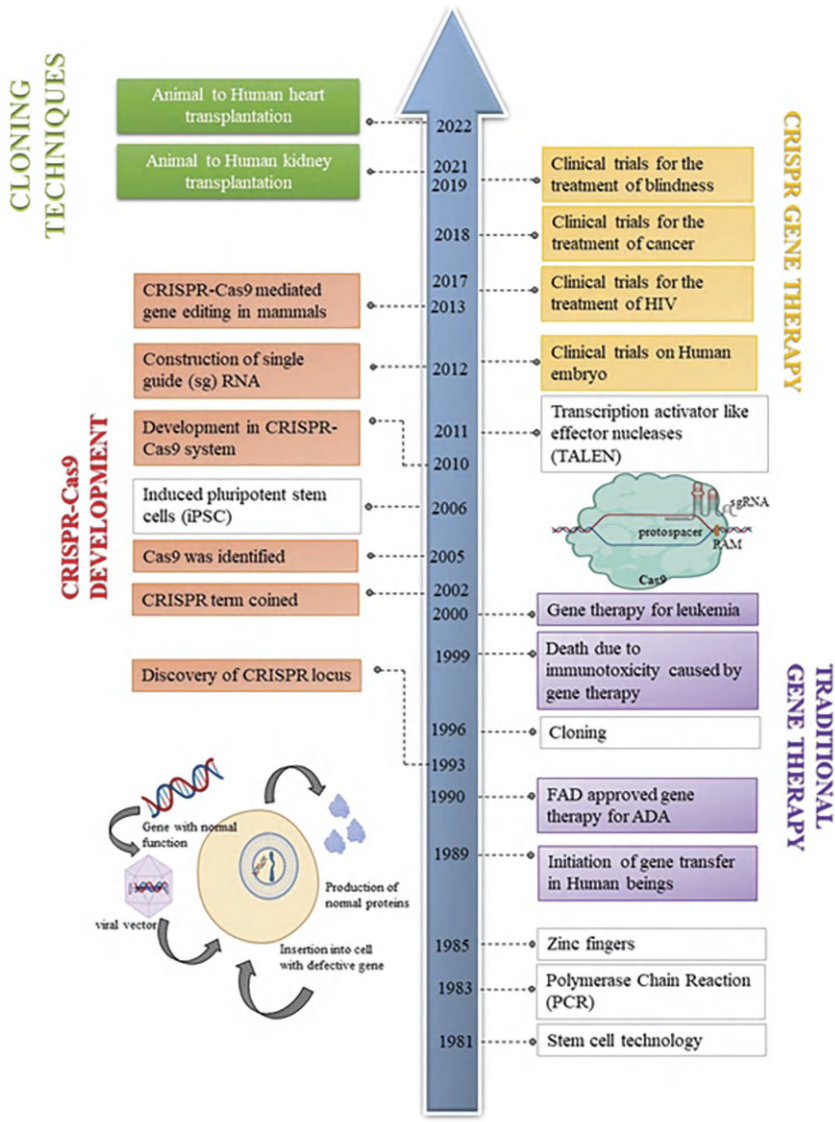


FIGURE 1.1 Major milestones achieved regarding genetic manipulation. ⏎

1.2.1 CLONING

Cloning is the process of making perfect genetic duplicate of a living organism. In biomedical research, cloning refers to the duplication of biological material. Gene cloning is one of the most remarkable discoveries of the time that is used to produce identical genes and DNA sequences that can be utilized for genetic engineering (Yu, 2022). The evolution of recombinant DNA technology has changed dramatically the biological field and allowed the scientific community to understand the functional and cellular aspects of various genes. Generally, gene cloning involves the insertion of desired DNA fragment in a suitable vector and its expression in the host. Advancement in recombinant DNA technology has led scientists to the development of living clones. The development of cloned embryos that are genetically identical to the parent, especially human embryos, and their use in scientific studies is a hot topic in the current era (Niemann and Lucas-Hahn, 2012).

1.3 SOMATIC CELL NUCLEAR TRANSFER (SCNT)

Duplicating living organisms via somatic cell nuclear transfer (SCNT) technique is emerging as an advance biotechnological tool that provides great opportunities. SCNT is a sequential process consisting of production of cytoplast through removing nucleus from oocyte, choosing appropriate donor cell, recreation of embryo, activation, and culturing. Somatic cloning has great significance in reproductive and therapeutic cloning. Reproductive cloning signifies the development of a new living organism with same genetic information as the donor cell. On the other hand, in therapeutic cloning, mature cell of a living organism is grown for a short period (Matoba and Zhang, 2018). Therapeutic cloning is useful for the generation of new tissues and cells for the treatment of various diseases, including leukemia, Alzheimer's, and Parkinson's diseases, and even to produce whole complex tissues and organs.

1.4 MECHANISM OF SOMATIC CELL CLONING

The classical methods of cloning mainly use micromanipulators. In handmade cloning, all the procedure has to be carried out without the aid of micromanipulators and embryos with removed zona pellucida which were

then cultured under control conditions until the formation of blastocysts which was then transferred to recipients. Oocytes at the stage of metaphase in second meiosis are the most suitable recipient to produce viable clones. These mature oocytes then enucleated by removing small portion of cytoplasm and then inserted into donor cell under zona pellucida. Various somatic cells and fetal fibroblasts are also used in somatic cloning. After the insertion of enucleated oocyte into donor cell, it gets fused. After their fusion, the cell gets activated by providing short electric pulses or by exposing the cell to certain chemical compounds. The activated cell is then cultured *in vitro* into blastocysts which is then transferred into recipient animal (Fig. 1.2) (Niemann and Lucas-Hahn, 2012; Czernik et al., 2019).

1.5 DOLLY—A BREAKTHROUGH IN BIOTECHNOLOGY

Cloning was initially considered impossible, but the advent of advance bioengineering tools led scientist to accomplish this out of question activity by the creation of Dolly, the cloned sheep. Dolly was the first somatic cell-mediated cloned mammal. The generation of Dolly has answered many questions regarding cellular and genetic equivalence in the adult organism (Alberio and Wolf, 2021). The generation of Dolly has opened many doors in the field of agriculture, livestock farming and medicine. Nuclear transfer performed under controlled conditions can remove epigenetic memory of somatic cells which has led to various rounds of cloning without losing its ability (Wakayama et al., 2013). This describes the strength of oocytes reprogramming ability. Moreover, physiological characters of cloned Dolly sheep were equivalent to the control animals of the same age which specify that nuclear transfer does not effects the aging. Following the generation of Dolly, many other organisms were cloned including dogs, pigs, rabbits, horses, mules, goats, and cats (Yu, 2022). In 2006, Japanese scientists became successful in creating heart and brain cells using adult stem cells (Mader and Windelspecht, 2017). After this, in 2019, scientists successfully created an animal embryo containing human cells and transplanted it into a surrogate animal (Cyranoski, 2019). These advancements suggest that scientists are continuously trying to circumvent the restrictions associated with cloning to introduce improved techniques in this field. However, the generation of chimera is controversial and has ethical concerns, but new laws permit transplantation of embryos in animals for the advantage of humans.

1.5.1 BENEFITS AND DRAWBACKS OF CLONING

Previously, genetic modification in animals relies on embryonic stem cells (ESC) of mouse and production of chimeras. These methods were not used for genetic modifications in livestock due to nonavailability of germline competent ESC. Later, pronuclear microinjections were used for genetic modifications in large animals. These are inefficient and facilitate random insertions of transgenes. However, nuclear transfer techniques are efficient for the production of bioengineered animals (Yu et al., 2022). The SCNT technique was first applied in sheep and soon extended to other species and remained a frequently used genetic engineering tool. It has advantage over ESC technology that results in the production of chimeric offspring. It is a significant tool for the generation of animal model diseases, generation of genetically modified animal donors for xenotransplantation, production of nutraceuticals, preservation of genetic resources, and for enhancing the productivity of livestock animals. However, in SCNT technology gene targets in somatic cells are less proficient than gene targets in pluripotent stem cells. This has led to the development of nuclease-mediated genome genetic modification that expands the technologies used for genetic modifications (Klinger et al., 2021). Figure 1.2 summarizes SCNT for genetic modifications in various fields.

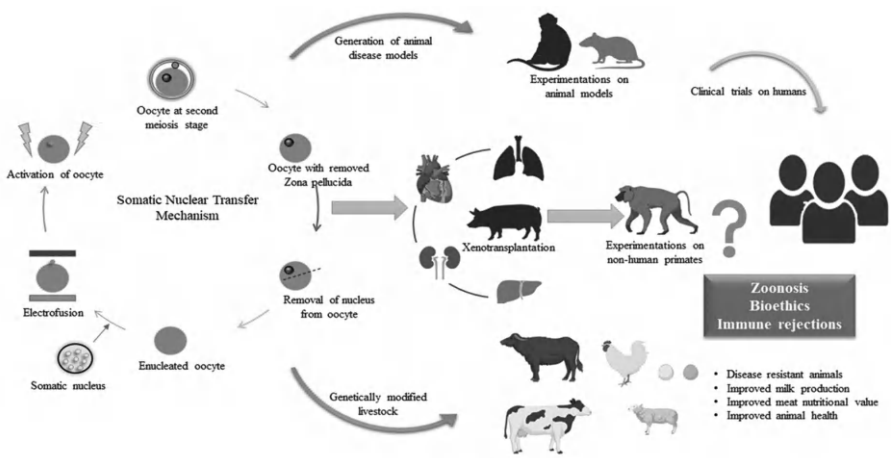


FIGURE 1.2 Somatic nuclear transfer technology (SCNT) for the creation of GMOs. ⏎

1.5.2 ANIMAL MODEL FOR HUMAN DISEASES

Scientists have developed animal models to study human diseases due to similarity in anatomy and metabolic characteristics. Genetically engineered animal models of various cardiovascular diseases, diabetes, cystic fibrosis, various types of cancer, Duchenne muscular dystrophy (DMD), and neurodegenerative disorders have been developed that closely reiterate the physio-pathological characters of human diseases (Schneider et al., 2020; Caballero et al., 2021). Scientists use these models for the development of their cure and other diagnostic techniques (Moretti et al., 2020). Initially, small rodents and large animals (pigs) were cloned for scientific investigations. Primates act as exemplary models to study various diseases associated with human diseases like brain disorders and cognitive functions. Healthy monkey clones have been produced for this purpose (Liu et al., 2018). The technique of creating gene targeted animal models is laborious, requires systematic cell screening and sequential cycles of nuclear transfusions, and, consequently, low number of healthy viable offspring are obtained. Thereby, the development of an efficient method is needed which drastically increases efficiency of gene modification. SCNT cloning of non-primates combined with advance genome-editing technique such as CRISPR-CAS9 serves as the beginning of a new age of human clinical science.

1.5.3 XENOTRANSPLANTATION

High demand and lesser provision of donors have led the scientist to use the SCNT technique for the production of modified progenitor animals for xenotransplantation. However, organs derived from genetically modified animals require precise protocols that follow safe and ethical sustainable source for organs (Kogel and Marckmann, 2020). Heart transplantation from pig to non-human mammals has already been done. However, for human transplantation, nuclear transfer technology was utilized to produce donor pigs that express protein CD46 and human thrombomodulin instead of alpha-1,3-galactosyltransferase (GGTA1). This transplanted heart survived for 900 days after transplantation (Alberio and Wolf, 2021). However, this has a limitation due to continued growth of pig heart. Inactivating growth regulatory hormone receptor gene is a promising method to solve this dilemma (Hinrichs et al., 2021). Recently, first animal to human heart transplantation was carried out at University of Maryland, USA. A transgenic pig heart was transplanted proficiently into a 57-year-old man. After the successful surgery, the recipient was able to survive 61 days freely without

the aid of cardiopulmonary bypass (Kobashigawa, 2022). This event is considered a great milestone in genetic engineering for organ transplantation; however, animal organs would come up with great challenges such as impact of immunosuppressants to control the host rejection response during xenotransplantation (Wang et al., 2022a, Kavarana et al., 2022). Before this transplantation event surgeons used same line of genetically modified pigs in 2021 for kidney transplantation into two legally dead humans with no visible cerebral activity. The recipients showed no organ rejection and kidneys functioned normally and the patients were sustained successfully on ventilators (Reardon et al., 2022). However, long-term effects of xenografting and longevity of transplanted organs still need to be focused due to lifespans difference between animals and humans. In addition, there are many challenges in animal-to-human transplantation that needs to be answered such as immune rejection, transmission of heterogenous pathogens, and mainly ethical concerns (Silverman and Odonkor, 2022).

Cross-species xenotransplantation has significant bioethical concerns. Some of the ethical considerations regarding xenotransplantation include hazards involved during clinical trials, animal rights concern, psychological issues, xenograft rejections, community health risk of zoonosis, and self-estrangement (Cengiz and Wareham, 2019; Tuck, 2020). Moreover, potential of animal mistreatment, violation of animal rights, and objectification of animals diverts the attention of scientific community for ethical consideration of animals (Oldani et al., 2019). This is also a challenge for humanity involving threat to human dignity, loss of recipient identity, disrupt moral distinction, etc. The xenograft recipient may suffer from psychological stress and disorder of identity, and various self-image issues arise that have impact on religious dimensions (Thomas et al., 2020). Moreover, it goes beyond the accustomed natural orders (Cengiz and Wareham, 2019; Kwisda et al., 2020). Another bioethical issue with this technique is the use of chimeras in xenotransplantation which exploits human and animal biological material. However, this is generally acceptable because of potential benefits to humans (Mann et al., 2019).

1.5.4 IMPROVED LIVESTOCK PRODUCTION

Genetic engineering is also used for genome editing to improve quality and productivity of livestock and to meet the increasing demand of healthier and nutritious animal products. Previously, microinjections of foreign DNA into zygote were used for the generation of transgenic animals. However, SCNT

technology is now preferred due to its efficiency and precision of gene transfer. This technology is employed for the production of improved meat by modifying proportion of unsaturated fats since these fat-containing feeds are not environment friendly. Nuclear transfer-mediated removal of Caenorhabditis elegans fat-1 desaturase gene in cattle makes them high source of omega-3 fatty acids which is significant to produce nutraceutical products (Zhang et al., 2013). Moreover, cloning for propagation of desirable traits in livestock is primarily restricted to meat and milk-producing animals. Generation of hypoallergic milk by terminating β-lactoglobulin, lactoferrin, and lysozyme in cows is widely used (Kaiser et al., 2017). Mutations at MSTN gene caused double muscle phenotype in some cattle resulting in 20% more muscular mass which is significant for the health of cattle. Similarly, genetic engineering is used to produce disease-resistant animals. Genetically modified pigs with reduced susceptibility to PRRSV infections were produced by causing the mutation of CD163 which results in the prevention of viral infections (Burkard et al., 2017). Similarly, using this technique, animals resistant to various diseases such as transmissible gastroenteritis virus (TGEV), African swine fever were also produced (Whitworth et al., 2019; McCleary et al., 2020). Hence, researchers are struggling to introduce favorable traits to cattle for improved production, health, and welfare of animal.

1.5.5 ABNORMALITIES IN ANIMAL CLONES

The primary concern for SCNT is the reversion of differentiated somatic cells to create embryonic stem cell-derived blastocysts that could be used for therapeutic cloning. However, it has a low efficiency of 1–5% in creating normal viable clones and a high number of abnormalities have been seen due to epigenetic reprogramming failures (Srirattana et al., 2022). Reproductive cloning results in loss of clone at any stage during post- and preimplantation and post- and prenatal development (Long et al., 2014). The first defect in the detained clone is genomic instability caused by ceasing cell cycle program (Chia et al., 2017). This has revealed that epigenetic processes are involved in differentiation affecting transcription and preventing DNA replication. These are due to instability of genome which has led to the failure of transcriptional reprogramming. Miscarriage and fetal mortality were frequently observed in living clones. These developmental defects are due to the incomplete somatic cell reprogramming by SCNT. The inadequate reprogramming leads to abnormal gene expression, which can impact placental and fetal development (Malin et al., 2022). Abnormal fetal development results in large offspring

syndrome that has various phenotypes. At the time of birth, abnormally high birth weight, loss of motor control, enlarged tongue, abnormal organ size, development of respiratory disorders, and weak immune responses in young clones have been observed. Moreover, obesity is also observed in adult clones (Gouveia et al., 2020). Contrary to that, in some cases, underweight and underdeveloped placentas are also observed such as in pig and murine clones. In mice, aberrant histone modifications and DNA methylations are associated with epigenetic abnormalities (Chen et al., 2022).

1.6 CLONING SUPPORT HUMAN HEALTH

SCNT technology is now implemented in medicine for the treatment of various human health issues. During the current global outbreak of COVID-19, scientists have come up with the use of SCNT technology to overcome this situation. SCNT technology has capacity to use regenerative medicines for biomedical applications. The potential genome reprogramming ability makes SCNT a suitable therapeutic strategy for COVID-19. This technology is ideal for the production of cloned embryo and embryonic stem cells with enhanced drug testing ability in COVID-19 (Singh et al., 2021).

1.7 HUMAN CLONING

Human cloning is one of the most recent techniques which would greatly impact various aspects of human life in the future. In human cloning, genetically identical individuals are produced through cloning technology. It can be accomplished either by embryo cleaving or by nuclear transfer approach. In the embryo splitting method, the developing embryo is split into separate parts and every split part will then develop into blastocyst which will evolve into a child. However, only a finite number of cloned human beings have been produced through the embryo splitting technique (Jasim, 2022). Human cloning has provided a gateway to answers relevant to human infertility, organ transplantation, selecting desired human characteristics, and developing humans. Due to extreme sociological and physiological considerations, human reproductive cloning is prohibited. A cloned embryo must undergo a series of molecular tests before its implantation to determine the liability of clone. However, this is considered unethical due to high death risk (Nasrullah et al., 2020; Jain et al., 2021). Human cloning has raised several important questions worldwide due to various scientific and religious challenges like origin of reproduction and clone identity, safety of designed children, and

psychological issues of cloned individual in society. Human cloning will affect various social ethics and will violate human rights (Kiyani and Munir, 2020). The dilemma of acceptance of clone in the society is not answered yet. Unlike normal embryo clone is considered abnormal as its genome is the combination of various sources which means that the original qualities of child are mingled. However, these are some expected issues related to human cloning. Its long-term consequences are still vague. This nonsexual human production will involve multidisciplinary issue that is still under debate of many scholars of various fields, including ethics, scientists, religions, and medical sciences (Nwoye, 2019; Sarkar, 2022).

1.7.1 GENOME EDITING VIA CRISPR/CAS

Gene modulation is a kind of genetic engineering technique in which DNA is manipulated at single-base pair level for the identification of gene affecting specific phenotype of organism through protein expression. This approach is used recently to produce super-species with selected phenotypic traits (Panesar and Marwaha, 2013). Genome-editing strategies were initially developed in yeast in which the homologous recombination repair pathway was used for the substitution of small portion of genomic DNA with exogenous piece of DNA. This approach was later used in mouse to study human diseases, for drug discovery and development. However, low efficiency and unwanted genome editing limits this approach. Later targeted genomic modifications via mega-nucleases or DNA scissors were discovered. Nucleases are highly site specific and cleave DNA at rare recognition sites to stimulate homology-directed repair (HDR) leading to accurate genetic modifications. However, besides HDR this strategy also leads to some nonhomologous end joining (NHEJ) by introducing random insertions or deletions in genome causing random mutations (Van-Eenennaam, 2017). HDR needs 5` resection and causes partial 5`-3` nucleolytic degradation of 5` end to generate 3` single-stranded overhangs to initiate repair mechanism. Hence, the homologous recombination technique can be used for precise insertion, deletion or substitution of specific genetic bases in genetic code at nicked position (Tröder and Zevnik, 2022). NHEJ is preferred for gene knockouts to study frameshift mutations, whereas HDR introduces properly designed templates (Anzalone et al., 2022). For more precisions, scientists have developed next-generation CRISPR/Cas9 system, i.e., Cas9D10A or Cas9n that only produces single-stranded nicks and does not switch on the NHEJ pathway (Cong et al., 2013). The alternative approaches like the use of DNA-protein

site-specific nucleases, i.e., Zinc finger (ZFs) and transcription-activator-like effector nucleases (TALENs), were taken in consideration to avoid problem of NHEJ. These are specially designed proteins that virtually cleave any target gene sequence in genomic DNA with high specificity, but their regular use is hindered by the difficulty encountered with their designing and validation. Recent innovations in genome modification systems especially CRISPR (clustered regularly interspaced short palindromic repeats) are of great interest due to their flexibility and accessibility for genome editing and visualization. The CRISPR-Cas9 system depends on a 20-nucleotide long single-stranded guided RNA (sgRNA) that pairs with target DNA and DNA protospacer adjacent motif (PAM) (Doudna et al., 2014). This technique has been gaining its success in various fields including agriculture, aquaculture, and humanitarian applications (Fig. 1.3).

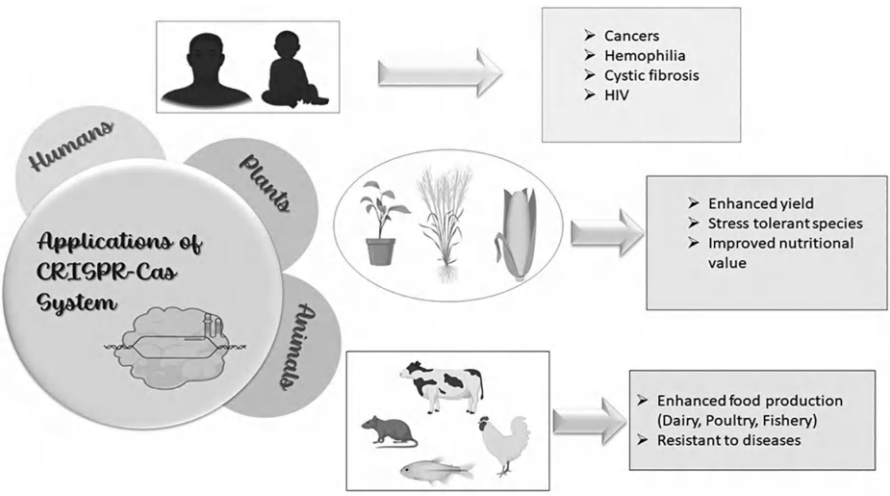


FIGURE 1.3 Applications of CRISPR-Cas system. ⏎

1.8 CRISPR/CAS SYSTEM

Genomic stability is essential for the survival of organisms, therefore, multiple mechanisms to repair and protect the genomic machinery have evolved within organisms. Generally, CRISPR/Cas locus has number of Cas genes and CRISPR array with several unique repetitive sequences, upto 30-40 bp long, Regularly interspersed with short spacer sequences (Jacinto et al., 2020). These protospacer guide Cas nuclease to produce double stranded breaks on invading DNA. Based on the protein effectors, the

CRISPR/Cas system is divided into two classes. Class 1 is specified by the existence of multi-protein effector complex and has three main types (I, III, IV) and various subtypes. Class 2 has single protein effector module and has three types (II, V, VI) (Wright et al., 2016; Makarova et al., 2015). General mechanism of CRISPR/Cas has three basic stages. During first stage, the Cas proteins excise the fragment of invading DNA or mobile genetic elements (MGE) which are then incorporated into CRISPR array. The CRISPR array is transcribed as single RNA known as precursor crRNA (pre-crRNA). In the second stage, this pre-crRNA is further treated to produce crRNA which is a then integrated with Cas nucleases. In last stage, the crRNA identifies the invading sequence and guide Cas nuclease to cleave the invading mobile genetic element. The Cas nuclease activity requires the existence of PAM on only invading DNA (not on CRISPR array) to avoid self-targeting and degradation (Hille et al., 2018). The CRISPR system type-II uses Cas 9 effector, bacterial RNase III and transactivating (tracrRNA) RNA molecule, Types V and VI, Cas12 and Cas13 effector proteins respectively (Jinek et al., 2014; East-Seletsky et al., 2016). This type II system is widely adopted by scientific community as gene editing tool due to its simplicity. The CRISPR type II is unique because only one Cas protein (Cas 9) is used for gene slicing. Moreover, two domains of nuclease Cas 9 i.e., NHN, RuvC are involved in cleaving the target DNA (Nishimasu et al., 2014). The gene editing mechanism via CRISPR Cas9 has been described in Figure 1.4.

1.9 APPLICATIONS OF CRISPR/CAS-9 SYSTEM

In recent years, CRISPR Cas-9 is known as one of the most successful techniques for gene manipulation. The advancement in this technique has led to a wide range of applications in different fields and opened the innovative doors of research (Tanaka and Gomi, 2021). CRISPR Cas-9 has diverse applications, especially in the field of medicine. It has been applied for the cure of various diseases such as cancer, infectious viral diseases, genetic diseases and detection of various pathogens, HIV, hepatitis, hereditary diseases, and liver infections. This system has now been successfully used for the development of mice monogenic disease model for the detection of mutations (El-Mounadi et al., 2020). In human cell models, this system is successful in replacing the viral antibodies with antibodies resistant to various viral infections like Epstein-Barr virus (EBV), human immunodeficiency virus (HIV), influenza virus, Respiratory Syncytial Virus (RSV) (Moffett et al., 2019).

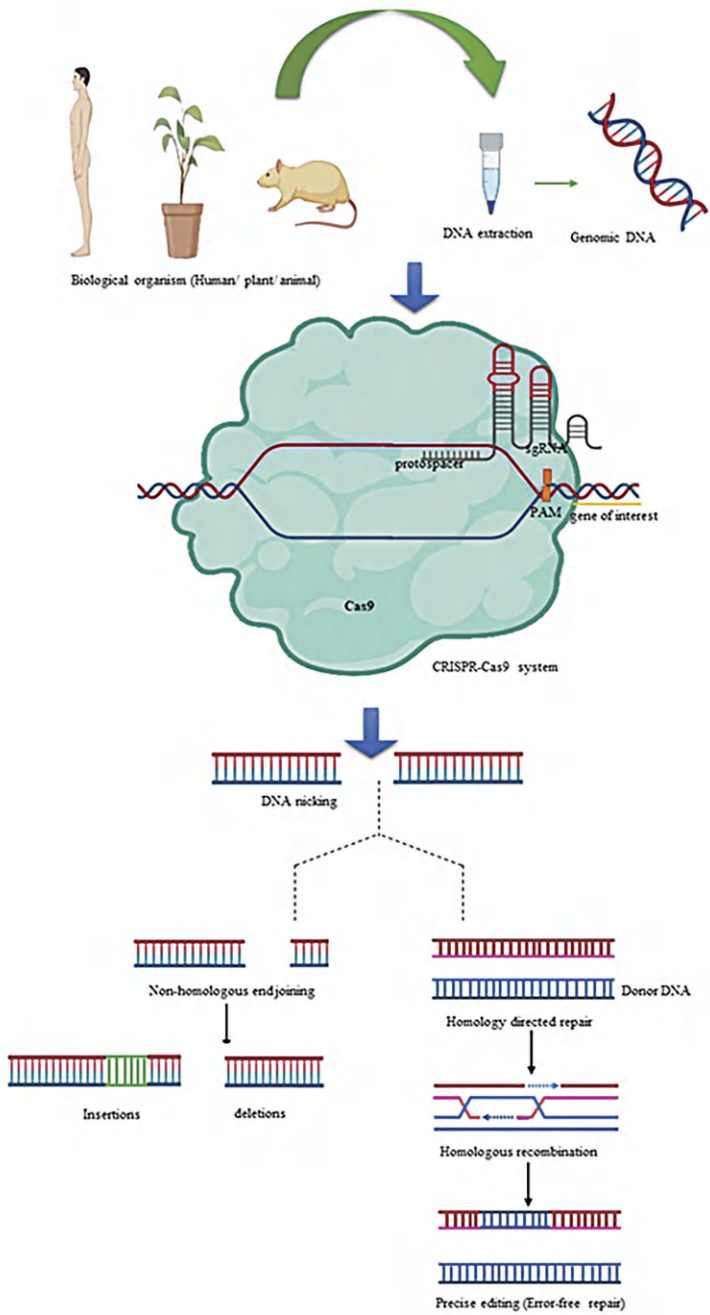


FIGURE 1.4 CRISPR-Cas9 system as genome editing tool for homologous and non-homologous repairs resulting in precise error free repairs or insertions/ deletions.

1.9.1 CRISPR-CAS 9 SYSTEM IN CANCER TREATMENT

In recent years, many people have been adversely affected by cancer. CRISPR Cas-9 has been considered an effective technology for cancer treatment. CRISPR-Cas 9 has drawn more attention as an advance technique in cancer therapy by quickly introducing gain-of-function (GOF) and loss-of-function (LOF) mutations in oncogenes or tumor suppressor genes (Nelson et al., 2022). It has transformed cancer research by identifying promising tumor-associated genes and modeling cancer *in vivo*. For instance, introduction of frequent LOF and GOF through CRISPR-Cas 9 system into human colorectal cancer (CNC) suggested that additional epigenetic studies are required. CRISPR-Cas9 system delivery and selective activation needs more research to ensure that modifications occur only where it is required (Shademan et al., 2022; Wen et al., 2020). To express the target genes in cancer therapy, dCas9 is conscripted by specific guide RNA at a specific target site (Jacinto et al., 2020). The CRISPR Cas9 system can be employed against oncolytic viruses to enhance the immunity, for example, deletion of ICP6 (ribonucleotide reductase) and ICP34.5 neuro-virulence genes. Another application is epigenome editing which involves alternation of target DNA by the fusion of Cas9 to histone modifiers and proteins. Cancer can be treated by this process since many types of cancers involve epigenetic factors. Use of CRISPR Cas9 is one of the challenging approaches to decrease the proliferation of tumors (Mansouri et al., 2021). *In vitro* editing of T cells for the expression of chimeric antigen receptors (CARs) has been studied and results showed that genetically engineered T cells express CARs that precisely identify and kill onco-cells. These findings suggested CRISPR-Cas9 as an effective tool for the treatment of diseases like leukemia and lymphomas (Linguanti et al., 2022). Various clinical trials are evaluating CAR T-associated treatments targeting expression of antigen, CD19, by various tumor cells, i.e., Mesothelin cMet, EGFR, Her2/neu, interleukin-13 receptor, GD2, and IL13Rα2 (Khorasani et al., 2022). However, using CRISPR technology to treat cancer has one limitation that it can be performed in immune cells as *ex-vivo* gene therapy. It is a time-taking mechanism with costly steps such as isolation of gene of interest, activation, and cell expansion (Buquicchio and Satpathy, 2021). Cancer immunotherapy is one of the best techniques to increase the immunity against tumors as it is more beneficial than chemotherapy and radiotherapy due to low cost and more efficiency. The viruses that are responsible for cancer are oncolytic viruses which can be modified genetically as an antiviral defense. Other methods involve direct lysis of tumor

cells which further decrease the proliferation of tumor cells and trigger the immune stimulation through antigens (Martinez-Lage et al., 2018).

1.9.2 CRISPR-CAS 9-MEDIATED IPSC MODELS

Somatic cells have the ability to switch to pluripotent state. These induced pluripotent stem cells (iPSC) are promising in the development of human disease, medicine, and drug discovery models (Hovland et al., 2022). The CRISPR-Cas9-based technique is commonly used for the generation of iPSC-based human disease models. Cas-9-mediated generation of isogenic disease models is used to study gene functions. Knockout events via the NHEJ pathway identify whether specific mutation is responsible for disease (Smith et al., 2015). Alternatively, disease-associated mutations are introduced through the CRISPR-Cas9 system into iPSC- or HDR-mediated editing for the development of human disease models (Okano and Morimoto, 2022). These models decipher genomic mechanisms related to specific disease. Researchers studied pathophysiological processes using Cas9-based iPSC models to understand the underlying mechanisms involved in cardiomyopathy and mitochondrial disorders of Barth syndrome (BTHS) and Tafazzin gene (TAZ), respectively (Caudal et al., 2022). Through this, researchers were able to access potential therapies for these diseases.

1.9.3 CRISPR-CAS9 IN GENE THERAPY

Gene therapy is the process of getting normal genes either by substituting or by removing the defective genes. It is used to treat various genetic illnesses. The first FDA-approved gene therapy treatment was regulated in 1990 to prevent the production of adenosine deaminase (ADA) from defective gene. It is an inherited condition of having defective immune system leading to the ability to protect against foreign agents (viruses and bacteria) (Lange and Kappel, 2022). Gene therapy is one of the most interesting benefits of the CRISPR-Cas system. CRISPR-Cas-mediated therapy induces genetically manipulated mutations to provide definitive cure. Gene editing and gene additions through the CRISPR-Cas9 system could be employed to treat many genetic diseases that affect proteins and cells used for blood formation. Many diseases are addressed by gene addition or gene editing, for example, protein deficiency disease hemophilia treated by using *in vivo* adeno-associated virus (AAV) and diseases of blood cells that is sickle cell anemia by *ex vivo* gene therapy

(Kohn, 2019; Walters et al., 2019). Researchers used the CRISPR-Cas tool to study genetic variations in humans by using patient-derived primary cells such as Fanconi anemia, beta thalassaemia, Duchenne muscular dystrophy (DMD), cystic fibrosis, and hemophilia (Fajac and Sermet-Gaudelus, 2022; Sipe et al., 2022; Wang et al., 2022b). CRISPR-Cas-mediated gene therapy aims to restore normal gene function and is an ideal treatment for genetic illnesses, caused by one or more abnormalities in the genome while gene editing enables targeted modifications of the genome to fix or change genetic sequences. Gene therapy achieves this goal by inserting the duplicate gene of interest to the target genome. The existence of a changed gene permits the expression of a functional protein, eradicates the root cause of disease, and enhances the function of the entire organ. Monogenic hereditary illnesses that are caused by abnormalities in just one gene are best treated with CRISPR-Cas-mediated gene therapy (Dwivedi et al., 2022). Gene therapy and gene-editing technologies generally involve complex processes, making it challenging for the public to comprehend their workings, possible advantages, and effective dissemination of new technology (Ginn et al., 2018).

1.9.4 CRISPR-CAS 9 SYSTEM TO PRODUCE GMO

The CRISPR-Cas9 system has shown interesting results in the generation of genetically modified organisms (GMO). The already available genetically modified animal models were slow and time-consuming which involve several labor-intensive processes. However, the CRISPR-Cas9 system makes it an efficient gene-targeting process that can generate mutant-organism with multiple genes in single zygote editing. The CRISPR/Cas system can result in effective gene knockouts at various loci detouring targets in embryonic cells. Multiple genes can be target simultaneously through CRISPR technology. Microinjections through CRISPR-Cas 9 into zygote can produce multiple mutations in existing GMO without deriving embryonic stem cells. CRISPR has revolutionized the development of genetically modified organisms by producing animal models for potential applications by lowering the cost and time consumption (Shimatani et al., 2017).

1.9.5 CRISPR-CAS 9 SYSTEM IN CROP PRODUCTION

Modification of plant cells with guided RNAs and Cas-9 is similar as a formation of genetically modified plants. The expression cassettes incorporate

promoters, transcription terminators, and resistance markers for the targeted phenotype of interest. The vector is introduced into *Agrincorporaatobacterium tumefaciens* which is further used to transform plants. This would lead scientists to modify genes present in plant. Its simplicity and efficiency allow faster genetic modifications for crop production and of sustainable agriculture (Toda et al., 2019; Wurtzel et al., 2019). CRISPR-Cas9 targets to improve plant metabolic processes to develop resistance against various ecological stresses, improvement of nutritional content, enhanced yield, grain quality, etc. Plants modifies via CRISPR-Cas9 include soybean, tomato, rubber producing plant, thermos-sensitive genic male sterility in maize and wheat, enhanced nutritive content in sorghum and wheat, tolerance to various stresses, pathogenic resistant species (Iaffaldano et al., 2016; Pyott, 2016; Li et al., 2018; Okada et al., 2019). Some other examples of achievements through CRISPR-Cas9 in agriculture are the development of amylopectin-producing potato species by eliminating gene responsible for the synthesis of granule-bound starch synthase (GBSS). This is a highly attractive commercial attribute (Andersson et al., 2017). Similarly, in cucumber production of varieties resistant to cucumber vein yellowing virus (Genus Ipomovirus) and potyviruses zucchini yellow mosaic virus and papaya ring spot mosaic virus are one of the great achievements of CRISPR-Cas9 system in agriculture (Chandrasekaran et al., 2016). Engineering of genetic resistance of plants against pathogens for which natural resistance is not detected has great potential to manage diseases to produce healthier plants such as maize lethal necrosis (Wamaitha et al., 2018; Zaidi et al., 2020).

1.10 PERILOUS OF CRISPR-CAS SYSTEM

CRISPR-Cas being a highly efficient genome-editing tool exhibits the possibility of unintended editing. Literature survey showed various off-target consequences of CRISPR-Cas9-based genetic modifications in humans, rat, and mouse (Anderson et al., 2018; Aryal et al., 2018; AlJanahi et al., 2022). In addition, RNA-guided CRISPR-Cas-9 also showed off-targeted effects in mammals (Grunewald et al., 2019; Li et al., 2020). This off-targeted editing has potential of causing loss of gene functions by causing mutations due to wrong insertions influencing the gene sequence. It may cause incorrect results due to the breaking and binding at the nonspecific sequence sites (Cox et al., 2015; Xu et al., 2021). This may lead to mutations due to chromosomal rearrangements and DNA mismatches into protospacer

adjacent motif (PAM)-distal position of the single-guide RNA sequence (Ghosh et al., 2019). However, there are very few reports on these off-site editings using CRISPR-Cas9 in plants. Zhang and his coworkers recently exposed off-target genome editing in *Arabidopsis thaliana* (Zhang et al., 2018; Es et al., 2019). Despite rare and sometimes even undetectable events of off-targeted edits, researchers are focusing on reducing the potential risk of off-site edits by CRISPR-Cas9 (Lee et al., 2019; Matres et al., 2021; Nagamine and Ezura, 2022). Moreover, these off-target editing by CRISPR-Cas9, and *Agrobacterium*-mediated transformations have low off-targeted mutations to cause genetic variations as compared to other gene-mutating tools such as radiations (Li et al. 2019; Anderson et al., 2016). CRISPR-Cas9 technology has shown low rate of off-site mutations in plants. For instance, out of expected 4413 potential off-targets, it caused only four off-site edits. However, despite its low risk in plants, in case of humans and animals, careful screening is necessary to identify off-site mutations (Manghwar et al., 2020). These off-target mutations have potential risk of inheritance to descendants (Wang et al., 2015; Xue and Greene, 2021).

1.11 HUMANITARIAN CONCERNS FOR GENETIC MODIFICATIONS

A potential risk of genetic modifications in humans through the CRISPR-Cas system might be the misconduct of technology through illegal experimentations. In November 2018, a Chinese scientist, He Jiankui, has become widely recognized after he claimed of creating genetically edited humans toddlers Lulu and Nana (Gouw, 2019). He reported that he had changed a key gene related to HIV resistance. The amendment is probably surpassed directly to the offspring of youngsters. He recruited couples wherein the father became inflamed with the HIV and the mother not affected. In a communication on the international Summit on Human Genome Editing in Hong Kong, China, he stated that he desired to spare the toddlers the opportunity of getting HIV later in life. The approach can be used to lessen the HIV/AIDS sickness burden in lots of Africans, he also argued, in which the ones inflamed regularly face extreme discrimination (Ye et al., 2019). This is highly criticized by scientific community worldwide declaring their opposing statements on exploiting genomic modifications on human embryo. They added that any clinical trial of genome editing on human embryo for reproduction purpose is against ethical norms and laws (Ma et al., 2019; Wang et al., 2019). According to China's guideline playing with human genome to produce genetically edited

human embryos is prohibited, so work of He breaches Chinese regulatory norms (Li et al., 2019). This technology is not suitable for human germ-line gene editing until there are proper guidelines for handling its consequences.

1.12 ETHICAL ISSUES

Cloning, in the eyes of ethicists, breaches human dignity, robs individuals of their individuality, endangers humanity, and treats humans as mere tools (Ayanoğlu et al., 2020; Shafique, 2020). A human clone existence would resemble the donor to a specific degree and could not be better or even normal from a psychological or philosophical standpoint (Beriain and Marcos, 2018). Many films portraying human clones have been made in the last ten years and showed the clones as weak, oppressed creatures that are used primarily as organ donors (Hansen and Wöhlke, 2016). In addition, the surge in CRISPR/Cas9 usage highlights the potential benefits and ethical controversies of permanent genetic alterations (Jacinto et al., 2020). A ban on this sort of study may be imposed due to a variety of moral, ethical, political, and scientific issues. It is critical to pinpoint this issue and clarify the idea and encourage dialogue among all parties involved (Zhang et al., 2020). The creation and use of new technology poses ethical threats since it has an impact on the living situations of present-day or hypothetical moral beings, such as the non-human environment (Bartkowski et al., 2018). CRISPR/Cas has changed how we think about genetic engineering's moral implications. Since GMOs may be the product of natural evolution, the dangers of the novel organisms are comparable to those of their traditionally cultivated counterparts (Marangon et al., 2021; Francesco et al., 2022). Although some scientists are concerned about the unforeseen dangers of off-target mutations, the primary justification for opposing CRISPR currently appears to be out of date. It suggests that nature and its order are intrinsically beneficial and may be used as the individual and system levels. In relation to CRISPR/Cas, the term naturalness is used to describe the final product as well as the breeding process itself. Both supporters and detractors of CRISPR/Cas utilize naturalness as a key justification in public discussions. The ethical argument is intertwined with the analysis of contemporary sustainable agriculture and its socioecological implications, as well as a consideration of the repercussions and ethical considerations for biotechnology. The introduction of CRISPR/Cas genome editing raises a few ethical issues, including fundamental societal issues of justice and intergenerational fairness, in addition

to concerns about what outcomes and side effects are acceptable. The social discussion on the guiding principles for the creation and use of each new technology is necessary in society (Clapp and Ruder, 2020). It is necessary to develop international platform for multidisciplinary efforts to establish ethical guidelines for risk management (Jacinto et al., 2020).

KEYWORDS

- **genetic engineering**
- **cloning**
- **SCNT**
- **CRISPR-Cas9**
- **bioethics**

REFERENCES

Alberio, R.; Wolf, E. 25th Anniversary of Cloning by Somatic-Cell Nuclear Transfer: Nuclear Transfer and the Development of Genetically Modified/Gene Edited Livestock. *Reproduction* **2021,** *162*, no. 1, 59–68.

AlJanahi, A. A.; Lazzarotto, C. R.; Chen, S.; Shin, T. H.; Cordes, S.; Fan, X.; Jabara, I.; Zhou, Y.; Young, D. J.; Lee, B. C.; Yu, K. R. Prediction and Validation of Hematopoietic Stem and Progenitor Cell Off-Target Editing in Transplanted Rhesus Macaques. *Mol. Ther.* **2022,** *30*, no. 1, 209–222.

Anderson, J. E.; Michno, J. M.; Kono, T. J.; Stec, A. O.; Campbell, B. W.; Curtin, S. J.; Stupar, R. M. Genomic Variation and DNA Repair Associated with Soybean Transgenesis: A Comparison to Cultivars and Mutagenized Plants. *BMC Biotechnol*. **2016,** 16, 41.

Anderson, K. R.; Haeussler, M.; Watanabe, C.; Janakiraman, V.; Lund, J.; Modrusan, Z.; Stinson, J.; Bei, Q.; Buechler, A.; Yu, C.; Thamminana, S. R. CRISPR Off-Target Analysis in Genetically Engineered Rats and Mice. *Nat. Methods* **2018,** 15, 512.

Andersson, M.; Turesson, H.; Nicolia, A.; Falt, A. S.; Samuelsson, M.; Hofvander, P. Efficient Targeted Multiallelic Mutagenesis in Tetraploid Potato (*Solanum tuberosum*) by Transient CRISPR-Cas9 Expression in Protoplasts. *Plant Cell Rep*. **2017,** 36, 117–128.

Anzalone, A. V.; Gao, X. D.; Podracky, C. J.; Nelson, A. T.; Koblan, L. W.; Raguram, A.; Levy, J. M.; Mercer, J. A., Liu, D. R. Programmable Deletion, Replacement, Integration and Inversion of Large DNA Sequences with Twin Prime Editing. *Nat. Biotechnol.* **2022,** *40*, no. 5, 731–740.

Aryal, N. K.; Wasylishen, A. R.; Lozano, G. CRISPR/Cas9 Can Mediate High-Efficiency Off-Target Mutations in Mice *in vivo*. *Cell Death Dis*. **2018,** 9, 1099.

Ayanoğlu, F. B.; Elçin, A. E.; Elçin, Y. M. Bioethical Issues in Genome Editing by CRISPR-Cas9 Technology. *Turk. J. Biol.* **2020,** *44*, no. 2, 110–120.

Bartkowski, B.; Theesfeld, I.; Pirscher, F.; Timaeus, J. Snipping Around for Food: Economic, Ethical and Policy Implications of CRISPR/Cas Genome Editing. *Geoforum* **2018,** 96, 172–180.

Beriain, M. I. D.; Marcos del Cano, A. M. Gene Editing in Human Embryos. A Comment on the Ethical Issues Involved. In *The Ethics of Reproductive Genetics*; 2018; Springer: Cham, pp 173–187.

Buquicchio, F. A.; Satpathy, A. T. Interrogating Immune Cells and Cancer with CRISPR-Cas9. *Trends Immunol* **2021,** *42*, no. 5, 432–446.

Burkard, C.; Lillico, S. G.; Reid, E.; Jackson. B.; Mileham, A. J.; Ait-Ali, T.; Whitelaw, C. B.; Archibald, A. L. Precision Engineering for PRRSV Resistance in Pigs: Macrophages from Genome Edited Pigs Lacking CD163 SRCR5 Domain are Fully Resistant to both PRRSV Genotypes While Maintaining Biological Function. *PLoS Pathog.* **2017,** 13 e1006206.

Caballero, I.; Ringot-Destrez, B.; Si-Tahar, M.; Barbry, P.; Guillon, A.; Lantier, I.; Berri, M.; Chevaleyre, C.; Fleurot, I.; Barc, C. and Ramphal, R. Evidence of Early Increased Sialylation of Airway Mucins and Defective Mucociliary Clearance in CFTR-Deficient Piglets. *J. Cyst. Fibros.* **2021,** 20, 173–182.

Caudal, A.; Ren, L.; Tu, C.; Wu, J. C. Human Induced Pluripotent Stem Cells for Studying Mitochondrial Diseases in the Heart. *FEBS Lett.* **2022,** *596*, no. 14, 1735–1745.

Cengiz, N.; Wareham, C. S. Ethical Considerations in Xenotransplantation: A Review. *Curr. Opin. Organ Tran.* **2019,** *25*, no. 5, 483–488.

Chandrasekaran, J.; Brumin, M.; Wolf, D.; Leibman, D.; Klap, C.; Pearlsman, M.; Sherman, A.; Arazi, T.; Gal-On, A. Development of Broad Virus Resistance in Non-Transgenic Cucumber Using CRISPR/Cas9 Technology. *Mol. Plant Pathol*. **2016,** 17, 1140–1153.

Chen, C.; Gao, Y.; Liu, W.; Gao, S. Epigenetic Regulation of Cell Fate Transition: Learning from Early Embryo Development and Somatic Cell Reprogramming. *Biol. Reprod.* **2022,** *107*, no. 1, 183–195.

Chia, G.; Agudo, J.; Treff, N.; Sauer, M. V.; Billing, D.; Brown, B. D.; Baer, R.; Egli, D. Genomic Instability During Reprogramming by Nuclear Transfer is DNA Replication Dependent. *Nat. Cell. Biol*. **2017,** 19, 282–291.

Clapp, J.; Ruder, S. L. Precision Technologies for Agriculture: Digital Farming, Gene-Edited Crops, and the Politics of Sustainability. *Glob. Environ. Polit.* **2020,** *20*, no. 3, 49–69.

Cong, L.; Ran, F. A.; Cox, D.; Lin, S.; Barretto, R.; Habib, N.; Hsu, P. D.; Wu, X.; Jiang, W.; Marraffini, L. A.; Zhang, F. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **2013,** *339*, no. 6121, 819–823.

Cox, D. B. T.; Platt, R. J.; Zhang, F. Therapeutic Genome Editing: Prospects and Challenges. *Nat. Med*. **2015,** 21, 121.

Cyranoski, D. Japan Approves First Human-Animal Embryo Experiments. *Nature* **2019,** *26*, doi: 10.1038/d41586-019-02275-3.

Czernik, M.; Anzalone, D. A.; Palazzese, L.; Oikawa, M.; Loi, P. Somatic Cell Nuclear Transfer: Failures, Successes and the Challenges Ahead. *Int. J. Develop. Biol.* **2019,** *63*,123–130.

Doudna, J. A.; Charpentier, E. Genome Editing. The New Frontier of Genome Engineering with CRISPR-Cas9. Sci. **2014,** *346*, 1258096–1258106.

Dwivedi, S.; Purohit, P.; Vasudeva, A.; Kumar, M.; Agrawal, R.; Sheikh, N. A.; Misra, R.; Kishore, S.; Misra, S. Gene Therapy and Gene Editing in Healthcare. In: *Biotechnology in Healthcare*; Vol. 1, 2022; Academic Press, pp 147–175.

East-Seletsky, A.; O′Connell, M. R.; Knight, S. C.; Burstein, D.; Cate, J. H.; Tjian, R.; Doudna, J. A. Two Distinct RNase Activities of CRISPR-C2c2 Enable Guide-RNA Processing and RNA Detection. *Nat*ure **2016,** *538*, 270–273.

El-Mounadi, K.; Morales-Floriano, M. L.; Garcia-Ruiz, H. Principles, Applications, and Biosafety of Plant Genome Editing Using CRISPR-Cas9. *Front. Plant Sci.* **2020,** *11*, 56.

Eş, I.; Gavahian, M.; Marti-Quijal, F. J.; Lorenzo, J. M.; Khaneghah, A. M.; Tsatsanis, C.; Kampranis, S. C.; Barba, F. J. The Application of the CRISPRCas9 Genome Editing Machinery in Food and Agricultural Science: Current Status, Future Perspectives, and Associated Challenges. *Biotechnol*. *Adv.* **2019,** *37*, 410–421.

Fajac, I.; Sermet-Gaudelus, I. Emerging Medicines to Improve the Basic Defect in Cystic Fibrosis. *Expert Opin. Emerg. Drugs* **2022,** *27*, 229–239.

Francesco, M.; Stefania, T.; Carzedda, M.; Federico, N. Consumers' acceptance of genome edited food and the role of information. *Italian Rev. Agric. Econ*. **2022,** *76*, no. 3, 5–21.

Marangon, F.; Troiano, S.; Carzedda, M.; & Nassivera, F. Consumers' Acceptance of Genome Edited Food and the Role of Information. *Italian Rev. Agric. Econ.* **2021,** *76*, no. 3, 5–21.

Ghosh, D.; Venkataramani, P.; Nandi, S.; Bhattacharjee, S. CRISPR– Cas9 a Boon or Bane: The Bumpy Road Ahead to Cancer Therapeutics. *Cancer Cell Int*. **2019,** *19*, 12.

Ginn, S. L.; Amaya, A. K.; Alexander, I. E.; Edelstein, M.; Abedi, M. R. Gene Therapy Clinical Trials Worldwide to 2017: An Update. *J. Gene Med.* **2018,** *20*, no. 5, e3015.

Gouveia, C.; Huyser, C.; Egli, D.; Pepper, M. S. Lessons Learned from Somatic Cell Nuclear Transfer. *Int. J. Mole. Sci.* **2020,** *21*, no. 7, 2314.

Gouw, A. The CRISPR Advent of Lulu and Nana. *Theol. Sci.* **2019,** *17*, no. 1, 9–12.

Grunewald, J.; Zhou, R.; Garcia, S. P.; Iyer, S.; Lareau, C. A.; Aryee, M. J.; Joung, J. K. J. N. Transcriptome-Wide Off-Target RNA Editing Induced by CRISPR-Guided DNA Base Editors. *Nature* **2019,** 569, 433–437.

Hansen, S. L.; Wöhlke, S. Contrasting Medical Technology with Deprivation and Social Vulnerability. Lessons for the Ethical Debate on Cloning and Organ Transplantation through the Film Never Let Me Go (2010). *Nano Ethics* **2016,** *10*, no. 3, 245–256.

Hille, F.; Richter, H.; Wong, S. P.; Bratovic, M.; Ressel, S.; Charpentier, E. The biology of CRISPR-Cas: Backward and forward. *Cell* **2018,** *172*, 1239–1259.

Hovland, A. S.; Bhattacharya, D.; Azambuja, A. P.; Pramio, D.; Copeland, J.; Rothstein, M.; Simoes-Costa, M. Pluripotency Factors are Repurposed to Shape the Epigenomic Landscape of Neural Crest Cells. *Develop. Cell*. **2022,** *57*, no. 19, 2257–2272.

Iaffaldano, B.; Zhang, Y.; and Cornish, K. CRISPR/Cas9 Genome Editing of Rubber Producing Dandelion Taraxacum kok-saghyz Using Agrobacterium Rhizogenes Without Selection. *Ind. Crops Prod.* **2016,** *89*, 356–362. doi: 10.1016/j.indcrop.2016.05.029.

Jacinto, F. V.; Link, W.; Ferreira, B. I. CRISPR/Cas9-Mediated Genome Editing: From Basic Research to Translational Medicine. *J. Cell Mol. Med.* **2020,** *24*, no. 7, 3766–3778.

Jain, R.; Human Cloning: Ethical Issues and Legal Implications. *Indian J. Law Legal Res.* **2021,** *2*, no. 1, 35–46.

Jasim, Y. A. A Genetic Disaster of Human Cloning. *Cihan Univ. Erbil Sci. J.* **2022,** *6*, no. 1, 53–56.

Jinek, M.; Jiang, F.; Taylor, D. W.; Sternberg, S. H.; Kaya, E.; Ma, E.; Anders, C.; Hauer, M.; Zhou, K.; Lin, S.; Kaplan, M. Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation. *Science* **2014,** *343*, 1247997.

Joy, B. Why the Future Doesn't Need Us: Our Most Powerful 21st-Century Technologies-Robotics, Genetic Engineering, and Nanotech-are Threatening to Make Humans an Endangered Species. In *Emerging Technologies: Ethics, Law and Governance*; 2020; Routledge, pp 47–63.

Kaiser, G. G.; Mucci, N. C.; Gonzalez, V.; Sanchez, L.; Parron, J. A.; Perez, M. D.; Calvo, M.; Aller, J. F.; Hozbor, F. A.; Mutto, A. A. Detection of Recombinant Human Lactoferrin and Lysozyme Produced in a Bitransgenic Cow. *J. Dairy Sci.* **2017,** *100*, 1605–1617.

Kavarana, S.; Kwon, J. H.; Zilinskas, K.; Kang, L.; Turek, J. W.; Mohiuddin, M. M.; Rajab, T. K. Recent Advances in Porcine Cardiac Xenotransplantation: From Aortic Valve Replacement to Heart Transplantation. *Expert Rev. Cardiovasc. Ther.* **2022,** *20*, no. 8, 597–608.

Khorasani, A. B. S.; Yousefi, A. M.; Bashash, D. CAR NK Cell Therapy in Hematologic Malignancies and Solid Tumors; Obstacles and Strategies to Overcome the Challenges. *Int. Immunopharmacol*. **2022,** *110*, 109041.

Kiyani, S.; Munir, Y. The Issue of Human Cloning: A Review in Semitic Religions' Context. *Al-Milal J. Religion Thought* **2020,** *2*, no. 2, 31–49.

Klinger, B.; Schnieke, A. 25th Anniversary of Cloning by Somatic-Cell Nuclear Transfer Twenty-Five Years After Dolly: How Far Have We Come? *Reproduction* **2021,** *162*, no. 1, 1–10.

Knott, G. J.; Doudna, J. A. CRISPR-Cas Guides the Future of Genetic Engineering. *Science* **2018,** *361*, no. 6405, 866–869.

Kobashigawa, J. Pig-to-Human Heart Transplantation: Culmination of Technology and Ingenuity. *Ann. Thorac. Surg*. **2022,** *113*, no. 3, 711.

Kohn, D. B. Gene Therapy for Blood Diseases. *Curr. Opin. Biotechnol.* **2019,** *60*, 39–45.

Kwisda, K.; White, L.; Hubner, D. Ethical arguments concerning human-animal € chimera research: a systematic review. *BMC Med. Ethics* **2020,** 21, 1–4.

Lange, V.; Kappel, K. CRISPR Gene-Therapy: A Critical Review of Ethical Concerns and a Proposal for Public Decision-Making. *Can. J. Bioeth.* **2022,** *5*, no. 2, 78–87.

Lanigan, T. M.; Kopera, H. C.; Saunders, T. L. Principles of Genetic Engineering. *Genes* **2020,** *11*, no. 3, 291.

Lee, K.; Zhang, Y.; Kleinstiver, B. P.; Guo, J. A.; Aryee, M. J.; Miller, J.; Malzahn, A.; Zarecor, S.; Lawrence-Dill, C. J.; Joung, J. K.; Qi, Y. Activities and Specificities of CRISPR/Cas9 and Cas12a Nucleases for Targeted Mutagenesis in Maize. *Plant Biotechnol. J.* **2019,** *17*, 362–372.

Li, A.; Jia, S.; Yobi, A.; Ge, Z.; Sato, S. J.; Zhang, C.; Angelovici, R.; Clemente, T. E.; Holding, D. R. Editing of an Alpha-Kafirin Gene Family Increases, Digestibility and Protein Quality in Sorghum. *Plant Physiol*. **2018,** *177*, 1425–1438. doi: 10.1104/pp.18.00200.

Li, H.; Yang, Y.; Hong, W.; Huang, M.; Wu, M.; Zhao, X. Applications of Genome Editing Technology in the Targeted Therapy of Human Diseases: Mechanisms, Advances and Prospects. *Signal Transduct. Target. Ther.* **2020,** *5*, no. 1, 1–23.

Li, J. R.; Walker, S.; Nie, J. B.; Zhang, X. Q. Experiments that Led to the First Gene-Edited Babies: The Ethical Failings and the Urgent Need for Better Governance. *J. Zhejiang Univ. Sci. B*. **2019,** *20*, 32–38.

Linguanti, F.; Abenavoli, E. M.; Berti, V.; Lopci, E. Metabolic Imaging in B-Cell Lymphomas During CAR-T Cell Therapy. *Cancers* **2022,** *14*, no. 19, 4700.

Liu, Z.; Cai, Y.; Wang, Y.; Nie, Y.; Zhang, C.; Xu, Y.; Zhang, X.; Lu, Y.; Wang, Z.; Poo, M.; Sun, Q. Cloning of Macaque Monkeys by Somatic Cell Nuclear Transfer. *Cell* **2018,** *172*, 881–887. doi: 10.1016/j.cell.2018.01.020.

Long, C. R.; Westhusin, M. E.; Golding, M. C. Reshaping the Transcriptional Frontier: Epigenetics and Somatic Cell Nuclear Transfer. *Mol. Reprod. Dev.* **2014,** *81*, 183–193.

Ma, Y.; Zhang, L. and Qin, C. The First Genetically Gene-Edited Babies: It's "Irresponsible and Too Early". *Animal Model Exp. Med*. **2019,** *2*, 1–4.

Mader, S.; Windelspecht, M. *Human Biology*; McGraw-Hill Education, 2017.

Makarova, K. S.; Wolf, Y. I.; Alkhnbashi, O. S.; Costa, F.; Shah, S. A.; Saunders, S. J.; Barrangou, R.; Brouns, S. J.; Charpentier, E.; Haft, D. H.; Horvath, P. An Updated Evolutionary Classification of CRISPR-Cas Systems. *Nat. Rev. Microbiol*. **2015,** *13*, 722–736.

Malin, K.; Witkowska-Piłaszewicz, O.; Papis, K. The Many Problems of Somatic Cell Nuclear Transfer in Reproductive Cloning of Mammals. *Theriogenology* **2022,** *189*, 246–254.

Manghwar, H.; Li, B.; Ding, X.; Hussain, A.; Lindsey, K.; Zhang, X.; Jin, S. CRISPR/Cas Systems in Genome Editing: Methodologies and Tools for sgRNA Design, Off-Target Evaluation, and Strategies to Mitigate Off-Target Effects. *Adv. Sci*. **2020,** *7*, 1902312.

Mann, S. P.; Sun, R.; Hermere´n, G. Ethical Considerations in Crossing the Xenobarrier. *Methods Mol. Biol*. **2019,** *2005*, 175–193.

Mansouri, V.; Beheshtizadeh, N.; Gharibshahian, M.; Sabouri, L.; Varzandeh, M.; Rezaei, N. Recent Advances in Regenerative Medicine Strategies for Cancer Treatment. *Biomed. Pharmacother.* **2021,** *141*, 111875.

Marangon, F.; Troiano, S.; Carzedda, M.; Nassivera, F. Consumers' Acceptance of Genome Edited Food and the Role of Information. *Italian Rev. Agric. Econ.* **2021,** *76*, no. 3, 5–21.

Martinez-Lage, M.; Puig-Serra, P.; Menendez, P.; Torres-Ruiz, R.; Rodriguez-Perales, S. CRISPR/Cas9 for Cancer Therapy: Hopes and Challenges. *Biomedicines* **2018,** *6*, no. 4, 105.

Matoba, S.; Zhang, Y. Somatic Cell Nuclear Transfer Reprogramming: Mechanisms and Applications. *Cell Stem Cell* **2018,** *23*, no. 4, 471–485.

Matres, J. M.; Hilscher, J.; Datta, A.; Armario-Nájera, V.; Baysal, C.; He, W.; Huang, X.; Zhu, C.; Valizadeh-Kamran, R.; Trijatmiko, K. R.; Capell, T. Genome Editing in Cereal Crops: An Overview. *Transgenic Res.* **2021,** *30*, no. 4, 461–498.

McCleary, S.; Strong, R.; McCarthy, R. R.; Edwards, J. C.; Howes, E. L.; Stevens, L. M.; Sánchez-Cordón, P. J.; Núñez, A.; Watson, S.; Mileham, A. J.; Lillico, S. G. Substitution of Warthog NF-kappaB Motifs into RELA of Domestic Pigs is not Sufficient to Confer Resilience to African Swine Fever Virus. *Sci. Rep.* **2020,** *10*, 8951.

Moffett, H. F.; Harms, C. K.; Fitzpatrick, K. S.; Tooley, M. R.; Boonyaratankornkit, J.; Taylor, J. J. B Cells Engineered to Express Pathogen-Specific Antibodies Protect Against Infection. *Sci. Immunol*. **2019,** *4*, eaax0644.

Moretti, A.; Fonteyne, L.; Giesert, F.; Hoppmann, P.; Meier, A. B.; Bozoglu, T.; Baehr, A.; Schneider, C. M.; Sinnecker, D.; Klett, K.; Fröhlich, T. Somatic Gene Editing Ameliorates Skeletal and Cardiac Muscle Failure in Pig and Human Models of Duchenne muscular dystrophy. *Nat. Med*. **2020,** *26*, 207–214.

Nagamine, A.; Ezura, H. Genome Editing for Improving Crop Nutrition. *Front. Genome Edit.* **2022,** *4*, 850104.

Nasrullah, R.; Iqbal, K.; BiBi, S.; Muneer, S.; BiBi S.; Anwar. F. N.; Ethical Issues of Human Cloning. *J. Med. Sci.* **2020,** *40*, no. 3, 103.

Nelson, J. W.; Randolph, P. B.; Shen, S. P.; Everette, K. A.; Chen, P. J.; Anzalone, A. V.; An, M.; Newby, G. A.; Chen, J. C.; Hsu, A.; Liu, D. R. Engineered pegRNAs Improve Prime Editing Efficiency. *Nat. Biotechnol.* **2022,** *40*, no. 3, 402–410.

Niemann, H.; Lucas-Hahn, A. Somatic Cell Nuclear Transfer Cloning: Practical Applications and Current Legislation. *Reprod. Domest. Anim.* **2012,** *47*, 2–10.

Nishimasu, H.; Ran, F. A.; Hsu, P. D.; Konermann, S.; Shehata, S. I.; Dohmae, N.; Ishitani, R.; Zhang, F.; Nureki, O. Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell* **2014,** 156, 935–949.

Nwoye, L. Ethical Issues in Human Cloning. *Int. J. Humanit. Innov.* **2019,** *2*, no. 4, 125–128.

Okada, A.; Arndell, T.; Borisjuk, N.; Sharma, N.; Watson-Haigh, N. S.; Tucker, E. J.; Baumann, U.; Langridge, P.; Whitford, R. CRISPR/Cas9-Mediated Knockout of Ms1 Enables the Rapid Generation of Male-Sterile Hexaploid Wheat Lines for Use in Hybrid Seed Production. *Plant Biotechnol. J.* **2019,** *17*, 1905–1913. doi: 10.1111/pbi.13106.

Okano, H.; Morimoto, S. iPSC-Based Disease Modeling and Drug Discovery in Cardinal Neurodegenerative Disorders. *Cell Stem Cell* **2022,** *29*, no. 2, 189–208.

Oldani, G.; Lacotte, S.; Toso, C.; Chimeric Xenotransplantation. *Curr. Opin. Organ Trans.* **2019,** *24*, 543–546.

Panesar, P. S.; Marwaha, S. S. Biotechnology in Agriculture and Food Processing: Opportunities and Challenges. CRC Press: Boca Raton, Florida, 2013.

Pyott, D. E.; Sheehan, E.; Molnar, A. Engineering of CRISPR/Cas9- Mediated Potyvirus Resistance in Transgene-Free Arabidopsis Plants. *Mol. Plant Pathol*. **2016,** *17*, 1276–1288. doi: 10.1111/mpp.12417.

Reardon, S. First Pig-to-Human Heart Transplant: What Can Scientists Learn? *Nature* **2022,** *601*, no. 7893, 305–306.

Sarkar, B. K. Cloning and its Ethical Queries. *J. Posit. School Psychol.* **2022,** *6*, 8–17.

Schneider, J. W.; Oommen, S.; Qureshi, M. Y.; Goetsch, S. C.; Pease, D. R.; Sundsbak, R. S.; Guo, W.; Sun, M.; Sun, H.; Kuroyanagi, H.; Webster, D. A. Dysregulated Ribonucleoprotein Granules Promote Cardiomyopathy in RBM20 Gene-Edited Pigs. *Nat. Med.* **2020,** *26*, 1788–1800.

Shademan, B.; Masjedi, S.; Karamad, V.; Isazadeh, A.; Sogutlu, F.; Nourazarian, A. CRISPR Technology in Cancer Diagnosis and Treatment: Opportunities and Challenges. *Biochem. Genet.* **2022,** *60*, 1446–1470.

Shafique, S. Scientific and ethical implications of human and animal cloning. *Int. J. Sci. Technol. Soc*. **2020,** *8*, no. 1, 9–17.

Shimatani, Z.; Kashojiya, S.; Takayama, M.; Terada, R.; Arazoe, T.; Ishii, H.; Teramura, H.; Yamamoto, T.; Komatsu, H.; Miura, K.; Ezura, H. Targeted Base Editing in Rice and Tomato Using a CRISPR-Cas9 Cytidine Deaminase Fusion. *Nat. Biotechnol.* **2017,** *35*, 441.

Silverman, H.; Odonkor, P. N. Reevaluating the Ethical Issues in Porcine-to-Human Heart Xenotransplantation. *Hastings Center Rep.* **2022,** *52*, no. 5, 32–42.

Singh, B.; Mal, G.; Verma, V.; Tiwari, R.; Khan, M. I.; Mohapatra, R. K.; Mitra, S.; Alyami, S. A.; Emran, T. B.; Dhama, K.; Moni, M. A. Stem Cell Therapies and Benefaction of Somatic Cell Nuclear Transfer Cloning in COVID-19 era. *Stem Cell Res. Ther.* **2021,** *12*, no. 1, 1–16.

Sipe, C. J.; Kluesner, M. G.; Bingea, S. P.; Lahr, W. S.; Andrew, A. A.; Wang, M.; DeFeo, A. P.; Hinkel, T. L.; Laoharawee, K.; Wagner, J. E.; MacMillan, M. L. Correction of Fanconi

Anemia Mutations Using Digital Genome Engineering. *Int. J. Mol. Sci.* **2022,** *23*, no. 15, 8416.

Smith, C.; Abalde-Atristain, L.; He, C.; Brodsky, B. R.; Braunstein, E. M.; Chaudhari, P.; Jang, Y. Y.; Cheng, L. and Ye, Z. Efficient and Allele-Specific Genome Editing of Disease Loci in Human iPSCs. *Mol Ther.* **2015,** *23*, 570–577.

Sohn, S.; Buskirk, M. V.; Buckenmeyer, M. J.; Londono, R.; Faulk, D. Whole Organ Engineering: Approaches, Challenges, and Future Directions. *Appl. Sci.* **2020,** *10*, no. 12, 4277.

Sood, P.; Singh, R. K.; Prasad, M. Millets Genetic Engineering: The Progress Made and Prospects for the Future. *Plant Cell Tissue Organ Culture* **2019,** *137*, no. 3, 421–439.

Srirattana, K.; Kaneda, M.; Parnpai, R. Strategies to Improve the Efficiency of Somatic Cell Nuclear Transfer. *Int. J. Mol. Sci.* **2022,** *23*, no. 4, 1969.

Tanaka, M.; Gomi, K. Induction and Repression of Hydrolase Genes in *Aspergillus oryzae*. *Front. Microbiol.* **2021,** *12*, 677603.

Thomas, A.; Hawthorne, W. J.; Burlak, C. Xenotransplantation Literature Update. *Xenotransplantation* **2020,** *27*, e12582.

Toda, E.; Koiso, N.; Takebayashi, A.; Ichikawa, M.; Kiba, T.; Osakabe, K. An Efficient DNA- and Selectable-Marker-Free Genome-Editing System Using Zygotes in Rice. *Nat. Plants* **2019,** *5*, 363–368.

Tröder, S. E.; Zevnik, B. History of Genome Editing: From Meganucleases to CRISPR. *Lab. Anim.* **2022,** *56*, no. 1, 60–68.

Tuck, N.; Animals in Moral Limbo: How Literary Pigs May Help Lab-Generated Ones. *Animals* **2020,** *10*, 629.

Van-Eenennaam, A. L. Genetic Modification of Food Animals. *Curr. Opin. Biotechnol*. **2017,** *44*, 27–34.

Wakayama, S.; Kohda, T.; Obokata, H.; Tokoro, M.; Li, C.; Terashita, Y.; Mizutani, E.; Kishigami, S.; Ishino, F.; Wakayama, T. Successful Serial Recloning in the Mouse Over Multiple Generations. *Cell Stem Cell.* **2013,** *12*, 293–297.

Walters, M.; Magis, W.; DeWitt, M.; Wyman, S.; Vu, J. T.; Romero, Z. G.; Fernandez, B. C.; Kohn, D. B.; Boffelli, D.; Corn, J. E.; Martin, D. I. S852 *In Vivo* Selection for Corrected Beta-Globin Alleles After Crispr/Cas9 Editing in Human Sickle Hematopoietic Stem Cells Enhances Therapeutic Potential. *HemaSphere* **2019,** *3*, no. S1, 380.

Wamaitha, M. J.; Nigam, D.; Maina, S.; Stomeo, F.; Wangai, A.; Njuguna, J. N.; Holton, T. A.; Wanjala, B. W.; Wamalwa, M.; Lucas, T.; Djikeng, A. Metagenomic Analysis of Viruses Associated with Maize Lethal Necrosis in Kenya. *Virol. J.* **2018,** *15*, 90. doi: 10.1186/s12985-018-0999-2.

Wang, C.; Zhai, X.; Zhang, X.; Li, L.; Wang, J.; Liu, D. P. Gene-Edited Babies: Chinese Academy of Medical Sciences' Response and Action. *Lancet* **2019,** *393*, 25–26.

Wang, D. N.; Wang, Z. Q.; Jin, M.; Lin, M. T.; Wang, N. CRISPR/Cas9-Based Genome Editing for the Modification of Multiple Duplications that Cause Duchenne Muscular Dystrophy. *Gene Ther.* **2022b,** *29*, 730–737.

Wang, W.; He, W.; Ruan, Y.; Geng, Q. First Pig-to-Human Heart Transplantation. *Innovation* **2022a,** *3*, no. 2, 100223.

Wang, X.; Wang, Y.; Wu, X.; Wang, J.; Wang, Y.; Qiu, Z.; Chang, T.; Huang, H.; Lin, R. J.; Yee, J. K. Unbiased Detection of Off-Target Cleavage by CRISPR-Cas9 and TALENs Using Integrase-Defective Lentiviral Vectors. *Nat. Biotechnol*. **2015,** *33*, 175.

Wen, L.; Zhao, C.; Song, J.; Ma, L.; Ruan, J.; Xia, X.; Chen, Y. E.; Zhang, J.; Ma, P. X.; Xu, J.; CRISPR/Cas9-Mediated TERT Disruption in Cancer Cells. *Int. J. Mol. Sci.* **2020,** *21*, no. 2, 653.

Whitworth, K. M.; Rowland, R. R.; Ewen, C. L.; Trible, B. R.; Kerrigan, M. A.; Cino-Ozuna, A. G.; Samuel, M. S.; Lightner, J. E.; McLaren, D. G.; Mileham, A. J.; Wells, K. D. Gene-Edited Pigs are Protected from Porcine Reproductive and Respiratory Syndrome Virus. *Nat. Biotechnol.* **2016,** *34*, 20–22.

Wright, A. V.; Nuñez, J. K.; Doudna, J. A. Biology and applications of CRISPR systems: Harnessing Nature's Toolbox for Genome Engineering. *Cell* **2016,** *164*, 29–44.

Wurtzel, E. T.; Vickers, C. E.; Hanson, A. D.; Millar, A. H.; Cooper, M.; Voss-Fels, K. P.; Nikel, P. I.; Erb, T. J. Revolutionizing Agriculture with Synthetic Biology. *Nat. Plants* **2019,** *5*, 1207–1210.

Xu, C. F.; Chen, G. J.; Luo, Y. L.; Zhang, Y.; Zhao, G.; Lu, Z. D.; Czarna, A.; Gu, Z.; Wang, J. Rational Designs of in Vivo CRISPR-Cas Delivery Systems. *Adv. Drug Deliv. Rev.* **2021,** *168*, 3–29.

Xue, C.; Greene, E. C. DNA Repair Pathway Choices in CRISPR-Cas9-Mediated Genome Editing. *Trends Genet.* **2021,** *37*, no. 7, 639–656.

Ye, Z. J.; Zhang, X. Y.; Liang, J.; Tang, Y. The Challenges of Medical Ethics in China: Are Gene-Edited Babies Enough? *Sci. Eng. Ethics* **2019,** *26*, 123–125.

Yu, K. *Therapeutic Cloning: Its Applications, Advantages and Disadvantages*, International Conference on Biotechnology, Life Science and Medical Engineering (BLSME 2022), 2022.

Zaidi, S. S.; Mahas, A.; Vanderschuren, H.; Mahfouz, M. M. Engineering Crops of the Future: CRISPR Approaches to Develop Climate-Resilient and Disease-Resistant Plants. *Genome Biol.* **2020,** *21*, no. 1, 1–9.

Zhang, P.; Liu, P.; Dou, H.; Chen, L.; Chen, L.; Lin, L.; Tan, P.; Vajta, G.; Gao, J.; Du, Y.; Ma, R. Z. Handmade Cloned Transgenic Sheep Rich in Omega-3 Fatty Acids. *PLoS ONE* **2013,** *8*, e55941.

Zhang, D.; Hussain, A.; Manghwar, H.; Xie, K.; Xie, S.; Zhao, S.; Larkin, R. M.; Qing, P.; Jin, S.; Ding, F. Genome Editing with the CRISPR-Cas System: An Art, Ethics and Global Regulatory Perspective. *Plant Biotechnol. J.* **2020,** *18*, no. 8, 1651–1669.

Zhang, Q.; Xing, H. L.; Wang, Z. P.; Zhang, H. Y.; Yang, F.; Wang, X. C.; Chen, Q. J. Potential High-Frequency Off-Target Mutagenesis Induced by CRISPR/Cas9 in *Arabidopsis* and its Prevention. *Plant Mol. Biol*. **2018,** *96*, 445–456.

CHAPTER 2

Biotechnology and Its Applications

TOOBA MOMIN and MALABIKA BHOWMIK

Department of Biology, University of Dayton, Dayton, OH, USA

ABSTRACT

Biotechnology is a dynamic and multidisciplinary field that integrates biological principles with technological advancements to create innovative solutions for medicine, agriculture, and environmental sustainability. This chapter explores the broad applications of biotechnology, focusing on its role in drug development, diagnostics, and gene therapy, alongside its transformative impact on agriculture through genetically modified crops and improved crop yields. Additionally, biotechnology contributes significantly to the development of biofuels and bioremediation processes, reducing environmental pollution. Ethical considerations and regulatory frameworks are also discussed to provide a balanced overview of biotechnology's promise and challenges. This chapter aims to offer a comprehensive understanding of biotechnology's pivotal role in shaping modern industries and addressing global challenges.

2.1 INTRODUCTION

Biotechnology is a multidisciplinary field that involves the application of biological principles and techniques to develop new products, processes, and technologies for improving human health, agriculture, and the environment. It encompasses a broad range of applications, including genetic engineering, molecular biology, genomics, proteomics, and cell culture technology (Gupta et al., 2016). Biotechnology has a lot of significance in medicine, including drug discovery, diagnostics, gene therapy, and personalized medicine (Dugger et al., 2018). Through the creation of genetically modified crops with improved yields, nutritional content, and resource efficiency, it has also

Molecular Genetics, Structures, Mechanisms, and Functions, Volume 1: Principles of Gene Manipulation and Genomics. Tariq Ahmad Bhat, Abdul Rauf Shakoori, and Jameel M. Al-Khayri (Eds.)
© 2026 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

revolutionized agriculture (Dugger et al., 2018). Moreover, biotechnology has led to the development of biofuels and bioremediation technologies that can help reduce dependence on fossil fuels and mitigate environmental pollution (Youssef et al., 2009; Bachmann et al., 2014).

In this book chapter, we will explore the biotechnology and its applications in medicine, agriculture, and the medicine and diagnostics. Specifically, we will explore how biotechnology is being used to develop new drugs and therapies, improve crop yields and food quality, and advance our ability to diagnose and treat disease. We will also discuss the challenges and ethical considerations associated with biotechnology, as well as the regulatory frameworks that govern its use. By examining these topics, we hope to provide a comprehensive overview of the role that biotechnology is playing in shaping the world around us.

2.1.1 DEFINITION OF BIOTECHNOLOGY

The use of biological systems and living beings to create new technologies, goods, and services for the benefit of people is known as biotechnology. The field merges knowledge from biology and technology to create innovative processes and products with practical applications. Biotechnology encompasses a wide range of disciplines, including genetics, molecular biology, and biochemistry.

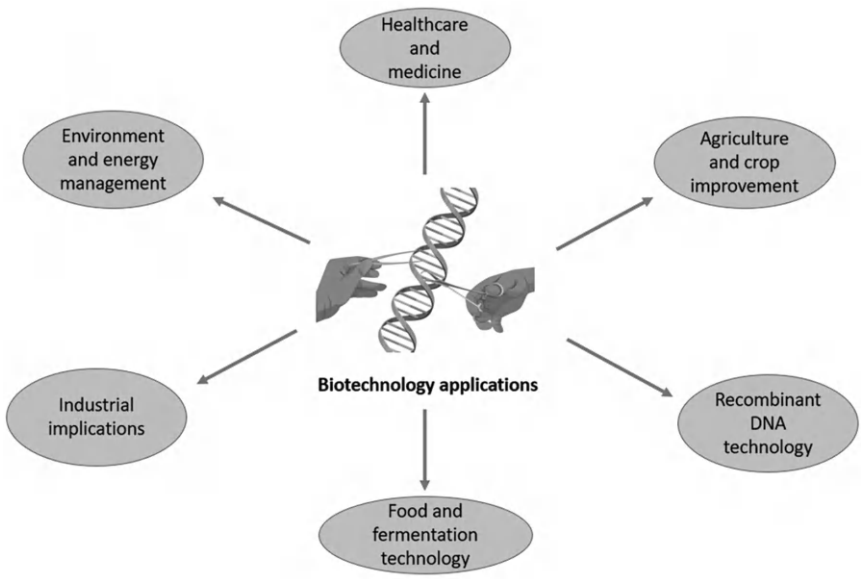


FIGURE 2.1 Various applications of biotechnology.

2.1.2 HISTORY OF BIOTECHNOLOGY

The roots of biotechnology date back to ancient civilizations, where fermentation processes were employed to produce food and beverages. The development of microbiology in the 19th century opened new avenues for research in biotechnology. Engineer Karl Ereky from Hungary introduced the word "biotechnology" in 1919 and predicted that biological processes would play a significant role in future industrial production (Ereky, 1919).

In the mid-20th century, advances in molecular biology and genetics led to the development of genetic engineering, which allowed for the manipulation of an organism's genetic material. This was a crucial turning point in the field of biotechnology and opened new avenues for research in areas such as medicine and agriculture.

2.2 MAJOR ADVANCES IN BIOTECHNOLOGY

Throughout the history of biotechnology, numerous significant advances have shaped the field into what it is today. Some of the most notable advancements include:

a. The 1953 work of Watson and Crick on the structure of DNA (Watson and Crick, 1953).
b. The 1970s saw the invention of recombinant DNA technology, which made it possible to alter an organism's genetic makeup (Cohenet al., 1973).
c. PCR (polymerase chain reaction), a method for amplifying particular DNA sequences for assessment, was developed (Mullis et al., 1992).
d. The creation of genetically modified organisms (GMOs) in the 1980s, which opened up new avenues for research in agriculture and medicine (Wöstemeyeret al., 1984).

2.3 APPLICATION OF BIOTECHNOLOGY IN AGRICULTURE

Introduction: Agriculture plays a vital role in the global economy by meeting the food demands of a rapidly growing population. However, with increasing population growth, there is a need for innovative ways to improve crop yields, reduce losses due to disease and pests, and increase the efficiency of the agriculture sector. Biotechnology offers a solution to these challenges,

as it has led to the development more disease, insect, and environmental stress-resistant genetically engineered crops, and new methods for efficient crop production. In this section, we will discuss the benefits and limitations of biotechnology in agriculture.

2.3.1 BENEFITS OF BIOTECHNOLOGY IN AGRICULTURE

Biotechnology has numerous benefits in agriculture, including:

2.3.1.1 INCREASED CROP YIELDS

Biotechnology has led to the development of genetically modified crops that are more resistant to disease, pests, and environmental stress, resulting in higher crop yields (Altman and Hasegawa, 2011). For example, genetically modified crops such as Bt cotton and Bt corn produce a toxin derived from the bacteria Bacillus thuringiensis (BT), which provides resistance against pests, leading to increased crop yields (Huang, 2002).

2.3.1.2 REDUCED LOSSES DUE TO DISEASE AND PESTS

By engineering crops with resistance to specific diseases and pests, biotechnology has helped to reduce losses in the agriculture sector. For example, Bt cotton and Bt corn have been effective in controlling pests, leading to reduced losses and increased profits for farmers (Huang, 2002).

2.3.1.3 INCREASED EFFICIENCY

Biotechnology has also led to the development of new methods for efficient crop production, such as tissue culture, which allows for the rapid production of large numbers of identical plants (Bhatia and Sharma, 2015). This technology has been particularly useful for propagating crops with desirable traits, such as resistance to disease and pests.

Improved Nutrient Content: Biotechnology has enabled the development of crops with enhanced nutrient content, providing essential vitamins and minerals to consumers and addressing malnutrition in some areas. For example, Golden Rice, a genetically modified rice, contains beta-carotene,

which the body converts into vitamin A, addressing vitamin A deficiency in developing countries (Potrykus, 2010).

2.3.2 LIMITATIONS OF BIOTECHNOLOGY IN AGRICULTURE

Despite the numerous benefits of biotechnology in agriculture, there are also limitations that need to be addressed. These include:

2.3.2.1 HIGH COST OF RESEARCH AND DEVELOPMENT

The development of new biotechnology products is a long and costly process, and the high cost of research and development can make it difficult for small farmers to access these technologies. This has led to concerns about the potential for biotechnology to exacerbate existing inequalities in the agriculture sector.

2.3.2.2 ETHICAL AND SOCIAL CONCERNS

There are also ethical and social concerns associated with the application of biotechnology in agriculture, such as the potential impacts on the environment and on the livelihoods of farmers (Thompson and Hannah, 2008). For example, the use of genetically modified crops may have unintended effects on nontarget species and may lead to the development of pesticide-resistant pests.

2.3.2.3 LACK OF REGULATION AND STANDARDIZATION

Lack of regulation and standardization in various industries can have several negative consequences. For example:

Consumer safety risks: Without regulations and standards, there is a higher risk of unsafe products entering the market and potentially harming consumers. This is particularly true for industries such as food and pharmaceuticals.

Unfair competition: Companies that do not follow regulations or industry standards may have an unfair advantage over those that do. This can create an uneven playing field and harm businesses that operate ethically.

Environmental damage: Lack of regulation and standardization can also lead to environmental damage. Industries that do not follow environmental regulations can harm natural resources and ecosystems, leading to negative impacts on the environment and wildlife.

Quality control issues: Without standards, there is no consistent benchmark for quality control in certain industries. This can lead to lower-quality products, inconsistent manufacturing processes, and a lack of accountability.

In short, lack of regulation and standardization can cause significant problems in various industries, impacting consumer safety, business competition, environmental sustainability, and product quality control. It is essential to have regulations and standards in place to ensure ethical and safe practices in business operations (Smythet al., 2014).

2.4 APPLICATION OF BIOTECHNOLOGY IN MEDICINE

2.4.1 GENE THERAPY

Utilizing genetic material to treat or prevent disease is known as gene therapy. Gene therapy has advanced significantly thanks to biotechnology, particularly in the creation of the vectors that deliver therapeutic genes to cells. Some of the most commonly used vectors include viruses, such as retroviruses, adenoviruses, and adeno-associated viruses. These vectors can be engineered to target specific cells or tissues and can be used to replace, add, or delete genes to correct genetic defects or treat diseases such as cancer, cystic fibrosis, and HIV/AIDS (Ginn et al., 2018).

2.4.2 ANTIBODY ENGINEERING

Antibody engineering involves the modification or creation of antibodies to target specific antigens, such as cancer cells, viruses, or bacteria. Biotechnology has enabled the production of monoclonal antibodies, which are identical copies of a single type of antibody. These antibodies can be engineered to have specific properties, such as enhanced binding affinity or reduced immunogenicity, and utilized to create treatments for a variety of illnesses, such as cancer, autoimmune disorders, and infectious diseases (Reichert, 2017).

2.4.3 PERSONALIZED MEDICINE

Personalized medicine is an approach to healthcare that tailors medical treatments and interventions to an individual's unique characteristics, such as their genetic makeup, environment, and lifestyle. The scope of personalized medicine includes the development of diagnostic tools, such as genetic tests, that can identify genetic variations associated with disease risk or drug response, and targeted therapies that are designed to treat specific molecular abnormalities that drive disease progression (Collins and Varmus, 2015). Personalized medicine presents several challenges, including ethical and regulatory concerns, data privacy and security issues, and limitations in data interpretation and clinical implementation (de Jong et al., 2021).

2.5 APPLICATION OF BIOTECHNOLOGY IN DIAGNOSIS

Diagnostics has undergone a revolution thanks to biotechnology, providing highly sensitive and specific methods for detecting disease-causing agents, genetic mutations, and other molecular markers of disease. Two major areas in which biotechnology has had a significant impact on diagnosis are molecular diagnostics and imaging/medical devices.

Molecular diagnostics involves the use of various biotechnological tools such as polymerase chain reaction (PCR), next-generation sequencing (NGS), microarrays, and CRISPR-Cas systems to identify specific biomolecules such as DNA, RNA, and proteins. These tools have revolutionized the diagnosis of infectious diseases, genetic disorders, and cancer, among other conditions. For instance, PCR is frequently used to detect viruses such as HIV, hepatitis, and influenza, while NGS has enabled the identification of disease-causing mutations in the human genome (Farkas, 2002).

In addition to molecular diagnostics, biotechnology has also enabled the development of advanced imaging and medical devices, which have greatly enhanced the accuracy and speed of diagnosis. For example, positron emission tomography (PET) and magnetic resonance imaging (MRI) are generally used for the detection and staging of cancer, while medical devices such as blood glucose monitors and continuous glucose monitoring systems have transformed the management of diabetes (Ganguly et al., 2010).

Overall, biotechnology has provided powerful tools for the diagnosis of a wide range of diseases, enabling early detection, accurate diagnosis, and personalized treatment. However, challenges remain in terms of cost, accessibility, and standardization of these technologies.

2.6 TRANSGENIC ANIMALS

Transgenic animals are animals that have been genetically modified to carry foreign genes. The process of creating transgenic animals involves introducing foreign genes into their genomes, either by pronuclear microinjection or by embryonic stem cell transfer. Transgenic animals have been used extensively in biomedical research to study disease mechanisms and develop new treatments. They have also been used in agriculture to improve animal production and food security. In medicine, transgenic animals have been used to produce human proteins and antibodies for therapeutic use, as well as to create animal models of human diseases. Some examples of transgenic animals include mice, cows, pigs, and sheep.

One example of the use of transgenic animals in medicine is the production of human insulin using transgenic bacteria and transgenic pigs. The transgenic bacteria are used to produce the human insulin protein, which is then extracted and purified for use in diabetes treatment. Transgenic pigs have been used to produce human blood clotting factors for the treatment of hemophilia. In agriculture, transgenic animals have been used to improve

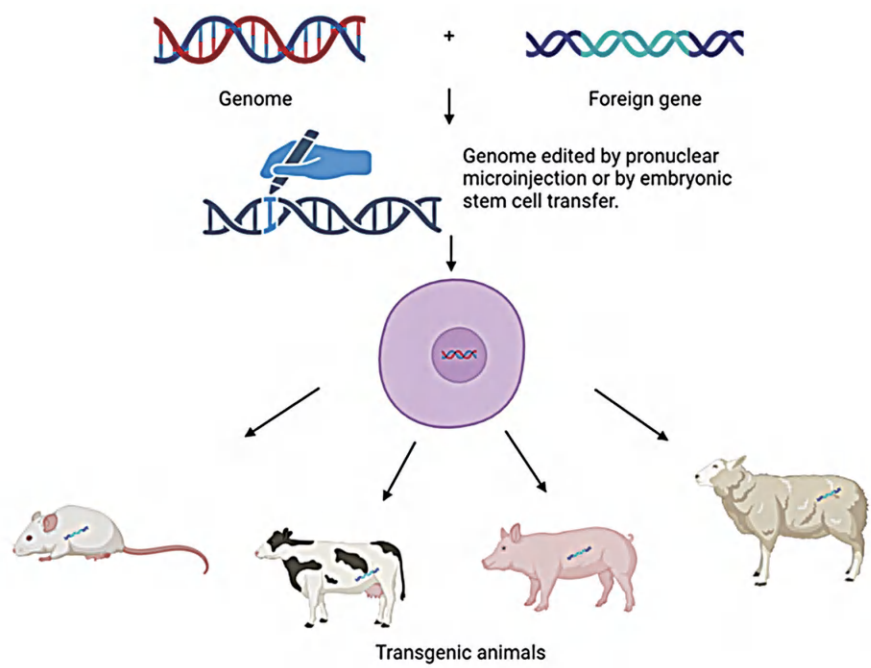


FIGURE 2.2 Flowchart of making a transgenic animal.

animal production by increasing milk production in cows, meat production in pigs, and wool production in sheep.

Houdebine LM (Houdebine, 2005) highlights the potential of transgenic animals in improving human health and animal production. The author discusses the applications of transgenic animals in producing recombinant proteins for therapeutic use, creating animal models of human diseases, and improving animal production. The author also notes the ethical and regulatory issues surrounding the use of transgenic animals and the need for responsible use and management of this technology.

2.7 CONCLUSION

In conclusion, biotechnology has revolutionized various fields including medicine, agriculture, and diagnostics. The development of novel techniques and technologies has enabled us to understand biological processes at a molecular level and manipulate them to improve human health and animal production. We can now make transgenic animals that can be used as models for researching human diseases, produce genetically modified crops that are resistant to pests and diseases, and create novel pharmaceuticals thanks to biotechnology. However, the application of biotechnology also brings up moral and security issues that call for careful regulation and oversight. Overall, biotechnology is a rapidly advancing field with immense potential to address some of the world's most pressing challenges.

KEYWORDS

- **biotechnology**
- **genetic engineering**
- **gene therapy**
- **agriculture**
- **bioremediation**
- **ethical considerations**
- **regulatory frameworks**

REFERENCES

Gupta, V.; Sengupta, M; Jaya Prakash, J.; and Tripathy, B. C. An Introduction to Biotechnology. *Basic Appl. Aspects Biotechnol.* **2016,** *23*, 1–21. doi: 10.1007/978-981-10-0875-7_1.

Dugger, S. A.; Platt, A.; Goldstein, D. B. Drug Development in the Era of Precision Medicine. *Nat. Rev. Drug. Discov.* **2018**, *17*, no. 3, 183–196.

Youssef, N.; Elshahed, M. S.; McInerney, M. J. Microbial Processes in Oil Fields: Culprits, Problems, and Opportunities. *Adv. Appl. Microbiol.* **2009,** *66*, 141–251.

Bachmann, R. T.; Johnson, A. C.; Edyvean, R. G. Biotechnology in the Petroleum Industry: An Overview. *Int. Biodeter. Biodegr.* **2014,** *86*, 225–237.

Ereky, K. *Biotechnology of Meat, Fat and Milk Production in an Agricultural Large-Scale Farm*; Budapest, Hungary, 1919.

Watson, J. D.; Crick, F. H. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953,** *171*, no. 4356, 737–738.

Cohen, S. N.; Chang, A. C.; Boyer, H. W.; Helling, R. B. Construction of Biologically Functional Bacterial Plasmids in Vitro. *Proc. Natl. Acad. Sci.* **1973,** *70*, no. 11, 3240–3244.

Mullis, K.; Faloona, F.; Scharf, S.; Saiki, R.; Horn, G.; Erlich, H. Specific Enzymatic Amplification of DNA in Vitro: the Polymerase Chain Reaction. *Biotechnol. Ser.* **1992,** 17–17.

Wöstemeyer, A.; Otten, L. A.; Schell, J. S. Sexual Transmission of T-Dna in Abnormal Tobacco Regenerants Transformed by Octopine and Nopaline Strains of Agrobacterium Tumefaciens. Mol. *General Genet.* **1984,** *194*, 500–507.

Altman, A.; Hasegawa, P. M. *Plant Biotechnology and Agriculture: Prospects for the 21st Century*; Academic Press, 2011.

Huang, J.; Hu, R.; Fan, C.; Pray, C. E.; Rozelle, S. Bt Cotton Benefits, Costs, and Impacts in China; 2002.

Bhatia, S.; Sharma, K. Chapter 11 - Micropropagation, In *Modern Applications of Plant Biotechnology in Pharmaceutical Sciences*, Bhatia, S., et al., Eds.; 2015; Academic Press: Boston, pp 361–368.

Potrykus, I. Lessons from the 'Humanitarian Golden RICE'PROJECT: Regulation Prevents Development of Public Good Genetically Engineered Crop Products. *New Biotechnol.* **2010,** *27*, no. 5, 466–472.

Thompson, P. B.; Hannah, W. Food and Agricultural Biotechnology: A Summary and Analysis of Ethical Concerns. *Adv. Biochem. Eng. Biotechnol.* **2008,** *111*, 229–264.

Smyth, S. J.; McDonald, J.; Falck-Zepeda, J. Investment, Regulation, and Uncertainty: Managing New Plant Breeding Techniques. *GM Crops Food* **2014,** *5*, no. 1, 44–57.

Ginn, S. L.; Amaya, A. K.; Alexander, I. E.; Edelstein, M.; Abedi, M. R. Gene Therapy Clinical Trials Worldwide to 2017: An Update. *J. Gene Med.* **2018,** *20*, no. 5, e3015.

Reichert, J. M. Antibodies to watch in 2017, in *MAbs*; 2017; Vol. 9, Taylor & Francis, pp 167–181.

Collins, F. S.; Varmus, H. A New Initiative on Precision Medicine. *N. Engl. J. Med.* **2015,** *372*, no. 9, 793–5.

de Jong, D. J.; Veldhuis, W. B.; Wessels, F. J.; de Vos, B.; Moeskops, P.; Kok, M. Towards Personalised Contrast Injection: Artificial-Intelligence-Derived Body Composition and Liver Enhancement in Computed Tomography. *J. Pers. Med.* **2021,** *11*, no. 3, 159.

Farkas, D. H. Molecular Diagnostics: The Best is Yet to Come. *Trends Mol. Med.* **2002,** *8*, no. 6, 245.

Ganguly, D.; Chakraborty, S.; Balitanas, M.; Kim, T. H. Medical Imaging: A Review. In *Security-Enriched Urban Computing and Smart Grid*; 2010; Springer: Berlin Heidelberg.

Houdebine, L. M. Use of Transgenic Animals to Improve Human Health and Animal Production. *Reprod. Domest. Anim.* **2005,** *40*, no. 4, 269–281.



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CHAPTER 3

Recombinant DNA Technology

TANVEER HUSSAIN and MUHAMMAD ISHTIAQ

Department of Botany, Mirpur University of Science and Technology (MUST), Mirpur, Pakistan

ABSTRACT

This study was focused on different methods of recombinant DNA technology (RDT). The formation of recombinant DNA was based on six steps. First step indicated DNA isolation from the desired organism cells. The cutting of the isolated DNA was made through enzyme. Third step involved after cutting of DNA from specific sites was amplification in PCR machine. After the amplification, DNA joining was completed by DNA ligase enzyme (fourth step). The recombinant deoxyribonucleic acid (rDNA) prepared in this step was added in the respective host (fifth step). The last step indicated the recombinant cell exclusion. The study was also depicted a virus produced by recombining DNA fragments using the RDT method is known as a recombinant virus. The role of gene therapy in recombinant DNA Technology was also under consideration. The gene therapy was expressed by two types such as germ line gene therapy (GLGT) and somatic gene editing (SGE). The Recombinant antibodies were also prepared by antibody technology (AT). The AT was used to bacterial expression of functional inhibitors and strategies for selecting genes from a catalog based on the phenotype of the translated protein. Hence, the rDNA preparation and their utilization in different purpose was compiled in this chapter.

3.1 INTRODUCTION

Recombinant DNA technology (RDT) is a hereditary designing method (HDM) that changes hereditary material external a life form to upgrade and

Molecular Genetics, Structures, Mechanisms, and Functions, Volume 1: Principles of Gene Manipulation and Genomics. Tariq Ahmad Bhat, Abdul Rauf Shakoori, and Jameel M. Al-Khayri (Eds.)
© 2026 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

want qualities in living organic entities or as their results. Innovations that are used to create fake DNA by the blend of various hereditary materials (DNA) from different sources are called recombinant DNA innovation. Genetic engineering is the common name for recombinant DNA technology (Saiki et al., 1988).

This approach entails adding DNA molecules from various sources with a desired DNA into an ideal carrier (Mingers and Leydesdorff, 2015). In this technique, put DNA strands from two separate species into a viral genome to create new genomic combos of enormous significance to research, healthcare, farming, and manufacturing. As a result, gene is main attention of experts. Geneticists' purpose is to collect, study, and modify genes (Smith and Welcox, 1970).

Although separating genetic material from a collection of cells is very simple, discovering a single copy inside this DNA specimen is equivalent to finding an object in a pile of straw. Assume that each living cell contains approximately two meters (6 feet) of DNA. As an outcome, a single tiny tissue specimen will contain many kilometers of DNA. But still, recombinant DNA innovation has empowered researchers to separate out one gene or other regions of DNA, establish its nucleotide process, analyze its result, modify it in a highly accurate manner, and then reinstall the new pattern into a life form (Kuhn, 2000).

In 1968, Swiss microbiologist Werner Arber found exclusion enzymes, which led to the development of artificial DNA technology. Placing the desired gene into the genetic code of the host is not as easy as it sounds. It comprises choosing the necessary gene for infusion into the host, then selecting the perfect vector inside which a gene merged and recombinant DNA created. Like a solution, the donor must have been injected with the recombinant DNA. As a result, the host must be introduced to the recombinant DNA. It must eventually remain in the host and be passed on to the offspring (Bosch, 2018).

Over the past few years, the importance of recombinant DNA has increased, and as genetic diseases become more common and agricultural land shrinks, rDNA will only increase in importance in the 21st century. rDNA will likely have a positive impact in the following areas: Better crops (resistant to drought and heat), synthetic vaccinations (hepatitis B), coagulation factors manufacturing, insulin synthesis, synthetic pharmaceutical manufacture, weedkiller plants, and germ line and somatic gene treatment (Knappik et al., 2000).

3.2 ROLE OF VECTORS IN RECOMBINANT DNA TECHNOLOGY

A virus produced by recombining DNA fragments using the RDT method is known as a recombinant virus. Vaccines against viruses and vectors for gene therapy could be made with this (Campbell and Stahl, 1998). During half of their cell replication, all viruses connect to their victims and inject their genetic traits into the recipient cell. As a result, extracting the viral genome and employing the virus like a carrier to deliver the remedial DNA has been identified as a viable technique for genome editing. Adeno-associated viruses, lentiviruses, the virus that causes herpes simplex, varicella virus, smallpox virus, and adeno-associated virus have all been employed in human gene therapy (Campbell, 2007).

3.3 ROLE OF GENE THERAPY IN RECOMBINANT DNA TECHNOLOGY

It has significant implications for the treatment of inherited and inherited disorders such as tumor and, potentially, acquired immune deficiency syndrome. There are two kinds of gene therapy: germ line gene therapy (GLGT) and somatic gene editing (SGE).

Germ line gene treatment involves the insertion of functional genes into germ cells (eggs and sperm). As a result, the alteration is heritable and will be handed down to future generations. This is hypothetically very helpful for managing genetic problems, but it is not currently being investigated for human application for a variety of ethical reasons.

Substantial quality treatment: The quality is presented exclusively in physical cells; however, it is not inherited as germline is not involved. Somatic gene therapy is further classified into two types: The first is where the functioning DNA is inserted constitutively alongside the genetic defect, so the transformed cell contains both the defective and normal (introduced) copies of the gene. This is referred to as augmentation therapy. The second method is specific gene transition, which employs homologous recombination to replace the indigenous gene with the functioning copy inserted (Gellert et al. 1968; Lobban and Kaiser, 1973).

3.4 RECOMBINANT ANTIBODIES

Advances in the disciplines of bacterial expression of functional inhibitors and strategies for selecting genes from a catalog based on the phenotype of the

translated protein have resulted in a breakthrough in antibody technology. To select *E. coli* recipient cells that produce specific antibody segments, phage display in conjunction with antibody gene libraries is commonly utilized nowadays. Such gene libraries are often derived from organic origins (for example, the liver of a vaccinated animal or blood cells from adult volunteers) or through GE. Later, it was utilized to generate libraries based on one or more antibody VH and VL gene segments that were varied using cassette mutation or other methods. These collections are usually impartial and can be utilized for any antigen. The great majority of ecological bioprocesses use naturally existing organisms (bacteria, fungus, etc.) to identify and filter production trash even before it enters the ecosystem. A biomonitoring programme including the employment of microorganisms is currently being implemented to clear up polluted air, ground routes, ponds, and waterways (Lomedico, 1982).

3.5 METHODS FOR RECOMBINANT DNA FORMATION

For the formation of recombinant DNA, there are three methods. nonbacterial transformation (NBT), phage introduction (PI), and transformation (TF) are their names. The following is a separate description of each protocol.

Transformation: The initial step in TF is to choose a fragment of DNA to insert into a carrier. The next procedure involves removing that section of DNA using an exclusion endonuclease enzyme and the DNA added into the carrier with an enzyme. The inclusion contains a specific protein that permits transgenic molecules to be identified. An antibiotic marker is frequently employed so that when a recipient cell lacking a carrier is exposed to that particular medication, that guest with the virus survives since it is resistant. In such a technique described as transformation, the antigen is deposited into a host cell. *E. Coli* itself is an example of a prospective host cell. To accept the invader DNA, the host cells must still be appropriately prepped. Selectable markers are able to identify altered hosts from wild-type hosts based on resistant bacteria, change in color, or any other trait. Various vectors have unique qualities that make them useful for a variety of applications. Asymmetrical cloning sites, size, and a high copy number are examples of attributes (Cederbaum et al., 1984).

Transformation that is not Bacterial: This is a technique that is nearly identical to the one described above. The sole distinction among the two is

that nonbacterial does not use bacteria like *E. coli* as a carrier. The DNA is incorporated intravenously into the nucleus of the cell being converted in inoculation. In biolistics, host cells are bombarded by high velocity micro-projectiles, such as gold or tungsten particles coated with DNA (Smith and Welcox, 1970).

Phage Introduction: The process of transfection known as phage introduction is similar to transformation, but instead of bacteria, phages are used. A vector is packaged in vitro in this case. Plaques of phage containing recombinants are produced using lambda or MI-3 phages in this method. Using a variety of selection techniques, the resulting recombinants can be distinguished from non-recombinants based on differences between them (Lobban and Kaiser, 1973).

3.6 STEPS IN FORMATION OF RECOMBINANT DNA (RDNA)

3.6.1 DNA ISOLATION

a. DNA is separated in its purest form, meaning it is free of all other polymers.
b. The first stage in recombinant DNA is to remove the necessary DNA in its truest form, that is, free of unwanted components.
c. Since DNA intermingle with other macromolecules within the cell membrane, it must be isolated and refined using enzymes such as lysozymes, cellulase, chitinase, ribonuclease, and proteases.
d. Different enzymes or therapies may be capable of eliminating additional macromolecules. Because of the presence of ethanol, the DNA finally precipitates out as stapling. The pure DNA is then rolled out (Knappik et al., 2000).

3.6.2 CUTTING OF DNA/LIMITATION CATALYST PROCESSING

a. Restriction enzymes are critical for this stage. It aids in determining where a specific gene is inserted into a carrier genome. The aforementioned reaction is referred to as restriction endonucleases digestions.

b. They involve digesting clean DNA with a restriction enzyme of selection in enzyme-specific conditions.
c. The "Agarose Gel Electrophoresis" method demonstrates the advancement of reduced enzyme hydrolysis.
d. This procedure involves running the DNA through an agarose gel. When a voltage is given to the negative electrode, DNA that is charged negatively runs to the positive electrode and is split into distinct lengths. This enables for the splitting and plucking of degraded DNA pieces.
e. The vector DNA is processed using the same manner.

3.6.3 DNA AMPLIFICATION

a. Gene duplicates are multiplied using PCR or polymerase chain reaction. Later the wanted gene of concern has been cut with restraint enzymes; it basically entails dividing a single copy of DNA into multiple copies.
b. According to Siegel (2002), it enables the amplification of a sole replica or a few copies of DNA into thousands or else loads of replicas.
c. The following components are utilized in "thermal cycler"-based PCR reactions:

    1. Template: DNA that requires amplification.
    2. Primers: Oligonucleotides are oligonucleotides that are matching to an area of DNA and are produced chemically. They are tiny.
    3. Enzyme: Polymerase of DNA.
    4. Nucleotides: To make the primers longer, the enzyme is needed.

d. Use PCR to amplify the cut DNA segments, which can then be fused with the chopped plasmid.

3.6.4 JOINING DNA

a. At this phase, the carrier and a portion of DNA are merged. It is conducted with the assistance of the enzyme DNA ligase.
b. The pure DNA and the vector of significance are chopped using the same restriction enzyme.

c. This produces the chopped DNA fragment and the cut carrier, which are now both free.
d. Occlusion is the process of joining these two pieces by using enzyme "DNA ligase."
e. The resultant DNA molecule is a cross between the target and vector DNA molecules. In genetics, recombination refers to the joining of distinct DNA strands (Kayser and Muller, 2004).

3.6.5 INCLUSION OF RDNA INTO A HOST

a. Here, rDNA is added to the beneficiary host cell, and the whole cycle is called change. Under favorable settings, the rDNA increases and shows as produced protein following insertion.
b. The new DNA is then moved into a beneficiary host cell, most ordinarily a bacterial cell, in this stage. "Transformation" is the name of this method.
c. Bacterial cells are difficult to accept as DNA. Consequently, treatments are administered to make them "capable" of accepting new DNA. Different procedures such as electroporation, Ca^{++} ion therapy, and thermal shock were used for inclusion of rDNA into a host (Brooke et al., 2008).

3.6.6 RECOMBINANT CELL EXCLUSION

a. The transformation process results in a mixed population of transformed and untransformed host cells.
b. Only the changed host cells are sifted during the determination method.
c. Recombinant cells and non-recombinant cells are distinguished by the plasmid vector's marker gene.
d. The PBR322 plasmid vector, for instance, contains two marker genes (the gene for resistance to ampicillin and the gene for resistance to tetracycline). Recombinant cells become ampicillin-sensitive as a result of the removal of the Ampicillin resistance gene from the plasmid by pst1 RE (Rosch, 1999).

KEYWORDS

- **recombinant DNA**
- **amplification**
- **gene therapy**
- **germ line gene therapy**
- **somatic gene editing**
- **translated protein**

REFERENCES

Mingers, J.; Leydesdorff, L. A Review of Theory and Practice in Scientometrics. *Eur. J. Oper. Res.* **2015,** *246*, no. 1, 1–19.

Bosch, G. Train PhD Students to be Thinkers not Just Specialists. *Nature* **2018,** *554*, 277.

Saiki, R. K.; Gelfand, D. H.; Stoffel, S.; Scharf, S. J.; Higuchi, R.; Horn, G. T.; Mullis, K. B.; Erlich, H. A. Primer-Directed Enzymatic Amplification of DNA with a Thermostable DNA polymerase. *Science* **1988,** *239***,** 487–491.

Kuhn, T. S. What is Scientific Revolution? In *The Road Since Structure: Philosophical Essays, 1970-1993, with an Autobiographical Interview*; Conant, J., Haugeland, J., Eds.; The University of Chicago Press: Chicago, 2000; p 13.

Campbell, A. Phage Integration and Chromosome Structure. A Personal History. *Ann. Rev. Genet.* **2007,** *41*, 1–11.

Campbell, A.; Stahl, F. W. Alfred D. Hershey. *Ann. Rev. Genet.* **1998,** *32*, no. 1, 1–6.

Gellert, M.; Little, J. W.; Oshinsky, C. K.; Zimmerman, S. B. Joining of DNA Strands by DNA Ligase of *E. Coli*. In *Cold Spring Harbor Symposia on Quantitative Biology*; 1968; Vol. 33, pp 21–26.

Cederbaum, S. D.; Fareed, G. C.; Lovett, M. A.; Shapiro, L. J. Recombinant DNA in Medicine. *Western J. Med.* **1984,** *141*, no. 2, 210–22.

Knappik, A.; Ge, L.; Honegger, A.; Pack, P.; Fischer, M.; Wellnhofer, G.; Hoess, A.; WoÈlle, J.; Plückthun, A.; Virnekäs, B. Full Synthetic Human Combinatorial Antibody Libraries (HuCAL) Based on Modular Consensus Frameworks and CDRS Randomized with Trinucleotides. *J. Mol. Biol.* **2000,** *296*, no. 1, 57–86.

Siegel, D. L. Recombinant Monoclonal Antibody Technology. *Transfus. Clin. Biol.* **2002,** *9*, no. 1, 15–22.

Brooke, B. S.; Jabashi, J. P.; Judge, D. P.; Patel, N.; Loeys, B.; Dietz, Harry C., I. Angiotensin II Blockade and Aortic-Root Dilation in Marfan's Syndrome. *N. Engl. J. Med.* **2008,** *358*, no. 26, 2787–2795.

Rosch, E. Principles of Categorization. In *Concepts of Core Readings*; Margolis, E., Laurence, S. Eds.; The MIT Press: Cambridge, Massachusetts, London, England, 1999.

Lomedico, P. T. Use of Recombinant DNA Technology to Program Eukaryotic Cells to Synthesize Rat Proinsulin: A Rapid Expression Assay for Cloned Genes. *Proc. Natl. Acad. Sci. USA* **1982,** *79*, 5798–5802.

Kayser, O.; Muller, R. H. A Primer on Pharmaceutical Biotechnology and Industrial Applications. In *Pharmaceutical Biotechnology, Drug Discovery and Clinical Applications*, 7th Ed.; *Wiley-VCH-Weinheim*, 2004, 1–8.

Lobban, P.; Kaiser, A. Enzymatic End-to End Joining of DNA Molecules. *J. Mol. Biol.* **1973,** *78*, no. 3, 453–471.

Smith, H. O.; Welcox, K. W. A Restriction Enzyme from Hemophilus Influenzae: I. Purification and General Properties. *J. Mol. Biol.* **1970,** *51*, no. 2, 379–391.



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CHAPTER 4

Construction and Screening of DNA Libraries and the Molecular Analysis of DNA, RNA, and Protein

MARIA BUTT, MAHNOOR IMRAN, TANZEELA REHMAN, and UZMA QAISAR

School of Biological Sciences, University of the Punjab, Lahore, Pakistan

ABSTRACT

In the molecular biology, with the help of powerful technologies, the isolation of any part of the DNA from any organism's genome and the transcribed gene after the process of reverse transcription produces complementary DNA is easily possible to construct a library. The cloning is the entrance of the DNA piece into a vector that is able to replicate in a microorganism, which then leads to the formation of huge quantity of the DNA parts for different physical techniques and the biological experiments. After having an idea of the exact position in the particular genome and from which a specific DNA part was obtained, that part has the characteristics of a DNA marker. These are required for genetic and physical mapping of an organism's genome. They are important for recognizing the genetic disorders. The vectors and library of clones are being used to clone different markers. For this purpose, vectors that are lambda based and gene libraries are most often in use. The variation of λ-based vectors having mix characteristics of any cloning vectors, for example, plasmids, M13 vectors, and P1 phages have produced. The usage is allowed to a particular technique, the formation of general genomic libraries (having the whole genome DNA parts) or particular genomic libraries (having only a specific subset of genomic DNA parts). For very high throughput techniques of DNA, RNA, and protein various

Molecular Genetics, Structures, Mechanisms, and Functions, Volume 1: Principles of Gene Manipulation and Genomics. Tariq Ahmad Bhat, Abdul Rauf Shakoori, and Jameel M. Al-Khayri (Eds.)
© 2026 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

molecular biology techniques are used and these are helpful in determining different roles of biological molecules in vitro as well as in vivo conditions.

4.1 INTRODUCTION

In molecular biology, "cloning" is defined as the introduction of gene of interest into a biological carrier, that is, the specific vector to produce there combinant DNA. The vector carrying gene of interest is now called as the recombinant vector in order to differentiate it from the host vector, which does not have the foreign fragment. After insertion of DNA fragment, recombinant DNA is allowed to transfer into an expression system either prokaryotic or eukaryotic where it can produce multiple copies accordingly. The important part in the study is the DNA fragment that has been inserted into the vector from other target organism (Wu, et al. 2006). DNA fragment has either a fragment of a gene or any sequence of the DNA that has not been characterized yet. The foreign gene can directly be obtained from the DNA and can also obtained by the reverse transcription of RNA with help of enzyme reverse transcriptase (Blattner, et al. 1977). The aim of cloning is to get a foreign piece of DNA or gene of interest in a larger quantity for molecular analysis at huge scale. The techniques and vectors used for cloning vary according to particular methods and objective of the study. This chapter highlights the study of using different vectors and building of genomic libraries that have an important part in molecular analysis of DNA, RNA, and protein (Tan and Yiap 2009).

A library is actually a group of recombinant vectors having isolated DNA parts either from the same organism or different organisms using biological techniques as shown in Figure 4.1. It can be either a small fragment that contains a specific information or the whole genome. Some of the fragments are obtained by sonication of whole genomes, whereas others are obtained after restriction digestion with specific restriction enzymes using their respective buffers. The restricted fragments may be having some specific repeats isolated from target genome. The genomic libraries contain complete information about the genomic fragments which they harbor (Yanisch-Perron, et al. 1985). A group or sets of complementary DNA (cDNA) libraries can also be prepared from RNA in a cell by the process of reverse transcription in vitro conditions with specific enzymes, that is, reverse transcriptase. Generally, cDNA library is prepared from messenger RNA (mRNA) (Murray and Murray 1974). Another library named expression library that is made in an expression vector in which the genes are directly expressed into the host cells. The small fragments can be inserted into a specific vector (plasmid,

phage, or cosmid) containing an insert size of up to 45 kbp (Kim, et al. 1992). Whereas the DNA libraries of larger insert are cloned in bacterial artificial chromosome (BAC), bacteriophage P1-derived artificial chromosome (PAC), plasmid-based clone (PBC), plant transformation competent binary BAC (BIBAC), and yeast artificial chromosome (YAC) vectors, having an insert size of 100 kb or more (Jenkins and Pennington 2001).

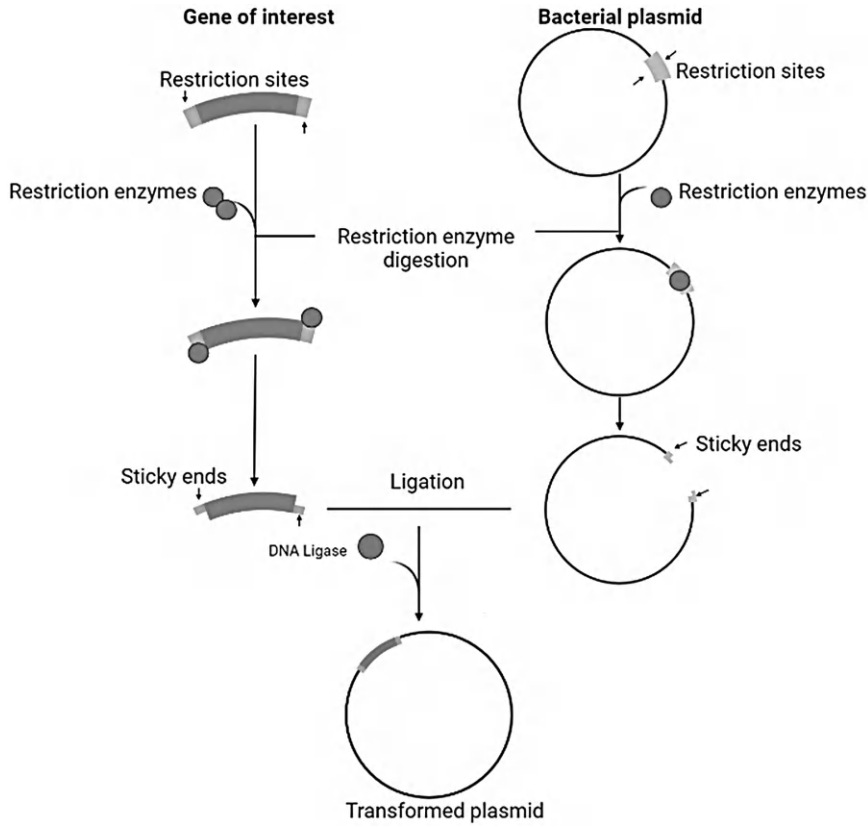


FIGURE 4.1 Gene cloning. ⏎

Source: Adapted from Eun, 1996

4.2 CLONING VECTORS

Plasmids are most commonly used cloning vectors of bacteria. It is a small circular genetic material that can replicate autonomously inside the host cell and can be easily transferred to the other host cells. These vectors have a special cloning site where foreign fragment can be inserted easily that is

called a multiple cloning site (MCS)(Henne, et al. 1999). This site has many restriction sites to which foreign DNA piece can be ligated as shown in Figure 4.2. Plasmids are being used as cloning vectors to insert outside fragment into both bacterial cells and yeast cells. After ligation of foreign DNA into the vector that is the process of cloning then the resultant recombinant plasmids are transferred into the bacteria. These plasmids have some selectable marker genes that are usually antibiotic resistance genes that provide an ability to the bacteria to grow and divide in a selection medium having that specific antibiotic (MacNeil, et al. 2001).

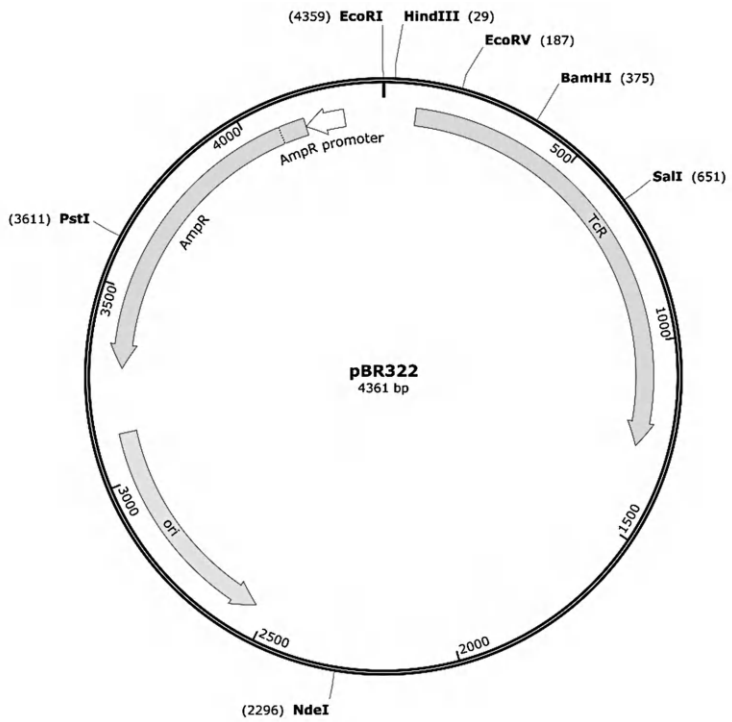


FIGURE 4.2 A schematic representation of the pBR322 plasmid. ⏎

Source: Adapted from Kim et al. (2009).

Plasmids are transferred to the competent bacterial cells through the process of transformation. Competent cells are the bacterial cells which have porous cell walls to accept foreign naked DNA (plasmid). These bacterial cells are made artificially competent through treatment with various chemicals. After the transformation, cells are grown on the selection media where only those clones survive that are having plasmid with selection marker. Usually,

antibiotics are added to the bacterial growth media to make it selective but sometimes nutrient deficient media and chromogenic media can also be used. Some vectors contain other reporter genes also known as marker genes to help in selecting the plasmids having cloned fragment. Cloning bacteria containing recombinant plasmids can divide vigorously multiplying plasmids in large quantities. Cells are collected and the plasmid of interest can be separated by using different protocols of plasmid isolation (Snapper, et al. 1990).

The bacteriophages that are used as cloning vectors include M13 phage and λ phage. They are having a specific limit to enter foreign DNA to be inserted into vector DNA with the size of approximately 53 kb. The phage vectors may contain some genes that are nonessential and they need to be removed. λ phage vectors have two types that are insertion and replacement vector (Lathe, et al. 1984).

The insertion vectors have a specific site called as cleavage site in which foreign fragment with a range of approximately 5–11 kb can be ligated. But in the case of replacement vectors, the flanking region are present in the cleavage site having the genes that do not need for the lytic cycle and that part should be excluded and exchanged with the DNA as an insert in the cloning procedure and large size of about 8–24 kb can be inserted (Chauthaiwale, et al. 1992).

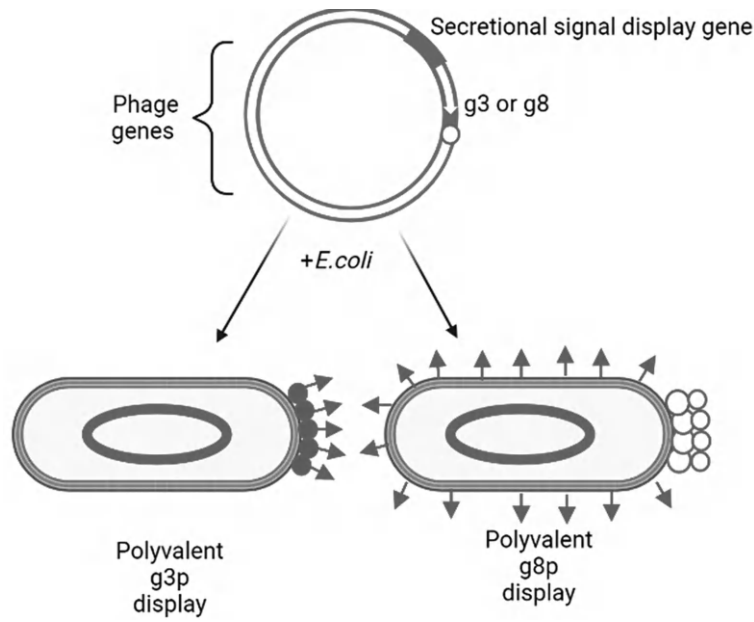


FIGURE 4.3 Phage vectors. ⏎

Source: Adapted from Larocca et al. (2002).

Phage vectors is having a complete phage genome mostly M13 phage. In this phage, the DNA that is inserted encodes the protein as shown in the Figure 4.3. The remaining is not changeable and able to provide different gene products for the phage life cycle. *E. coli* is used for the easy growth of phage, the function of the fusion protein in the life cycle of phage should be properly balanced and it depends on type of fusions prepared by using these vectors (Norrander, et al. 1983).

One of the interesting additions in these vectors is the cosmid as shown in Figure 4.4. These plasmids having specific characteristics like they are circular, double-stranded DNA having an OriC, also particular marker gene and cloning site which is having a specific portion for the λ cos site.

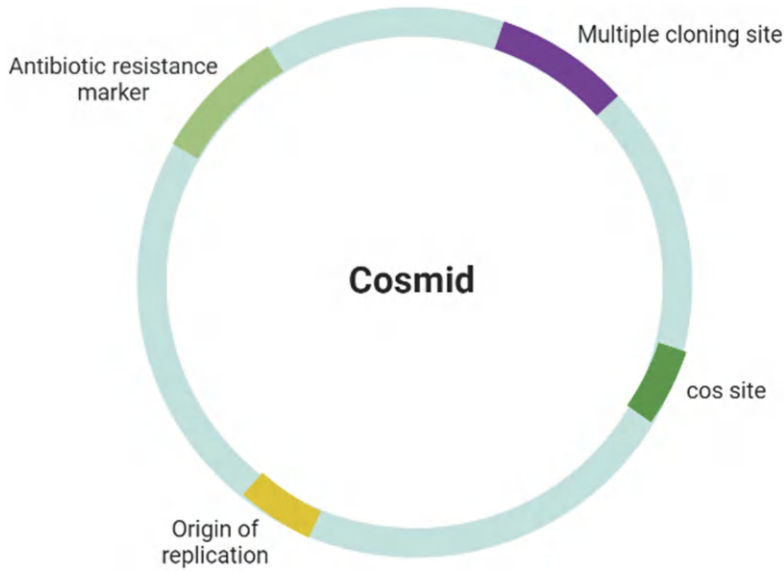


FIGURE 4.4 Schematic representation of Cosmid Vector. ⏎

Source: Adapted from Wahl et al. (1987).

The genome of λ is a straight DNA molecule that is 40–45 kb long. They are single stranded having complementary overhangs at opposite ends and these overhangs originate during the process of packaging by splitting of the cos site (Ish-Horowicz and Burke 1981). Once they are entered into the bacteria, both the ends find each other and are joined with the help of the enzymes of bacteria. In bacteria, this circular λ genome is copied and the most crucial enzyme DNA polymerase keeps on moving in a circle. Hence, a long

copy is produced having many λ genomes that have been combined together. The particular enzymes of λ-genome break this concatemer at the sites of cos and produce very linear λ genomes having a single-stranded overhang that can be easily packed into the phage bodies (Saito and Stark 1986).

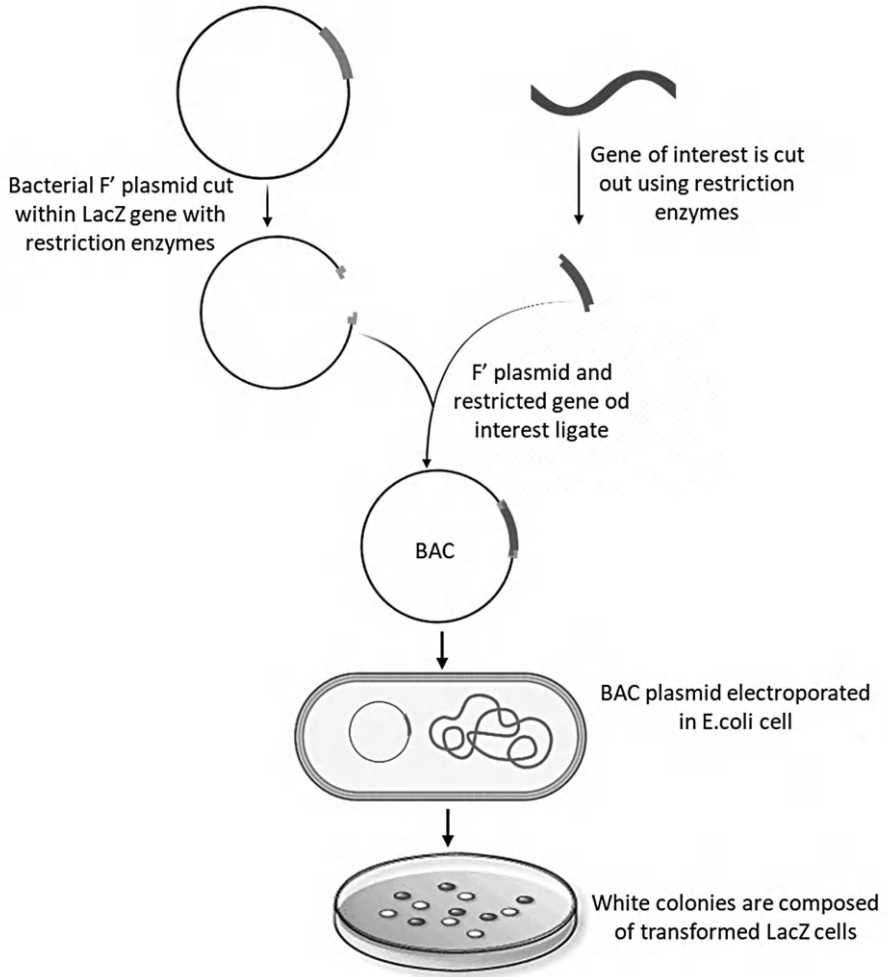


FIGURE 4.5 Gene cloning in BAC vector. ⏎

Source: Adapted from Berry et al. (2003).

BAC vectors were designed to have inserts of greater than 300 kb, although 100–200 kb is recommended due to the limitations of making large DNAs and then transferring them into the bacterial cell. One more

same cloning vector known as PAC is been made from the bacteriophage P1 as shown in Figure 4.5(Berry, et al. 2003). BAC libraries with the characteristics of having large insert capacity have been proven as the initial point for sequencing different model organisms having large genomic size that includes rice, human, and mouse. The DNA fragments that have been constructed into the most BAC libraries have been produced by the partial restriction digestion. For any given organism, many libraries are constructed using different fragments produced from different restriction enzymes. Those regions like centromeres and telomeres that actually lack restriction sites are most rarely present in these libraries. Cloning of the large DNA fragments is difficult and as well as challenging for producing BAC libraries from randomly generated DNA fragments.

YACs are prepared by genetic engineering and are obtained from yeast cells DNA mostly *Saccharomyces cerevisiae* and after that it is entered in a plasmid of bacteria. Typically, a huge DNA insert approximately 100–1000 kb can be ligated into the vector and physical mapping can be done by a process known as chromosome walking. This process was utilized for the Human Genome Project (Cossons, et al. 1997).

The main constituents of YAC are the replication origin sequences (ARS), telomeres from *S. cerevisiae* and centromere as shown in Figure 4.6. In addition, some selectable marker genes and visible marker gene helps in selecting only the transformed cells of yeast. In the absence of these additional genes, the chromosome will not stabilize itself during the cell replication and not be able to differentiate from the colonies that are without the vector. This vector is derived from a circular DNA plasmid, which is then chopped in straight DNA molecule with the help of restriction enzymes. Another enzyme, DNA ligase to join a DNA fragment or gene of choice into a linear DNA that will form one large, circular fragment of DNA (Riley, et al. 1990).

The main steps that are included in linearizing YACs are: the addition of a selectable marker gene into the plasmid vector that will help in differentiating the colonies having or not having the marker gene. Another gene known as the gene for antibiotic resistance allows YAC vector to divide and selected for *E. coli* cells by reverting the potential of mutant *E. coli* cells to produce leucine in the growth medium with all necessary components. The TRP1 and URA3 genes are the other selectable marker genes. The cloning site for the YAC vector having outside DNA is present within the gene SUP4. Its function is the compensation of any sudden change in the yeast cell that is responsible for gathering of red pigment. The cells of the

host are red in color and those that are transformed with YAC vector only produce the colorless colonies. The cloning of fragment of DNA in a YAC vector results in an inactivation by insertion of gene which will restore red color. That's why the colonies that are having a outside DNA fragment are red in color (Jakobovits 1994).

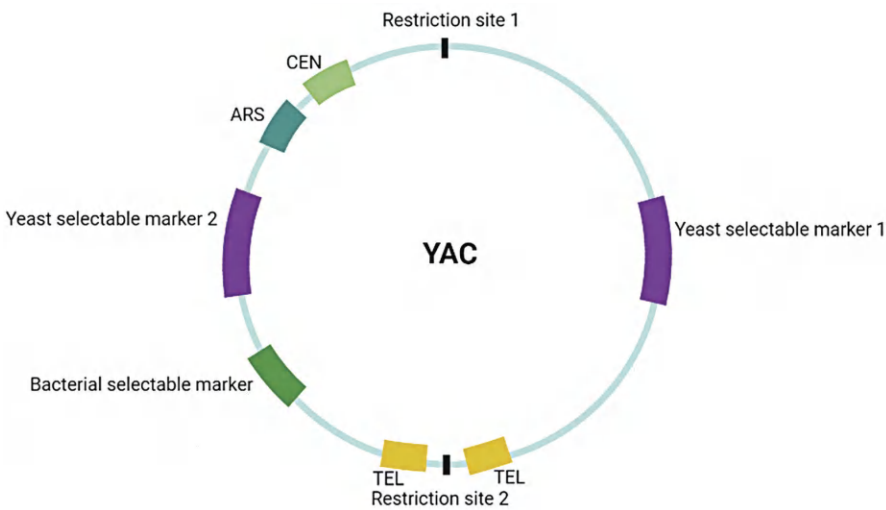


FIGURE 4.6 A schematic representation showing YAC vector. ⏎

Source: Adapted from from Nielsen et al. (2000).

4.3 SCREENING METHODS

In cloning experiments specially dealing with the vector based, the blue white screen is one of the methods that will do the rapid and easy identification of the recombinant bacterial specie. It permits the detection of final product of the reaction by visualizing only the color of the colonies of the bacteria as shown in Figure 4.7. This screening method is performed by using a selectable bacterial specie but some other organisms like the yeast cells can also give the results. The DNA of selected fragment is inserted in the vector. After that it enters in competent cell for the purpose of transformation. They are grown in presence of X-gal. The cells transformed with the vectors having recombinant DNA can produce white colonies while the cells transformed with only the vector that is actually nonrecombinant are going to produce blue colonies. It is the old method that is used to confirm the insertion into the vector.

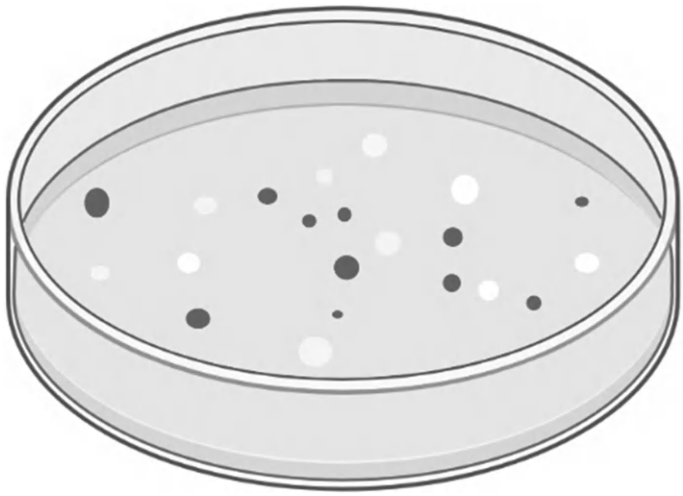


FIGURE 4.7 Blue white colonies. ⏎

Source: Adapted from Pan et al. (2014).

The basic principle of the method involving β-galactosidase gene hat undergoes α-complementation process. A very common lab bacterium is the *E. coli* DH5α cells that have been designed to enhance the transformation ability. One of the characteristic is having three types of mutations. One of them is called as the lacZM15 mutation, which deactivates the activity of lacZ gene in the cells that results in production of β-galactosidase enzyme which is in an inactive form as shown in Figure 4.8. The activation of the enzyme can be accomplished by the process of α-complementation by insertion of a plasmid containing a lacZ alpha subunit into the bacterial strain. After introducing, it causes the complementation of the truncated gene of lacZ and produces β-galactosidase enzyme in an active form (Chaffin and Rubens 1998).

The other method for screening of constructed DNA libraries is the restriction digestion method as shown in Figure 4.9. For the analysis, restriction enzymes are used to cut foreign DNA from the selected vector. A restriction enzyme known as endonuclease identifies a particular nucleotide pair region in DNA which is the site for restriction and it cuts the DNA. It hydrolyzes the linkage of phosphodiester backbones in the vicinity of that sequence. All these enzymes cleave the DNA at the carbon atom and the phosphate molecule of the phosphodiester bond. After that the fragments or pieces produced by this digestion contain only phosphates and hydroxyls (Sambrook and Russell 2006).

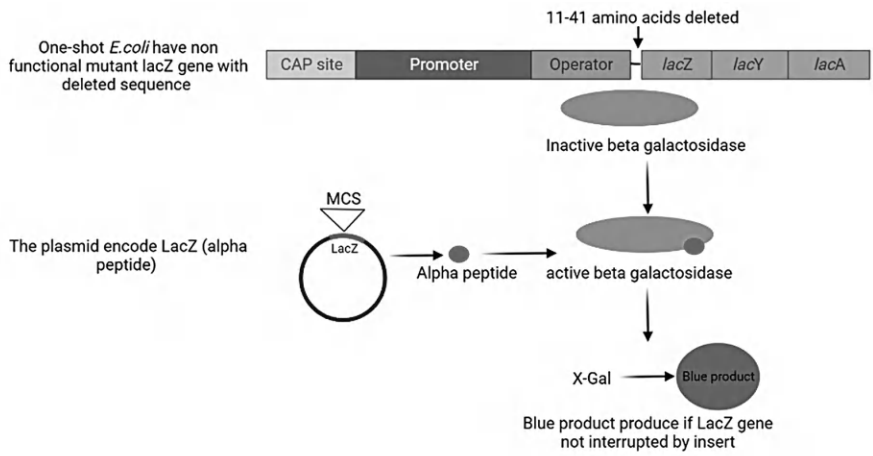


FIGURE 4.8 Mechanism of lacZ gene. ⏎

Source: Adapted from Sambrook and Russell (2016).

These identification regions are almost four, six, eight, ten, or twelve nucleotides in length and are also called as palindromic regions (i.e. the same sequence of nucleotide in the 5'–3' direction). Moreover, almost all artificial plasmids have a polylinker which is also known as "MCS" that has a plenty of restriction enzyme having different recognition regions present in a small fragment of DNA. This permits an insertion of any part of DNA into a vectors which is then easily cloned by inserting into the copied bacterial cells (Oakes, et al. 2009).

After restriction digestion of the DNA by the restriction enzymes, DNA can be visualized by agarose gel electrophoresis. In this technique, a sample is first "loaded" onto a slab of agarose gel. After that the gel is connected to an electric field, which moves the negatively charged DNA across it. All the molecules move at different speed and therefore cover different distances that actually depend on the net charge. With the increasing charge, the molecule travels further.

A fast method for screening of cloning genomic DNA is a PCR-based screening tool as shown in Figure 4.10. A cloning of DNA fragments in a construct is confirmed with the help of PCR method. It demands at least three nucleotides directly anneal to the required gene part and hence enhancing an ability of the screening method and also decreasing the possibility of getting false positive clones. A PCR requires probes to focus a particular region of DNA and a enzyme that is thermostable known as DNA polymerase which

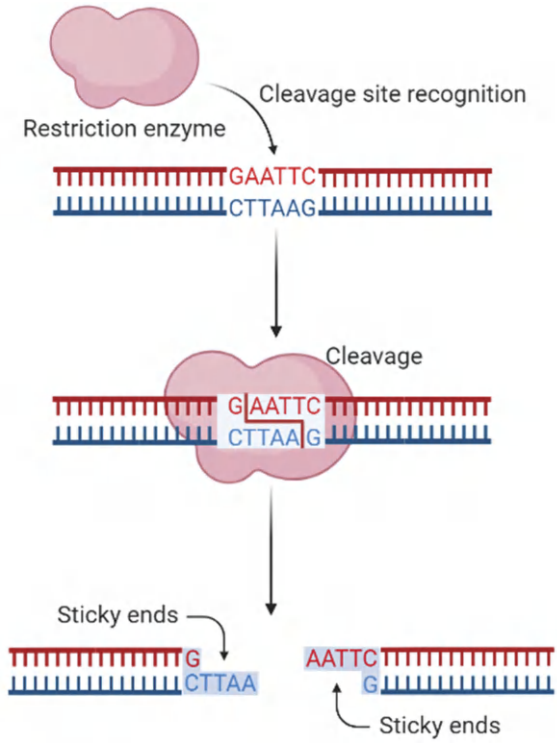


FIGURE 4.9 A schematic representation showing restriction digestion.

Source: Adapted from Khan et al. (2017).

multiplies that sequence so that visualization become easy. It always gives high sensitivity, specificity, and it is a fast technique (Milbury, et al. 2009).

The conditions for reaction mostly need an initial optimization for a specific amplified product and specific primers. One important component includes the concentration of the Mg^{2+}, the number of cycles, and range of temperature. The temperature for annealing is critical. The near temperature of melting for both primers is actually good. A thermostable enzyme known as Taq-polymerase performs catalysis at 72°C and the other polymerases that do the proofreading use other optimal temperature and causes catalysis at 60–68°C.

Most of the times PCR results depend on the concentration of primers and the selection of DNA polymerase. The specificity should be enhanced after a complex formation that is very thermosensitive. The complex is of polymerase and antibody that are used to start the hot start PCR. In this case, the polymerases are active at 95°C which is the denaturation temperature and the starting at the low temperature then mispriming problems are

reduced. The PCR specificity and amplicon concentration can be increased by enhancing substances like β-mercaptoethanol, etc.

A lengthy amplicon of PCR needs much time to prepare and require optimization. The collection of the proofreading enzymes can permit the amplification of the different lengthy templates. Additionally, the primers that bind internally can also be used thereby enhancing the specificity of reaction and the achievable amplicon size (Wolfe and Liston 1998).

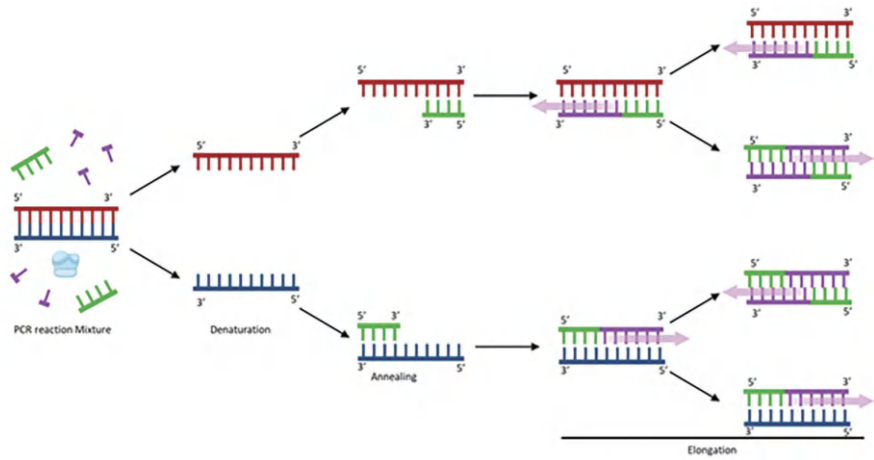


FIGURE 4.10 PCR showing cycles.

Source: Adapted from Edwards et al. (1991).

4.4 MOLECULAR ANALYSIS

In molecular biology, the isolation of biological molecules like DNA, RNA, and protein is the important method. It is actually an initial point for the later procedures and formation of the product that includes kits for diagnosis. Any molecule can be separated from the biological item like the live cells or any preserve tissues, cells, viral agent, or some other material for biological purposes (Liu 2019).

4.5 DNA

There are two categories that have the purification of the DNA that includes the separation of recombinant DNA plasmids and also separation of genomic DNA from different species. The isolated DNA or RNA should not have contaminants like proteins, carbohydrates, oils, or any other DNA or RNA. For example, DNA should not have RNA or RNA should not have DNA. The

quality and also the stability of the separated DNA or RNA can affect the experiments and results of whole scientific research.

4.6 EXTRACTION METHOD

For molecular analysis of DNA, first approach is to isolate DNA using different methods so that it will proceed to different molecular techniques for the confirmation of recombinant DNA libraries. There are two approaches for successful separation of DNA one is the conventional approach that includes phenol-chloroform extraction, CTAB method and alkaline extraction method and other one is the DNA extraction by kit. There are some steps involved in nucleic acid isolation like either by conventional or nonconventional method:

1. The disruption or lysis of cells.
2. After lysis, there is a process of denaturation of nucleoprotein complexes.
3. Removal of nucleases like RNase in case of extraction of RNA extraction and DNase in case of extraction of DNA.
4. Isolation of interested nucleic acid from the debris
5. Avoid contamination

A very common impurity in the DNA, RNA samples is salt. It's compulsory to remove salt from the samples before going to downstream techniques and further analysis. For this purpose, sometimes a single or a multiple isolation protocols are required to remove the salt from the sample. DNA precipitation can be done with the help of centrifugation and extra salt is washed away by 70% ethanol and after that centrifugation will be done to remove the supernatant. The DNA pellet obtained at the last is then mixed with autoclaved distilled water (Rogers and Bendich 1989).

4.7 ANALYSIS TECHNIQUES

4.7.1 GEL QUANTIFICATION METHOD

Gel electrophoresis isolates fragments of DNA depending on the size in a solid support medium that is known as agarose gel as shown in Figure 4.11. A DNA sample is loaded into the sample wells after that an electric current is applied which moves negatively charged DNA toward the anode that is positive end. The rate of migration of DNA fragments depend on the size, the shorter the fragments it moves more rapidly and shows band at the bottom of the gel. DNA can be seen because of an intercalating agent, ethidium

bromide that is present in a gel. Fragments of the DNA absorb the dye as they move within the gel. An illumination with the help of ultraviolet rays causes the intercalated agent to illuminate. The bigger fragments show an intense fluoresce. The smaller fragments have less DNA, absorb less dye, and therefore show less fluoresce. A ladder is a set of DNA fragments of known size loaded simultaneously and differentiate between the sizes of unknown fragments present in a gel.

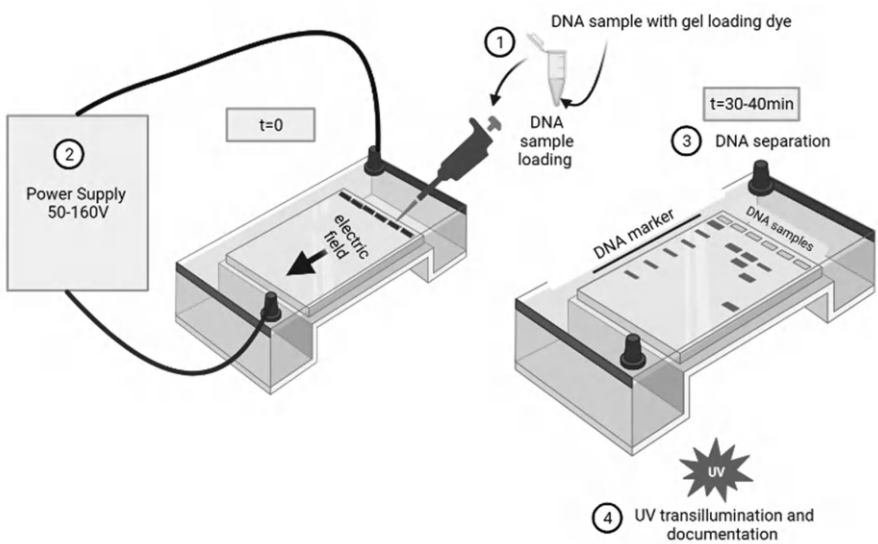


FIGURE 4.11 Horizontal gel electrophoresis. ⏎

Source: Adapted from Edwards, et al., 1991.

4.7.2 PCR

This technique makes it possible to obtain a huge number of DNA from any sample in vitro conditions. The template can be a genomic DNA or a complementary DNA that is obtained from RT-PCR from mRNA or a mitochondrial DNA. It is a technique that will produce a large amount of any sequence from the sample of DNA. The amplification depends on doubling rate of a double-stranded DNA template. This technique is divided into three stages: first a phase of denaturation, second a phase of hybridization with the primers, and last one is the elongation phase. The products of every previous synthetic step will be considered as a template for the further steps. In this way, the amplification is obtained from exponential phase.

It takes place in a reaction mixture which consists of the template, a thermostable enzyme like Taq polymerase, forward and reverse primers, and the four dNTPs in huge amount and mix thoroughly in a solution of buffer. The tube containing the whole mixture goes through repetitive cycles of temperature many times located in heating block of a thermal cycler. This apparatus is has the sample tubes and temperature can vary from 0 to 100°C.

4.7.3 SOUTHERN BLOTTING

This technique is an example restriction fragment length polymorphism (RFLP). It was developed by Edward M. Southern in 1975. It is a hybridizing approach for detection and analysis of specific DNA size from the mixture as shown in Figure 4.12. It depends on the technique of separating DNA fragments first by AGE and then it will be recognized by probe hybridization that is labeled. Firstly, the fragments of DNA are separated on the basis of size and charge during the procedure of gel electrophoresis. The isolated fragments are transferred onto the nylon membrane. In this way the specific DNA is visualized by DNA probe that shows complementation to the target DNA. A hybridized one is only 100–500 bp and a single-stranded DNA (Brown 1993).

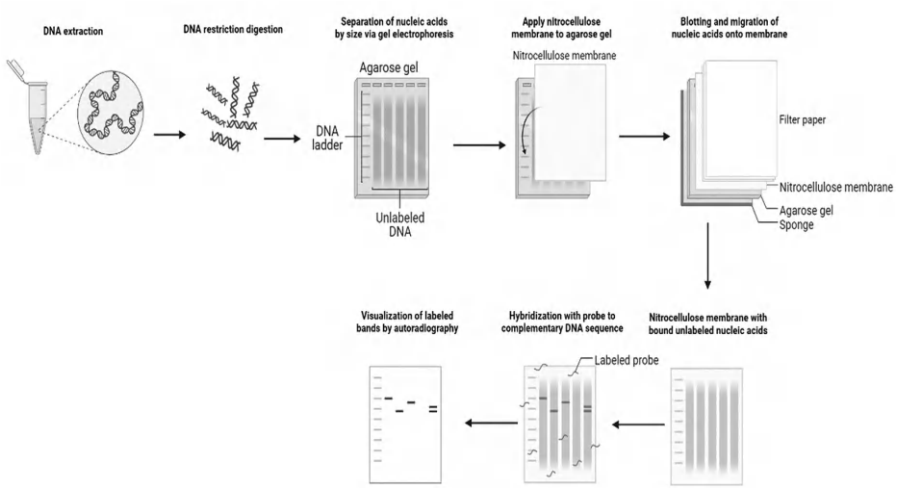


FIGURE 4.12 Southern blotting.

Source: Adapted from Southern (1991).

4.8 RNA

It is once isolated from the cell or tissues and its half life is short therefore it is an unstable molecule. There are different types of RNA that occur naturally that includes ribosomal RNA (80–90%), mRNA (2.5–5%), and transfer RNA (1–2%). RNA is a sensitive molecule therefore care is needed for its separation as it undergoes degradation as compared to the DNA. RNA is not stable because of having RNases which are the enzymes present in blood, all tissues as well as most bacteria and fungi in the environment. In order to avoid this issue, denaturing agents have been added in the isolation of pure RNA to stop inherent RNases. The extraction of RNA depends on optimized laboratory technique and most importantly to those techniques that are RNase-free. It is a heat-resistant enzyme and refolds itself after denaturation at high temperature. They can't be inactivated as they do not need any cofactors.

4.8.1 RNA-SEQ

It is a high throughput and next generation RNA sequencing and it is a quantification study for understanding the transcriptomics and gene expression. Complementary DNA is produced from total mRNA with the help of reverse transcription and then subjected to fragmentation. At the same time, the ligation with the adaptor and preparation of library are practiced before sequencing. The reading is done by the sequencer and it quantifies the cDNA complementary to the mRNA. This is basically a RNA analysis technique but based on NGS. It permits the analysis, quantification, and gene expression comparison in a target transcriptome.

4.8.2 NORTHERN BLOTTING

It is a molecular biology technique which is used to compare gene expression by detection of RNA or isolated mRNA in a given sample. RNA samples get separated on the basis of size by electrophoresis and detection is possible with a hybridization probe that is complementary to any part of or the entire target sequence as shown in Figure 4.13. The term "northern blot" is denoting the transfer of RNA from the gel to the blotting membrane. James Alwine, David Kemp, and George Stark at Stanford University developed this technique in 1977. Its name was taken because of its resemblance to the previous blotting

technique which is the southern blot on the name of scientist Edwin Southern (He and Green 2013).

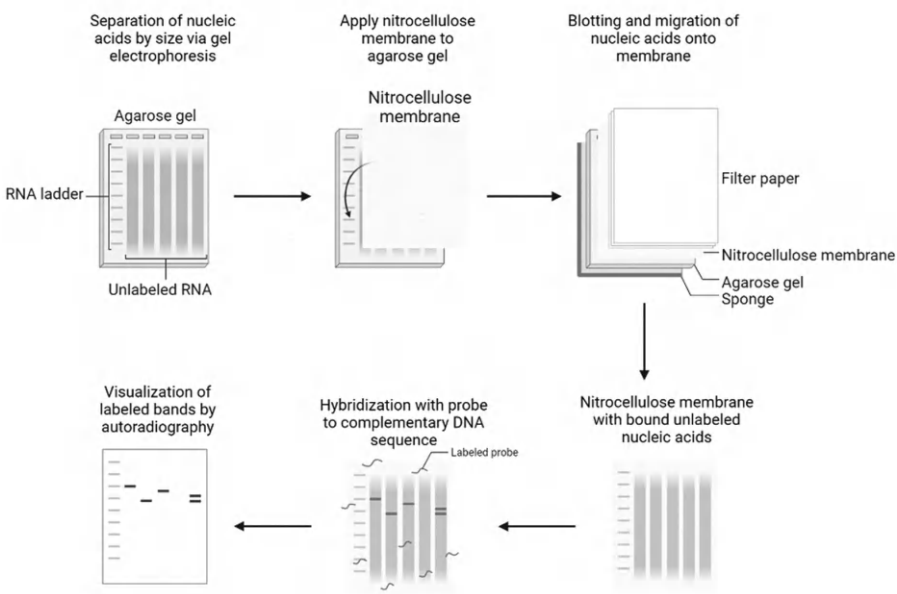


FIGURE 4.13 Northern blotting. ⏎

Source: Adapted from Josefsen and Nielsen (2011).

4.9 PROTEIN

In protein research, the purification stage is one of the most vital stages for understanding the chemistry of proteins, their unique properties, size, shape, behavior, charge, and function. The understanding of protein structure and function depends on all these factors. The extraction of cell is the initial step for the protein purification. The extraction of protein can be done by some of the methods like treatment with low ionic salt (salting out), lysis using detergent, by applying shearing force, and by applying different pressure. All these methods are used to break the membranes that are outside the cells that will permit the protein to come out of the cell. There are some factors that should be considered for isolating proteins. These are basically the extraction of protein should be performed at low temperature like 4°C. As protein gets denatured at high temperature and the condition with the buffer is an important factor. Specific buffer conditions are applied considering the sensitive nature of the proteins and also the change in pH. There are many methods

that are used in purifying the protein like the ion exchange chromatography, gel filtration, affinity chromatography, and gel electrophoresis (Preece, et al. 2017).

4.9.1 ANALYSIS

Some of the steps are involved in protein isolation are:

1. The lysis of cell is the first step in the purification of protein.
2. Release of protein but in soluble form from the native cell.
3. A physical lysis technique is used to get rid of the cell wall, for example, French press or glass beads.

Ion exchange chromatography isolates the protein on the basis of charge with the help of specific resins that are of two types, positively charged resins or negatively charged resins with their respective chemical groups. Most of the proteins are neutral that depends on their isoelectric point (pI) at a certain pH. If the protein net charge is positive at a pH that is below the pI value, then it will bind to a cation exchanger. A pH that is actually above the pI value and protein net charge is negative, then it will bind to an anion exchanger.

Gel filtration chromatography also known as the size exclusion or the gel permeation chromatography. It isolates the proteins on the basis of molecular sizes and shape. In this technique, bigger molecules pass the column faster as compared to the small molecules. The small molecules penetrate into the bead of matrix that is like small holes. So the small size molecules will take more time to come out of the column as they enter into the porous medium as compared to the larger protein molecules that do not enter into the holes so they elute faster.

Affinity chromatography deals with the separation of protein by its specific interaction with the solid phase in order to get rid of the contaminant molecules. It purifies a protein depending on the biological function. The proteins that are having more affinity with the ligand molecule will strongly attach to the matrix of the column while the other one is directly eluted from the column. There are some interactions that bind ligand with the final target protein like hydrogen bonding, van deer Waals forces, and hydrophobic interactions. For the elution of the bound proteins, a solution having concentrated form of ligand that is soluble will pass through the column. The attachment

of the ligand and target protein should be reversible so that it will permit proteins to elute inactive form (Urh, et al. 2009).

4.9.2 GEL ELECTROPHORESIS

It is a technique that is separating a protein on the basis of their size to charge ratio. The protein that is purified from chromatography is not fully purified so rest is purified by non-denaturing polyacrylamide gel electrophoresis (PAGE) or also called as native gel electrophoresis as shown in Figure 4.14.

In PAGE, the proteins move by applying a current to the gelated matrix. The protein movement in the gel depends on the charge density of the particles. The protein molecules that have more charge density move rapidly. The pore size of acrylamide is playing an important part to isolate different sizes of the protein molecules. The larger the protein, slow it will move because it gets stacked in the gel. A shape is another parameter because those proteins move quickly that are more compact and globular proteins as compared to

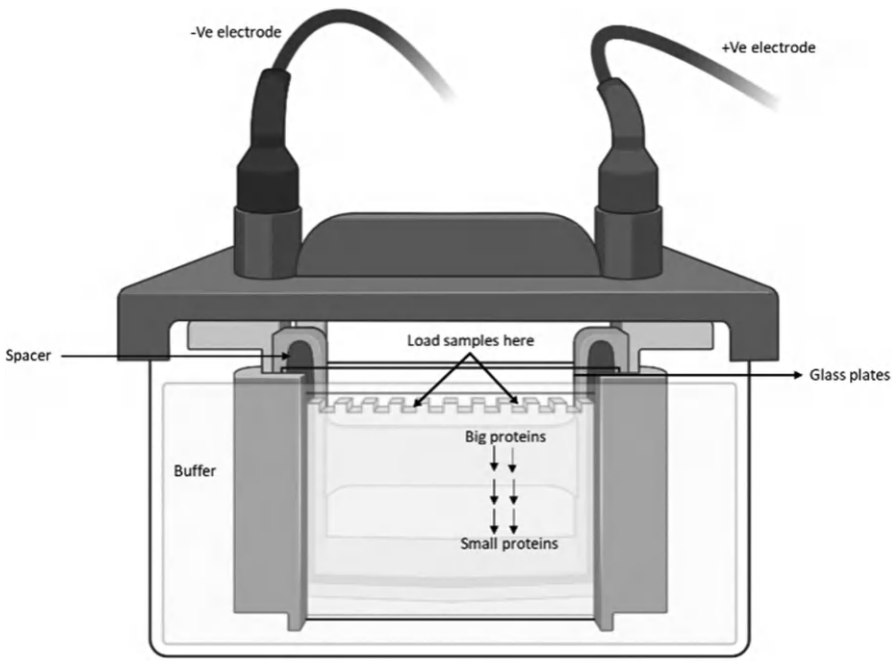


FIGURE 4.14 A schematic representation of vertical gel electrophoresis. ⏎

Source: Adapted from Weber and Osborn (1969).

those that are elongated and fibrous proteins having same molecular weight. PAGE is performed by the addition of sodium dodecyl sulfate (SDS). Treatment with SDS will remove the confirmations like secondary, tertiary, and quaternary. A protein becomes straight like a rod because of electrostatic repulsion between the SDS molecules. There is a direct proportionality between the number of SDS molecules and the molecular weight of the protein (Fritz 1987).

4.9.3 WESTERN BLOTTING

It is also called as immunoblotting because of one reason that a specific antigen is detected by an antibody. It utilizes SDS gel to separate different proteins from a sample. The separated proteins are transferred onto a membrane like a nitrocellulose or PVDF membrane then is incubated with antibodies that act as probe and they are very specific to the final protein as shown in Figure 4.15.

Western blotting technique could be helpful to detect the final protein having low concentration as approximately 1 ng because of gel's high resolution and great specificity with high sensitivity of the immunoassay (Bass, et al. 2017).

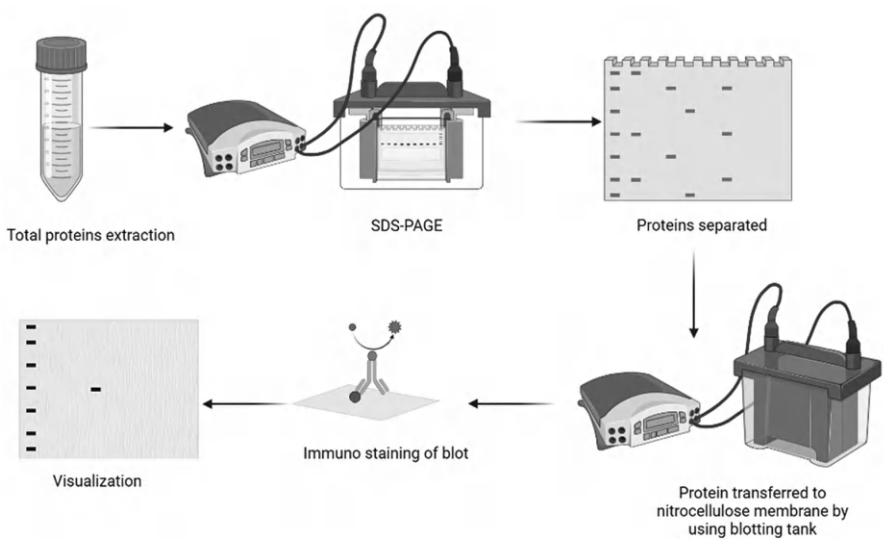


FIGURE 4.15 Western blotting. ⏎

Source: Adapted from Kurien and Scofield (2006).

4.10 CONCLUSION

The technologies that are used for the DNA libraries development and analysis are emerging very fast. The different sources of DNA and cloning vectors are used to make DNA libraries. The large insert size DNA libraries like BAC, PAC YAC, PBC, and BIBAC give rise to the revolution in physical mapping, chromosome walking, and sequencing of large genomes. The advantages in different clones of DNA having all genes showed that DNA libraries are tools for every kind of research purpose. Friedrich Miescher successfully isolated the first DNA in 1869 and in 1958, Meselson and Stahl developed the first extraction method for the DNA by density gradient centrifugation. There are different approaches for purifying biomolecules. The gradual development from extraction with phenol chloroform to the usage of column technique and also purification procedures by chromatography to immune blotting are used to obtain the proteins. An isolation of biological molecules has aided researchers in genetic engineering techniques and analysis in molecular biology for the better understanding of the science.

KEYWORDS

- **reverse transcription**
- **DNA marker**
- **physical mapping**
- **clones**
- **gene libraries**
- **vectors**

REFERENCES

Bass, J. J.; Wilkinson, D. J.; Rankin, D.; Phillips, B. E.; Szewczyk, N. J.; Smith, K.; Atherton, P. J. An Overview of Technical Considerations for Western Blotting Applications to Physiological Research. *Scand. J. Med. Sci. Sports* **2017,** *27*, 4–25.

Berry, A. E.; Chiocchini, C.; Selby, T.; Sosio, M.; Wellington, E. M. Isolation of High Molecular Weight DNA from Soil for Cloning into BAC Vectors. *FEMS Microbiol. Lett.* **2003,** *223*, 15–20.

Blattner, F. R.; Williams, B. G.; Blechl, A. E.; Denniston-Thompson, K.; Faber, H. E.; Furlong, L.-A.; Grunwald, D. J.; Kiefer, D. O.; Moore, D. D.; Schumm, J. W. Charon Phages: Safer Derivatives of Bacteriophage Lambda for DNA Cloning. *Science* **1977,** *196*, 161–169.

Brown, T. Southern Blotting. *Curr. Protocols Mol. Biol.* **1993,** *21*, 2–9.

Chaffin, D.; Rubens, C. Blue/White Screening of Recombinant Plasmids in Gram-Positive Bacteria by Interruption of Alkaline Phosphatase Gene (phoZ) Expression. *Gene* **1998,** *219*, 91–99.

Chauthaiwale, V. M.; Therwath, A.; Deshpande, V. Bacteriophage Lambda as a Cloning Vector. *Microbiol. Rev.* **1992,** *56*, 577–591.

Cossons, N.; Nielsen, T.; Dini, C.; Tomilin, N.; Young, D.; Riabowol, K.; Rattner, J.; Johnston, R.; Zannis-Hadjopoulos, M.; Price, G. Circular YAC Vectors Containing a Small Mammalian Origin Sequence Can Associate With the Nuclear Matrix. *J. Cell. Biochem.* **1997,** *67*, 439–450.

Edwards K, Johnstone C, Thompson C 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic acids research 19: 1349.

Eun H-M. 1996. Enzymology primer for recombinant DNA technology: Elsevier.

Fritz, J. S. Ion Chromatography. *Anal. Chem.* **1987,** *59*, 335A–344A.

He, S. L.; Green, R. Northern Blotting. In. *Methods in Enzymology*; Elsevier, 2013, pp 75–87.

Henne, A.; Daniel, R.; Schmitz, R. A.; Gottschalk, G. Construction of Environmental DNA Libraries in Escherichia coli and Screening for the Presence of Genes Conferring Utilization of 4-Hydroxybutyrate. *Appl. Environ. Microbiol.* **1999,** *65*, 3901–3907.

Ish-Horowicz, D.; Burke, J. Rapid and Efficient Cosmid Cloning. *Nucl. Acids Res.* **1981,** *9*, 2989–2898.

Jakobovits, A. Tools of the Trade YAC Vectors: Humanizing the Mouse Genome. *Curr. Biol.* **1994,** *4*, 761–763.

Jenkins, R. E.; Pennington, S. R. Arrays for Protein Expression Profiling: Towards a Viable Alternative to Two-Dimensional Gel Electrophoresis? *Proteomics* **2001,** *1*, 13–29.

Josefsen K, Nielsen H. 2011. Northern blotting analysis. In. RNA: Springer. p. 87–105.

Khan S, Qureshi MI, Alam T, Abdin M 2007. Protocol for isolation of genomic DNA from dry and fresh roots of medicinal plants suitable for RAPD and restriction digestion. African Journal of Biotechnology 6: 175–178.

Kim S-R, Lee D-Y, Yang J-I, Moon S, An G 2009. Cloning vectors for rice. Journal of Plant Biology 52: 73–78.

Kim, U.-J.; Shizuya, H.; de Jong, P. J.; Birren, B.; Simon, M. I. Stable Propagation of Cosmid Sized Human DNA Inserts in an F Factor Based Vector. *Nucl. Acids Res.* **1992,** *20*, 1083–1085.

Larocca D, Brug M, Jensen-Pergakes K, Ravey E, Gonzalez A, Baird A 2002. Evolving phage vectors for cell targeted gene delivery. Current pharmaceutical biotechnology 3: 45–57.

Lathe, R.; Kieny, M.; Schmit, D.; Curtis, P.; Lecocq, J. M13 Bacteriophage Vectors for the Expression of Foreign Proteins in Escherichia Coli: The Rabies Glycoprotein. *J. Mol. Appl. Genet.* **1984,** *2*, 331–342.

Liu, B. BioSeq-Analysis: A Platform for DNA, RNA and Protein Sequence Analysis Based on Machine Learning Approaches. *Brief. Bioinform.* **2019,** *20*, 1280–1294.

MacNeil, I.; Tiong, C.; Minor, C.; August, P.; Grossman, T.; Loiacono, K.; Lynch, B.; Phillips, T.; Narula, S.; Sundaramoorthi, R. Expression and Isolation of Antimicrobial Small Molecules from Soil DNA Libraries. *J. Mol. Microbiol. Biotechnol.* **2001,** *3*, 301–308.

Milbury, C. A.; Li, J.; Makrigiorgos, G. M. PCR-Based Methods for the Enrichment of Minority Alleles and Mutations. *Clin. Chem.* **2009,** *55*, 632–640.

Murray, N. E.; Murray, K. Manipulation of Restriction Targets in Phage λ to Form Receptor Chromosomes for DNA Fragments. *Nature* **1974,** *251*, 476–481.

Nielsen TO, Cossons NH, Zannis-Hadjopoulos M, Price GB 2000. Circular YAC vectors containing short mammalian origin sequences are maintained under selection as HeLa episomes. Journal of cellular biochemistry 76: 674–685.

Norrander, J.; Kempe, T.; Messing, J. Construction of Improved M13 Vectors Using Oligodeoxynucleotide-Directed Mutagenesis. *Gene* **1983,** *26*, 101–106.

Oakes, C. C.; Salle, S. L.; Trasler, J. M.; Robaire, B. Restriction Digestion and Real-Time PCR (qAMP). In *DNA Methylation*; Springer, 2009, pp 271–280.

Pan Y, Zhang R, Sun A, Xiao L, Li AS, Wang Q, Zhang J, Li K 2014. Effects of in-frame and frame-shift inserts on the colors of colonies in beta-gal based blue/white screening assay. Science of Advanced Materials 6: 817–821.

Preece, K.; Hooshyar, N.; Zuidam, N. Whole Soybean Protein Extraction Processes: A Review. *Innov. Food Sci. Emerg. Technol.* **2017,** *43*, 163–172.

Riley, J.; Butler, R.; Ogilvie, D.; Finniear, R.; Jenner, D.; Powell, S.; Anand, R.; Smith, J.; Markham, A. A Novel, Rapid Method for the Isolation of Terminal Sequences from Yeast Artificial Chromosome (YAC) Clones. *Nucl. Acids Res.* **1990,** *18*, 2887–2890.

Rogers, S. O.; Bendich, A. J. Extraction of DNA from Plant Tissues. In *Plant Molecular Biology Manual*; Springer, 1989; pp 73–83.

Saito, I.; Stark, G. R. Charomids: Cosmid Vectors for Efficient Cloning and Mapping of Large or Small Restriction Fragments. *Proc. Natl. Acad. Sci*. **1986,** *83*, 8664–8668.

Sambrook, J.; Russell, D. W. Screening Bacterial Colonies Using X-gal and IPTG: α-Complementation. *Cold Spring Harbor Protocols* **2006,** pdb. prot3945.

Snapper, S.; Melton, R.; Mustafa, S.; Kieser, T.; Isolation and Characterization of Efficient Plasmid Transformation Mutants of Mycobacterium Smegmatis. *Mol. Microbiol.* **1990,** *4*, 1911–1919.

Tan, S. C.; Yiap, B. C. DNA, RNA, and Protein Extraction: The Past and the Present. *J. Biomed. Biotechnol.* **2009,** *2009*, 574398.

Urh, M.; Simpson D, Zhao K. 2009. Affinity Chromatography: General Methods. *Methods Enzymol.* **2009,** *463*, 417–438.

Voytas D 1992. Agarose gel electrophoresis. Current protocols in immunology 2: 10.14. 11–10.14. 18.

Wahl GM, Lewis KA, RuIz JC, Rothenberg B, Zhao J, Evans GA 1987. Cosmid vectors for rapid genomic walking, restriction mapping, and gene transfer. Proceedings of the National Academy of Sciences 84: 2160–2164.

Weber K, Osborn M 1969. The reliability of molecular weight determinations by dodecyl sulfate-polyacrylamide gel electrophoresis. Journal of Biological Chemistry 244: 4406–4412.

Wolfe AD, Liston A. 1998. Contributions of PCR-based methods to plant systematics and evolutionary biology. In. Molecular systematics of plants II: Springer. p. 43–86.

Wolfe, A. D.; Liston, A. Contributions of PCR-Based Methods to Plant Systematics and Evolutionary Biology. In *Molecular Systematics of Plants II*; Springer, 1998; pp 43–86.

Wu, C.; Xu, Z.; Zhang, H. B. DNA Libraries. In *Reviews in Cell Biology and Molecular Medicine*; 2006.

Yanisch-Perron, C.; Vieira, J.; Messing J. Improved M13 Phage Cloning Vectors and Host Strains: Nucleotide Sequences of the M13mpl8 and pUC19 Vectors. *Gene* **1985,** *33*, 103–119.

CHAPTER 5

Genetic Engineering Tools: Restriction Enzymes and Vectors

ROOPASHREE M. S. and VENKATESH K. H.

Department of Life Science, Bangalore University, Jnana Bharathi Campus, Bengaluru, Karnataka, India

ABSTRACT

Genetic engineering techniques have been applied for various benefits of mankind, as it is the basis for modern-day scientific research. The technique involves cutting and ligating a part of the genetic material, specifically DNA from a particular organism, followed by the introduction of hybrid DNA into a host, which results in new combinations of genetic material that are heritable.

The bioprocessing industry employs genetic engineering pathways for the development of a number of products. Molecular and cellular manipulations are made possible by the use of enzymes such as restriction endonucleases, DNA ligases, and suitable vectors, such as plasmids, viruses, and hybrid vectors, and isolation and introduction of desirable DNA fragments to enable the expression of desired products.

Restriction enzymes cut through double-stranded DNA molecules precisely and form a vital tool in manipulating DNA molecules. A number of restriction enzymes have been identified (>19,000). These enzymes cut the DNA by identifying and recognizing a particular pattern of nucleotides. Based on the site of binding and cutting of DNA, restriction enzymes are classified into three types. Among them, type II binds and cuts the DNA at the same position and thus has gained much significance. It is used for splicing (cutting and insertion) of genes of interest into other segments of DNA, thereby providing a means to modify DNA to obtain new forms.

Molecular Genetics, Structures, Mechanisms, and Functions, Volume 1: Principles of Gene Manipulation and Genomics. Tariq Ahmad Bhat, Abdul Rauf Shakoori, and Jameel M. Al-Khayri (Eds.)
© 2026 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

Exogenous DNA obtained using restriction enzymes is delivered into a host organism through the vectors. Vectors are DNA molecules that carry the foreign gene and facilitate the integration and replication of the DNA molecule in the host. The use of vectors in genetic engineering enables the process of delivery of foreign genes into a host to be economical, simple, and rapid. Thus, vectors have been singled out as the most important tool for the transfer of foreign DNA into host cells. Organisms are genetically engineered using vectors to produce important products, such as insulin, employing *Escherichia coli* as the host. A specific gene sequence can be introduced within a genome, and nucleotide sequences can be determined through DNA sequencing by employing vectors. In molecular biology, vectors are vital components that can be used to study cell structure, composition of nucleic acid, production of proteins, antibiotics, diagnostic probes, recombinant antigens, recombinant vaccines, hormones, etc. They are employed to study additional traits of engineered plants and animals.

This chapter throws light on the types and functions of restriction enzymes and vectors used for the transfer of desired DNA fragments.

5.1 INTRODUCTION TO GENETIC ENGINEERING

Techniques involving DNA manipulation led to an era of DNA Science and technology which was made possible by the discovery of DNA structure and elucidation of DNA as universal genetic material. The term "biotechnology" encompasses the utilization of biological processes, organisms, or systems that can be used to obtain products that enhance human lives. It is a process of improving a specific trait by engineering organisms. In this process, living systems are utilized, or natural processes are influenced to obtain desirable products, systems, or environments that can help human development. A bacterial plasmid DNA molecule is spliced using fragments from naturally occurring plasmids that may be of distinct origin to obtain a genetically modified organism. Jack Williamson first employed the term "Genetic Engineering" in his novel Dragon's Island science fiction in 1951. It is a process in which the genetic makeup of the target cell is deliberately altered to add additional traits to the organism. It is also known as the genetic modification technique as the genome of an organism is manipulated by transferring genes within and across species barriers to construct recombinant DNA.

Various techniques are employed for the manipulation of genetic material so as to enhance specific functions and alter or repair the genome.

Reconstructed DNA strands are spliced into different hosts, such as bacteria or bacteriophages, by empowering vectors or by microinjection into the host. Synthetic biology is a process that employs new techniques to design and build novel organisms that have supplemented the traditional tools to engineer novel life forms.

Genetic engineering encompasses all the techniques that are employed to alter an organism artificially so as to obtain the desired substances (such as an enzyme or a metabolite), which are otherwise not naturally produced by the organism, or to enhance the production of a preexisting cellular process. The road map for the production includes, isolation of the desired DNA segment or extracting and purifying the total cellular DNA from a source organism. The next step involves the insertion of manipulated DNA using various techniques into a genetic carrier molecule, and it is delivered to the host strain. The strategy for gene delivery depends on the strain used. Methods of gene delivery include transformation, transfection, transduction, or conjugation. The transformed cell acquires desired characteristics such as antibiotic production, disease resistance, or to produce enzymes, hormones, vaccines, etc. Each modified cell will produce multiple copies of the host cell along with desired traits.

5.2 PRINCIPLES OF GENETIC ENGINEERING

DNA cloning technique essentially involves

1. A host organism is chosen.
2. Gene of interest is isolated from the total DNA using restriction enzymes.
3. Gene of interest is ligated to a DNA molecule called a vector to obtain a hybrid vector.
4. This hybrid vector is transferred into a host cell which multiplies to produce a clone of cells with additional traits.
5. Selection of transformed cells by screening for desired DNA inserts and biological properties.

5.3 COMPONENTS OF GENETIC ENGINEERING

Biological tools used for genetic transformation include the following:

a. Restriction enzymes.
b. Vector DNA.

c. Passenger DNA (gene of interest).
d. Host for cloning.

5.3.1 RESTRICTION ENZYMES

Restriction enzymes are protein molecules that recognize and bind to a specific, short nucleotide sequence and make a cut in the DNA. The region where it binds is known as the "recognition site."

Bacteria, archaebacteria, and a few other prokaryotes produce restriction enzymes in order to defend themselves against viral infections. They recognize and bind to specific nucleotide sequences on the double-stranded DNA molecule. The pattern of cut is at the site of binding or nearer to it. Foreign DNA is selectively cut without damaging the host DNA, and the process is known as restriction. By a process called methylation, host DNA is protected, and its cleavage activity is prevented, a process known as a restriction-modification system.

More than 600 restriction enzymes are used for the genetic engineering process. Salvador Luria and Giuseppe Bertani identified the role of these enzymes in the 1950s in *Escherichia coli* infected with lambda phage.

5.3.1.1 NOMENCLATURE

The role of restriction digestion caused by enzymatic cleavage of phage DNA was studied by Werner Arber and Mathew Meselson in the 1960s, who termed it a "restriction enzyme." Nomenclature of restriction enzymes depends upon the bacterial genus, species, and strain from which they are obtained.

***5.3.1.1.1** Pattern of Nomenclature*

The nomenclature of restriction enzymes follows a pattern wherein the first letter is written in capitals, and it indicates the name of the genus from which the enzyme is isolated. The next two letters are small alphabets and represent the species of the organism. The enzyme name is written in italics or underlined; for example, *Eco*RI where *E* represents the genus *Escherichia*, core presents species *E. coli*, R stands for strain, which is depicted as subscript Eco_{R}, and in practice, subscripts are ordinarily written online. If an organism produces more than one enzyme, they are identified by sequential Roman numerals, "I" represents the order of identification.

5.3.1.2 CLASSIFICATION OF RESTRICTION ENZYMES

Restriction enzymes are of three types and are classified based on their structure as well as the requirement of enzyme cofactors. Classification is also based on the recognition site, target sequence, and the region where they make a cut. It mainly depends upon the region where the DNA substrate cuts, either at their recognition site or if the recognition and cleavage sites are away from one another. Restriction enzymes bind at specific sites on both the strands of DNA and make two incisions at the sugar-phosphate backbone on both the strands of the double helix.

5.3.1.2.1 Type I

Complex type-I structure cleaves only one strand of DNA. Its recognition sequences are about 15-bp in length and require Mg^{++} ions, *S*-adenosyl-L-methionine, and ATP for their functioning. It has both methylase and restriction activities. It cleaves the DNA about 1000 bp away from the 5′-end of the sequence TCA located within the recognition site; for example, *Eco*B.

5.3.1.2.2 Type II

Type II restriction enzymes are commonly used in gene cloning techniques. They are more stable, have a single function, and are independent of methylase activity. They bind at the recognition site and cleave immediately outside the recognition sequences in both strands of DNA but require Mg^{++} ions for cutting specific lengths of DNA. They do not require ATP and recognize palindromic sequences having rotational symmetry. The first half of one strand is the mirror image of the second half of the other strand; for example, *Eco*RI, *Hae*III, etc.

5.3.1.2.2.1 Palindromic sequences with rotational symmetry

5′ …..GAT ATC…. 3′

3′ ….CTA TAG…. 5

5.3.1.2.3 Type III

They are intermediate between types I and II restriction endonucleases. The pattern of DNA cleavage is at well-defined sites in the immediate vicinity of

recognition sites, requires Mg^{++} ions and ATP for cleavage, and is not used for gene cloning; for example, *Hinf*III.

TABLE 5.1 Types of Restriction Enzymes and their Features.

Types	Type I	Type II	Type III
Examples	*Eco*KI, *Eco*AI	*Eco*RI, *Eco*RV	*Eco*P15I, *Eco*PI
Cofactors and activators	Mg^{2+}, AdoMet, ATP (hydrolyzed)	Mg^{2+}	Mg^{2+}, AdoMet (stimulates activity but not required), ATP (not hydrolyzed)
Recognition site	Interrupted bipartite	Palindromic or interrupted palindrome	Nonpalindromic
Cleavage site	Cleaves far from the recognition site	Defined, within restriction site, give sticky or blunt ends	Cleaves ~ 25 bases away from the restriction site
Subunits and structure	Pentameric complex (2R, 2M, and 1S)	Homodimer (2R–S)	Both R and M–S required
Methylase features	Heterodimer (1M and 1S) or heterotrimer (2M adn 1S)	Separate, single, monomeric (M–S)	Methylates adenines, only one strand independently

5.3.1.3 ARTIFICIAL RESTRICTION ENZYMES

Natural or engineered DNA binding domain is fused to a nuclease domain, often the cleavage domain of the type IIS restriction enzyme *Fok*I. Large DNA sites (up to 36 bp) are targeted and engineered to bind to desirable DNA sequences. The most commonly used artificial restriction enzymes are zinc finger nucleases. They are mainly used in genetic engineering applications, but can also be used for more standard gene cloning applications. Other artificial restriction enzymes are based on the DNA binding domain of transcription activator-like (TAL) effectors.

5.3.1.4 PATTERN OF CLEAVAGE BY RESTRICTION ENZYMES

Restriction enzymes can cleave DNA molecules in two patterns:

1. Blunt ends: In this method, both strands of DNA are cleaved within the recognition sequence at the same point. This type of cut generates sequences with blunt ends; for example, *Pvu*II, *Hae*lll, and *Alu*l.

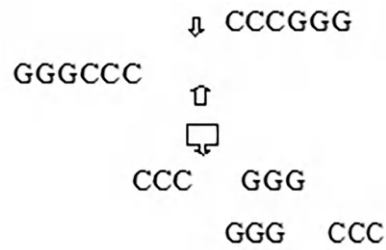


FIGURE 5.1 Blunt end cut by restriction enzymes.

1. Sticky ends: In this method, the two strands of DNA are cleaved at two different points, producing cuts that are termed "sticky" or staggered ends. This results in the formation of protruding ends, where one strand extends a few bases beyond the other strand. Example, *Eco*RI.

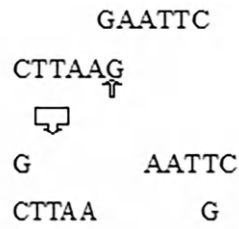


FIGURE 5.2 Sticky end cut by restriction enzymes.

TABLE 5.2 Examples of Restriction Enzymes.

Enzyme	Organism	Recognition site	Type of cut (sticky or blunt)
EcoRI	Escherichia coli	5'GAATTC	5'---GAATTC--3'
	Sticky ends	3'CTTAAG	**3'---CTTAAG---5'**
BamH1	Bacillus amyloliquefaciens	5'GGATCC	5'---GGATCC--3'
	Sticky ends	3'CCTAGG	3'---CCTAGG---5'
Taq1	Thermus aquaticus	5'TCGA	5'---TCGA---3'
	Sticky ends	3'AGCT	**3'---AGCT---5'**
Alu1	Arthrobacter luteus	5'AGCT	5'---AGCT---3'
	Blunt ends	3'TCGA	3'---TCGA---5'

Each restriction enzyme has a unique recognition sequence, producing fragments with differences in the length sequence and strand orientation (5' end or the 3' end) of a sticky-end "overhang" of an enzyme restriction. If the same sequences are recognized by different restriction enzymes, they

are known as "Neoschizomers" and they often cleave in different locales of the sequence, while different enzymes that recognize and cleave at the same location are known as "Isoschizomers."

5.3.1.4.1 Applications of restriction enzymes

1. Gene cloning and protein expression techniques commonly employ plasmid vectors to insert desired genes into host cells. Polylinker regions (multiple cloning sites [MCSs]) on plasmids are modified to contain a number of restriction enzyme recognition sequences, which gives the flexibility to ligate desirable genes into vectors. The same restriction enzymes are generally used to cut both plasmid DNA as well as vectors and DNA ligase are used to join them together.
2. They are used for the generation of DNA maps, DNA libraries, Restriction fragment length polymorphism (RFLP), Epigenetic modifications, etc.

5.3.2 VECTORS USED IN GENETIC ENGINEERING

5.3.2.1 INTRODUCTION

Vectors are DNA molecules that carry foreign DNA into a host organism. It also facilitates stable integration and replication of introduced genes inside the host cell system. In the molecular cloning technique, DNA molecules undergo restriction digestion (both target DNA and vector), ligation of target DNA with the vector, followed by introducing a hybrid vector into a host organism for replication.

Vectors are autonomously replicating DNA molecules that carry foreign DNA for transfer or propagation in an organism. The functional characteristics of vectors include

a. Ori- origin of replication region

This region carries a specific sequence of nucleotides to initiate replication, enabling the vector to replicate autonomously inside the host cell. Foreign DNA attached to the vector will also replicate alongside.

b. Cloning region/site

c. This is the region where foreign DNA is joined and is the point of entry or analysis for genetic engineering. Vector DNA is digested with the restriction enzymes at this site, followed by the ligation of

foreign DNA. Plasmids with MCS harboring up to 20 restriction sites have been produced for cloning works.

d. Marker region/selectable marker

There are certain specific genes on a plasmid that confer resistance to particular antibiotics or selective agents, which is fatal for the host organism under normal conditions. These genes confer the host cell with the property to survive and propagate in a specific culture medium containing specific antibiotics or enzymes.

e. Reporter Gene

Reporter genes permit successful screening of clones or recombinant cells. This technique is utilized extensively in blue–white colony selection.

Vectors used in genetic engineering techniques must not transfer the new traits to the natural population of bacteria via the conjugation process. A variety of highly specialized vectors that are designed for specific functions are available now.

5.3.2.2 CLONING VECTOR

DNA molecules that carry foreign DNA into a host cell are known as cloning vectors. They replicate and produce many copies of themselves and the foreign DNA inside the host cell. Vectors are chosen depending on the objective of the cloning experiment, ease of working, knowledge existing about them, suitability and reliability.

Vectors used for cloning purposes are given in Table 5.3.

TABLE 5.3 Types of Vectors and their Properties. ⏎

Sl. No.	Vectors	Uses	Insert size (kb)
1	Plasmid	DNA manipulations	10–20
2	λ phage (insertion)	Genomic libraries	~10
3	λ phage (replacement)	Genomic libraries	~23
4	M13	DNA manipulations, *in vitro* mutagenesis	8–9
5	Phagemid	DNA manipulations, *in vitro* mutagenesis	10–20
6	Cosmid	Genomic libraries	~44
7	Bacterial artificial chromosome (BAC)	Genomic libraries	130–150
8	Yeast artificial chromosome (YAC)	Genomic libraries	1000–2000

5.3.2.3 PLASMIDS AS CLONING VECTORS

Plasmids are extrachromosomal, circular, double-stranded DNA molecules occurring naturally in bacteria and a few other organisms. They are capable of replicating autonomously inside a cell. Their size ranges from 1.0 kb to over 250 kb. Plasmid vectors facilitate the replication of incorporated foreign DNA, isolation and identification of transformed cells. Plasmids can be designed in a specific manner which allows the cells to transcribe and translate the foreign gene into proteins.

Plasmid vectors used for cloning are 2–4 kb in length and are capable of carrying foreign DNA up to 15 kb in size. Plasmids are closed circular double-stranded DNA molecules that confer a specific phenotype to the host bacterial cells. Few of the plasmids integrate themselves into the bacterial chromosomes and are referred to as "episomes." Plasmids contain genes that encode proteins required for replication (Replication proteins), and such genes are located close to the origin of replication. All the other enzymes essential for replication, such as DNA polymerases, DNA ligase, helicase, etc., are synthesized by the host cell.

The *ori* region of plasmids will determine its host range. Plasmids having ColE1 *ori* region have a restricted host range. ColE1 plasmids replicate in enteric bacteria, such as *E. coli*, *Salmonella*, etc., whereas RP4-type plasmids can replicate in several Gram-negative bacteria. RP4-type plasmids are readily transmitted by conjugation, while RSF1010-type plasmids are not conjugative but can be transformed into a wide range of Gram-negative bacteria. Specific plasmids have a broad host range and can replicate in many other Gram-positive bacteria, such as plasmids isolated from *Staphylococcus aureus*.

5.3.2.3.1 Types of Plasmids

Based on the phenol type trait conferred, plasmids are divided into six major classes:

1. **R plasmids or Resistance plasmids:** These plasmids carry genes which give resistance to the bacteria from one or more chemical agents, such as antibacterial agents. Such plasmids are vital in clinical microbiology as they confer profound consequences in treating bacterial infections; for example, RP4 plasmid is commonly found in *Pseudomonas* and in many other bacteria.

2. **F plasmids or Fertility plasmids:** These are found in F^+ bacterium and are conjugative plasmids with a higher frequency of conjugation. They carry the transfer gene (*tra*) and form a conjugative bridge (F pilus) with F-bacterium; for example, *E. coli* with F plasmid.
3. **Col plasmids:** These are found in ColE1 of *E. coli* and have genes that code for colicins (proteins) that kill other bacteria.
4. **Degradative plasmids:** Host cells carrying this plasmid have the ability to metabolize unusual molecules, such as toluene and salicylic acid; for example, TOL plasmid of *Pseudomonas putida*.
5. **Virulence plasmids:** These plasmids confer pathogenicity on the host bacterium; for example, Ti plasmids of *A. tumefaciens* induce crown gall disease upon infecting dicotyledonous plants.
6. **Cryptic plasmids:** These do not confer any apparent effect on the phenotype of the cell but code for enzymes that are required for the replication and maintenance of the host cell.

Plasmids are divided into two major classes such as natural and artificial based on their origin or source:

a. **Natural plasmids:** They are present in prokaryotes or eukaryotes naturally; for example: ColE1.
b. **Artificial plasmids:** These plasmids are constructed by re-combining selected fragments of two or more naturally occurring plasmids or artificial plasmids in vitro. Example: pBR322.

Based on the number of copies of plasmids produced in a host cell, plasmids are classified by copy number. They are termed as high copy number (>20) plasmids or low copy number (<10) plasmids depending on the frequency.

Vectors are selected depending on the objective of cloning. High-copy number plasmids are preferred if the target is to clone and express a gene for the synthesis of distinct proteins or secondary metabolites in bacterial cell systems for enhanced production; low copy number plasmids are used in transformation experiments.

The advantage of utilizing plasmids as tools in genetic engineering and biotechnology is due to the ease with which they can be modified and replicated. Plasmids are constructed in the laboratory for genetic engineering purposes.

- They are constructed by using an origin of replication, cloning site and selection markers.
- Restriction enzymes act as molecular scissors, cut the DNA at specific points, and produce sticky ends of the DNA.
- The genes of interest are then inserted and ligated by using DNA ligase.
- By transformation process, recombinant DNA molecules are introduced into the host bacteria cell.
- Recombinant plasmids start replication and multiplication using host DNA polymerases.
- pBR322 and pUC18/19 are cloning plasmids commonly used for genetic engineering by transformation in *E. coli* host cells.

1. pBR322

pBR322 is the most commonly used *E. coli* cloning vector. It was constructed at the University of California, San Francisco, CA, USA, by Bolivar and Rodriguez in the year 1977 in Herbert Boyer Laboratory. The alphabet p stands for "plasmid" B for "Bolivar" and R for "Rodriguez."

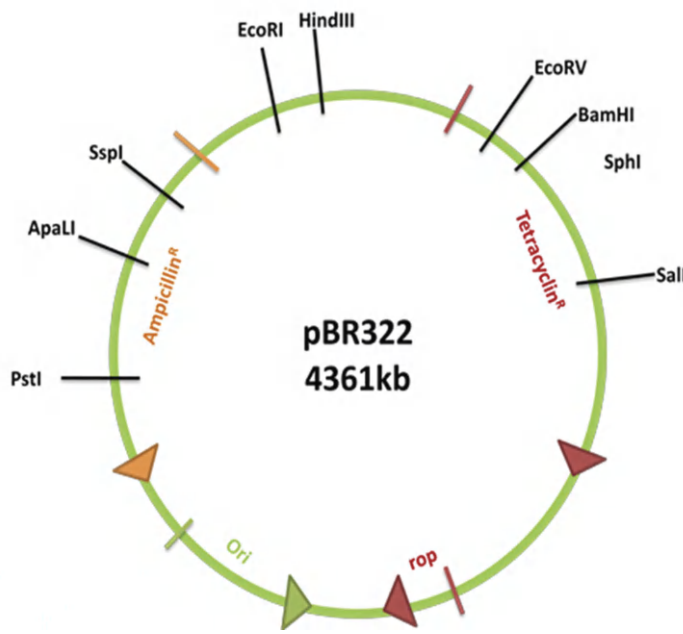


FIGURE 5.3 Structure of pBR322.

Source: https://mlsu.ac.in/econtents/150_Vectors%20NPETL%20Notes.pdf

pBR322 is 4361 base pairs in size and is constructed using the following elements:

- The "*rep*" gene is the replicon region obtained from pMB1plasmid. This region initiates the replication process.
- The "*rop*" gene is chosen from the pMB1 plasmid and codes for the Rop protein. The Rop proteins decrease copy number and confer stability.
- *Bam*HI, *Hin*dIII, *Sal*I, *Pvu*I, *Pvu*II, *Pst*I, *Eco*RI, and *Cla*I are the restriction sites attached.
- Ampicillin and tetracycline resistance genes are the two selectable markers.
- The "*tet*" gene codes for a protein that confers tetracycline resistance and is derived from pSC101 plasmid.
- "*bla*" gene codes for βlactamase that confers ampicillin resistance and is obtained from transposon Tn3.

Advantages of Plasmids as cloning vectors:

- Because of its small size, it is easy to manipulate and isolate.
- As it has a circular configuration, it is more stable.
- It replicates on its own, independent of the host.
- It replicates rapidly, producing a high copy number.
- Due to the presence of antibiotic-resistant genes, detection is easy.

Disadvantages of using Plasmids as vectors:

- The size of the foreign gene is a limiting factor, as large fragments cannot be cloned.
- Its size varies and is only up to 10 kb.
- In some instances, standard transformation methods are inefficient.

2. pUC plasmids

pUC series of cloning plasmid vectors were constructed at the University of California by Messing and co-workers. The letter "p" name stands for plasmid, and "UC" for the University of California. These plasmids are small in size, copy number is high, with a size of 2686 bp.

pUC vectors are constructed by ligating an MCS within the *lacZ* sequence. This region has restriction sites for a broad spectrum of restriction endonucleases and permits rapid visual detection of an insert.

Variations in pUC18 and pUC19 vectors are due to the fact that the MCS is arranged in opposite orientations.

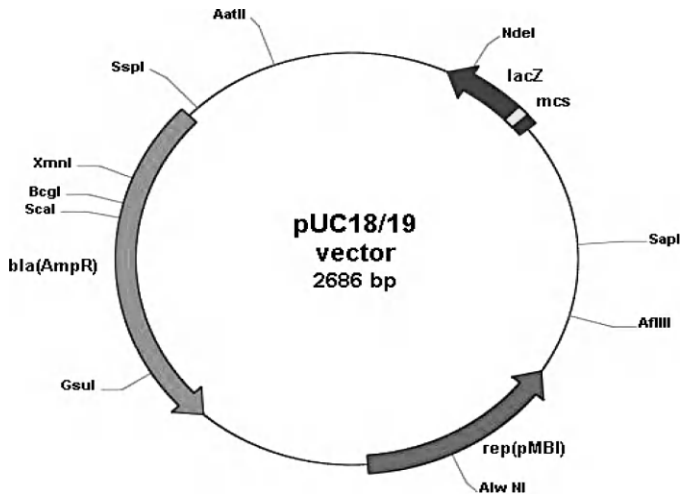


FIGURE 5.4 Structure of pUC plasmids.

Source: https://mlsu.ac.in/econtents/150_Vectors%20NPETL%20Notes.pdf

pUC vectors are constructed using the following elements:

- Replicon region pMB1 "*rep*" is derived from plasmid pBR322 with single point mutation to enhance the copy number.
- Ampicillin resistance providing region "*bla*" gene encoding for βlactamase is derived from pBR322. They differ from pBR322 by two-point mutations.
- Lacoperonsystem is derived from *E. coli*.
- To increase the copy number, "*rop*" gene is removed.
- A short DNA sequence called the MCS region consisting of restriction sites for many restriction endonucleases is attached.

A number of potential cloning strategies are made available by extending the range of enzymes that can be used to generate a restriction fragment suitable for cloning at the MCS region. By ligating the gene of interest

within an MCS, the sites are made adjacent. This enables any two sites within it to be cleaved simultaneously without cutting vector sequences. MCS is inserted next to the gene, which encodes for a promoter as well as α-peptide of β-galactosidase in the *lac Z* region. However, such insertion does not affect the ability of the α-peptide to mediate complementation while cloning DNA fragments into the MCS. Bacterial cells containing the recombinant plasmids can be easily visualized by the blue/white colony screening on culture medium incorporated with X gal and inducer IPTG molecules.

5.3.2.3.1.1 Ti plasmid

A. tumefaciens are gram-negative bacterial plant pathogens that infect and cause tumors in many dicotyledonous *plants. The bacterium* causes crown gall disease, *and it is due to the presence of a plasmid* known as tumor-inducing or Ti plasmid, which causes tumor induction. This plasmid is widely used as a cloning vector to carry through the desirable genes to the host plant to get **transgenic plants.**

Naturally occurring Ti plasmids consist of ~250 kbp of nucleotides.

Ti plasmids are of different types based on the variety of genes they possess, which code for different opines such as nopaline, octopine, and leucinopine, etc.

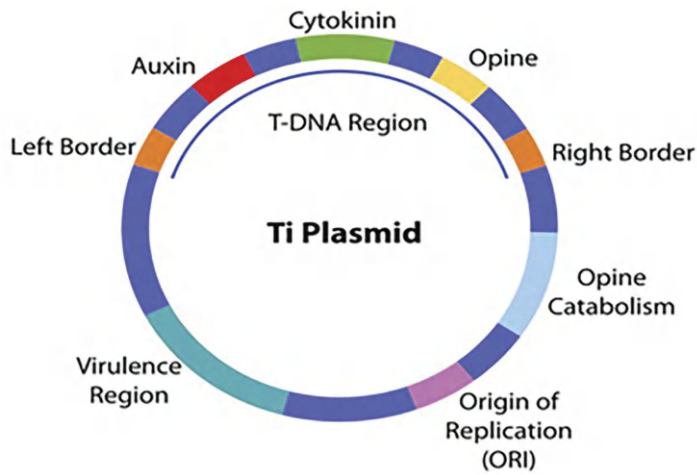


FIGURE 5.5 Structure of Ti plasmid.

Source: Public domain

A. tumefaciens contains one or more regions on T-DNA, which codes for hormones such as auxin and cytokinin that induce cell division. Gall disease occurs due to the transformation of normal plant cells into tumor cells by the insertion of a DNA piece known as T-DNA. Ti plasmids contain *vir* or virulence genes**,** which are responsible for the transfer of T-DNA region onto plant cells. This region gets integrated into the plant genome and starts producing chemicals that are required by the bacterium. Ti plasmids can be engineered as per the requirement to insert the desired gene at the T-DNA region. *A. tumefaciens* are referred to as "nature's genetic engineer."

5.3.2.4 BACTERIOPHAGES AS CLONING VECTORS

Viruses infecting bacterial cells are known as Bacteriophages or phages. Their genetic material can be engineered to deliver the gene of interest into the host bacterial cell. Up to 53 kb of foreign DNA can be packaged into the phage head. If the size of the vector DNA is too small, then it cannot be adequately packaged into the phage particle; for example, phage λ, phage M13, etc.

1. **Phage Lambda λ**

   The structure of phage λ consists of head, tail, and tail fibers. It has 48.5 kb of genomic DNA and 12 bp single-stranded DNA, which comprises sticky ends at both terminals. Single-stranded DNA sequences on either end are complementary to each other; they can form complementary base pairing and are known as "cohesive ends" or "cos sites." Tail fibers of the phage adhere to the bacterial cell surface, followed by contraction of the tail and injection of the DNA inside the host cell.

   Two types of phage vectors are used in genetic engineering techniques:

   i. **Insertion vectors:** Insertional vectors contain a specific cleavage site at which the foreign DNA of up to 5–11 kb can get inserted.

   ii. **Replacement vectors:** The genome of replacement vectors possesses a region flanked by genes that are not important for the host. These genes can be deleted and replaced with foreign DNA inserts.

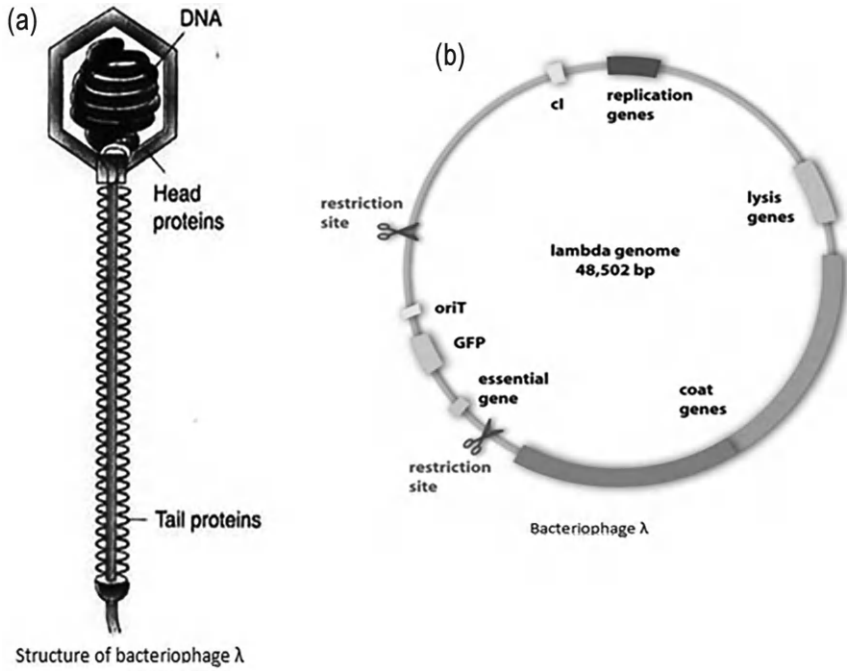


FIGURE 5.6 Structure of bacteriophages λ.

Source: (a) https://www.notesonzoology.com/genetics/cloning-vectors/6-main-types-of-cloning-vectors-biotechnology/5300. (b) https://2008.igem.org/Team:Heidelberg/Human_Practice/Phips_the_Phage/Project_Sensing

Insertion vector: These vectors possess a single restriction site which is used to insert the foreign DNA. Smaller foreign DNA can be packed here; for example, λgt10 and Charon16A.

a. **λgt10,** a double-stranded DNA phage with a total size of about 43 kb, can carry foreign DNA of up to 7 kb. A foreign gene is inserted into a unique *Eco*RI site located in the CI gene. This vector produces clear plaques. Insertional inactivation of this plaque-producing gene results in the recombinants, which produce clear plaques rather than turbid plaques.

b. **Charon 16A** can accommodate foreign DNA up to 9 kb. New DNA gets inserted into a unique EcoRI site. This leads to the inactivation of the lacZ gene carried by the vector. Recombinants produce clear plaques rather than blue plaques when host cells are grown on an X-galagar medium.

Replacement vectors: Their genome consists of a right arm, a left arm, and a central stuffer region. The stuffer region is flanked by two restriction sites on either side. Commonly, the replaceable region carries additional restriction sites that can be used to cut it up into small pieces. As a result, its own insertion is rather unlikely. It can carry large DNA fragments when compared with insertional vectors. For example, EMBL4, Charon 40, and λWES.

a. **EMBL4:** These phage genomes can take up to 9–23 kb of foreign DNA by replacing a segment that contains pairs of *Eco*RI, *Bam*HI, and *Sal*I sites.

b. **Charon 40:**They can accommodate up to 9–22 kb of gene of interest.

c. **λWESB1:** Its genome consists of two *Eco*RI sites situated on either end of the replacement fragment, and based on the size, recombinant selection can be carried out. DNA fragments up to 15 kb can be cloned using this vector.

2. **Phage M13**

Phage M13 has a circular genome with 6.4 kbp of genetic material and is derived from filamentous bacteriophage M13. Single-stranded

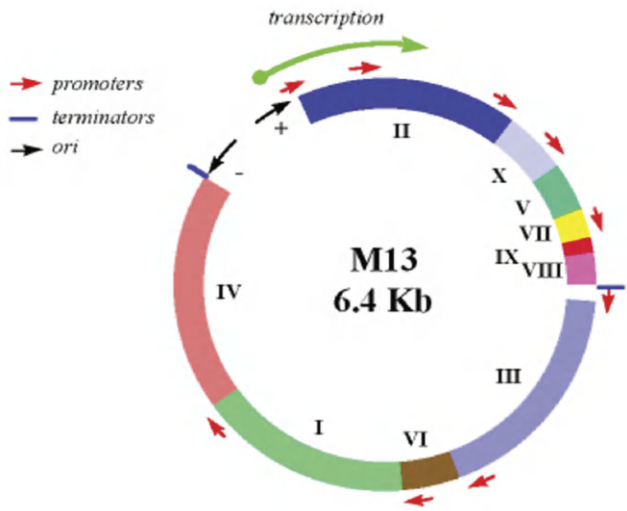


FIGURE 5.7 Structure of Phage M13.

Source: https://bio.libretexts.org/Bookshelves/Biochemistry/Supplemental_Modules_%28Biochemistry%29/4._Biotechnology_2/4.2%3A_M13_Phage

copies of the cloned foreign DNA can be obtained using this vector. As its genomic DNA is very small, large-size inserts of foreign DNA can be cloned, and hence it is used as a potential vector. These vectors are employed for DNA sequencing as well as in vitro mutagenesis studies. Pure single-stranded DNA copies can be obtained from double-stranded inserts.

After the infection of the host *E. coli* cells, phage DNA replicates as double-stranded molecules initially, but subsequently, single-stranded virions are produced. Single-stranded virions so produced further infect the neighboring host bacterial cells. As M13 genomic DNA is double-stranded in its replicative form, it behaves such as a plasmid and is used such as plasmids for invitro experimental purposes.

Advantages of Phage Genome as Vectors

- Large inserts can be cloned using phages and are highly efficient than plasmids.
- Identification of cloned phage plaques is much easier than that of recombinant bacterial colonies.

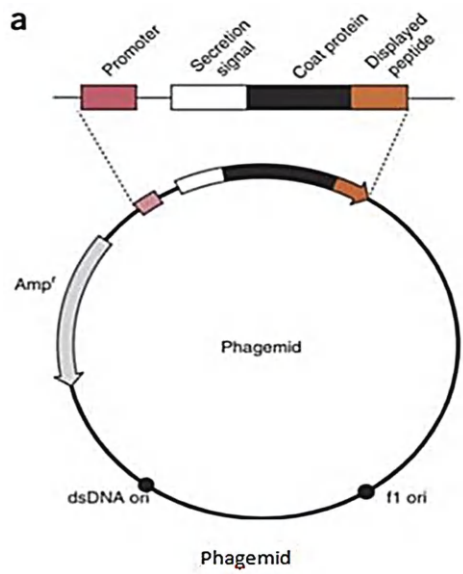


FIGURE 5.8 Structure of Phagemid.

Source: Reprinted with permission from Qi et al., 2012. Copyright © 2012 Elsevier Ltd. Published by Elsevier Ltd.

5.3.2.5 PHAGEMIDS OR PHASMID AS CLONING VECTORS

Phagemids are prepared by artificially engineering the F1 phage origin of replication (F1 region). In combination with the phage M13 genome, they are used as cloning vectors. In the bacterial host, they replicate such as a plasmid, and in the phage particle, they get packaged as single-stranded DNA.

The components of a typical phagemid are:

- An origin of replication from phage M13.
- A part of the lac Z' gene along with the lac promoter.
- Multiple cloning site (MCS) inserted in the lac Z' gene.
- Promoter sequences from T3 and T7 phage on either side of the MCS region.
- Origin of replication from ColE1.
- Ampicillin-resistant gene (amp^R) derived from pBR322.

Advantages of Phagemids as vectors:

- The presence of MCSs.
- Presence of an inducible lac gene promoter.
- Identification of cloned colonies by observing blue–white colonies.

5.3.2.6 COSMIDS AS CLONING VECTORS

Cosmids are constructed using plasmids with the incorporation of bacteriophage λDNA segment. The incorporated segment of DNA contains the cohesive sites (cos sites). For the effective packaging of DNA into phage λ particles, cos sites are essential. The first part of the name of the cosmids indicates it has a cos site from phages, and the last part of the name is due to an origin of replication from pBR 322 plasmid.

The components of a typical phagemid are

- An origin of replication from ColE1.
- amp^R derived from pBR322.
- A polylinker region.
- Cos sites from phage λ particles.
- A unique restriction site for ligation of foreign DNA.

After the cosmid is constructed, they are packaged into the capsid of phage λ to obtain recombinant phage particles and are made to infect *E. coli* cells.

The DNA material is linear when they are packaged inside the phage head and circularizes once it enters inside the *E. coli* cells and acts as a plasmid.

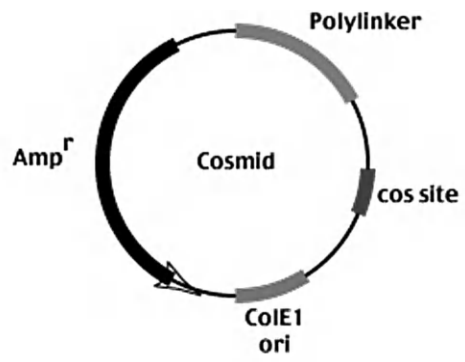


FIGURE 5.9 Structure of Cosmid.

Source: https://faculty.tru.ca/dnelson/courses/biol335/335notes/3recdna/2-vectors/lambda%20vectors/recDNA5c.html

Cosmids can be employed to clone large DNA fragments varying in size from 25 to 45 kb. As it can be packaged into phage λ, foreign DNA or genes can be delivered into the bacterial host by the process of transduction. For example, Col EI cosmid, pWE cosmid, pJB8, pHC 79, etc.

Advantages of cosmids as cloning vectors:

- Large number of clones can be obtained from a small amount of DNA as their transformation efficiency is very high.
- The ability to carry 45 kb of insert makes them a better candidate as vectors when compared with plasmids or phages.

Disadvantages of cosmids as cloning vectors:

- DNA inserts larger than 45 kb cannot be inserted into it.

5.3.2.7 BACTERIAL ARTIFICIAL CHROMOSOMES (BACS) AS CLONING VECTORS

BACs are similar to *E. coli* plasmid vectors in construction. BACs include the origin of replication genes derived from F' plasmid, and these genes code for ori-binding proteins. These proteins are vital for the replication of BACs and its insert carrying capacity varies from 150 to 350 kb.

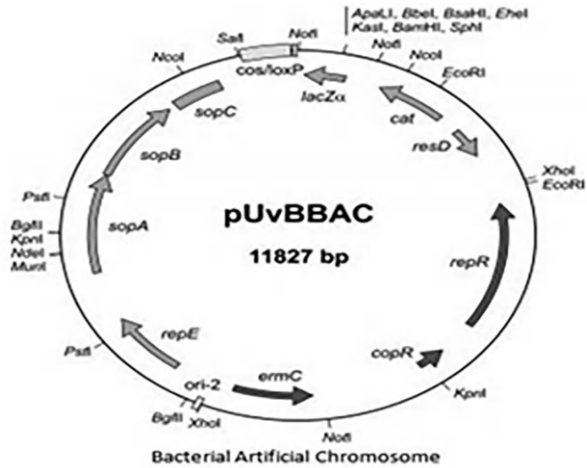


FIGURE 5.10 Structure of bacterial artificial chromosomes.

Source: Reprinted with permission from Hain et al., 2008. Copyright 2008.

The copy number of F-factor origin of replication (*OriS)* is maintained at approximately one copy per cell. Four-factor genes that are essential for replication, as well as maintenance of copy number, are present in BACs. They are *repE*, *parA*, *parB*, and *parC*. An antibiotic-selectable marker, the Lac Z gene with an MCS region that allows for the blue-white screening of BACs containing foreign DNA inserts is present. A *cos* site (*cos*N) and loxP sites that are used for unique cleavage of the insert containing BAC during restriction mapping are also there.

Advantages of BACs as cloning vectors:

- Large sequences of DNA can be cloned, and there is no risk of rearrangement.
- Studies on genetic or infectious disorders are carried out using BACs.
- The yield of DNA clones is very high.

Disadvantages of BACs as cloning vectors:

- Low copy number is a cause of concern.
- They cannot be used for cloning eukaryotic DNA inserts with repetitive sequences.
- Such inserts in BACs are structurally unstable and may result in rearrangement or deletion.

5.3.2.8 YEAST ARTIFICIAL CHROMOSOMES (YACS) AS CLONING VECTORS

YACs are vectors that can be propagated in eukaryotic cells such as yeast as chromosomes and can carry large DNA inserts of up to 200 kb. It essentially contains a centromere sequence (CEN), which aids in the proper segregation of nuclear material during meiosis. It also possesses a bacterial origin of replication (ori) and two telomeric sequences, one at each end. Yeast, as well as bacterial cells, can be used as hosts.

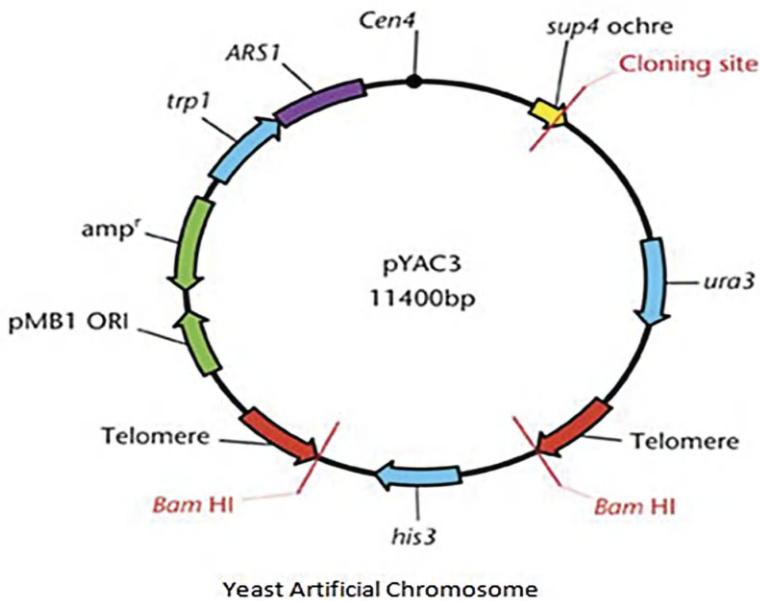


FIGURE 5.11 Structure of yeast artificial chromosomes.

Source: Reprinted with permission from Tosato et al., 2012. Copyright © 2012 John Wiley & Sons, Ltd.

Foreign genomic DNA of approximately 500 kbp in size can be ligated into YACs. The presence of yeast genomic components such as yeast centromere (CEN4), autonomously replicating sequence (ARS1), yeast telomere (TEL), genes for YAC selection in yeast, a bacterial origin of replication (ori), and selectable antibiotic resistance marker (amp^r), makes it a unique and most popular vector. A three-element consensus sequence of 78–86 bp containing more than 90% of AT residues with a conserved sequence on one

side and a short consensus sequence on the other side is present. At the end of each chromosome, multiple copies of telomeric sequences (5′-TGTGGGT-GTGGTG-3′) essential for replication and chromosome maintenance are present. YACs have a functional copy of URA3, a gene involved in uracil biosynthesis, and a Trp1 gene essential for tryptophan biosynthesis, which aids in the selection of transformed yeast cells.

Advantages of YACs as vectors:

- YACs can take up large amounts of foreign DNA.
- YACs can be used to construct physical maps of large genomes such as the human genome.

Disadvantages of YACs as vectors:

- YACs have very low efficiency of transformation.
- YACs have a low yield of cloned DNA.

KEYWORDS

- **restriction enzymes**
- **cloning**
- **vectors**
- **plasmids**
- **phages**
- **phagemids**
- **cosmids**
- **bacterial artificial chromosomes**
- **yeast artificial chromosomes**

REFERENCES

Arber, W.; Linn, S. DNA Modification Andrestriction. *Ann. Rev. Biochem.* **1969,** *38*, 467–500.

Asami, J.; et al. Bacterial Artificial Chromosomes as Analytical Basis for Gene Transcriptional Machineries. *Transgenic Res.* **2011,** *20* (4), 913–24. DOI: 10.1007/s11248-010-9469-3.

Bhatia, S.; Goli, D. In *Introduction to Pharmaceutical Biotechnology*; IOP Publishing Ltd., 2018.

Botstein, D.; White, R. L.; Skolnick, M.; Davis, R. W. Construction of a

Genetic Linkage Map in Man using Restriction Fragment Length Polymorphisms. *Am. J. Hum. Genet.* **1980,** *32* (3), 314.

Brown, T. A. In *Gene Cloning and DNA Analysis*: *An Introduction*, 6th Ed.; *Wiley- Blackwell,* 2010.

Elshire, R. J.; Glaubitz, J. C.; Sun, Q.; Poland, J. A.; Kawamoto, K.; Buckler, E. S.; Mitchell, S. E. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS One* **2011,** *6* (5), e19379.

Griffiths, A. J. F.; Gelbart, W. M.; Miller, J. H. In *Modern Genetic Analysis*; W. H. Freeman: New York, 1999.

Hain T., Sonja O., Ulrich B., Chatterjee S. S., Technow U, Billion A, Ghai R, Mohamed W, Domann E, Chakraborty T. Novel Bacterial Artificial Chromosome Vector pUvBBAC for Use in Studies of the Functional Genomics of Listeria spp. Appl Environ Microbiolgy. 2008, 74. https://doi.org/10.1128/AEM.00415–07

Kessler, C.; Manta, V. Specificity of Restriction Endonucleases and DNA Modification Methyl Transferases are View (Edition 3). *Gene* **1990,** *92* (1–2), 1–248.

Krüger, D. H.; Bickle, T. A. Bacteriophage Survival: Multiple Mechanisms for Avoiding the Deoxyribonucleic Acid Restriction Systems of their Hosts. *Microbiol. Rev.* **1983,** *47* (3), 345–360.

Ladish, H.; Berk, A.; Zipursky, S. L.; Matsudaira, P.; Baltimore, D.; Darnell, J. DNA Cloning with Plasmid Vectors. In *Molecular Cell Biology*, 4th Ed.; W H Freeman: New York, 2000.

Lewin, B. In *Genes VII*; Pearson Prentice Hall, Pearson Education, Inc.: Upper Saddle River, NT, 2004.

Lodish, H. In *Molecular Cell Biology*, 4th Ed.; Freeman, 2000.

Moo Young, M. In *Comprehensive Biotechnology*; Second Ed.; Elsevier, 2011.

Murray, A. W.; Szostak, J. W. Artificial Chromosomes. *Sci. Am.* **1987,** *257* (5), 62–68.

Old, R. W.; Primrose, S. B. In *Principles of Gene Manipulation - An Introduction to Genetic Engineering*; Blackwell Scientific Publications, 1993.

Qi H., Lu H., Qiu H., Petrenko V., Liu A. Phagemid Vectors for Phage Display: Properties, Characteristics and Construction; *J.Mol. Bio. 2012,* 417, 129–143.

Primrose, S. B.; Twyman, R. N.; Old, R. W. In *Principles of Gene Manipulation*, 5th Ed.; Blackwell Publishing, 2006.

Roberts, R. J.; Vincze, T.; Posfai, J.; Macelis, D. REBASE – Enzymes and Genes for DNA Restriction and Modification. *Nucleic Acids Res.* **2007,** *35*, D269–D270.

Rodriguez, R. L.; Denhardt, D. T. In *Vectors: A Survey of Molecular Cloning Vectors and Their Uses (Biotechnology)*, Kindle Edition; Butterwort Publishers, 1987.

Rosenberg, E. In *It's in Your DNA: From Discovery to Structure, Function and Role inEvolution, Cancer and Aging*; Academic Press, 2017.

Tosato, V.; Arnak, R.; Bruschi, C. Yeast Artificial Chromosomes. 2012.10.1002/9780470015902.a0000379.pub3.

Williamson, R. In *Genetic Engineering*; Academic Press, 1981; vol *I–IV*.

Van Embden, J. The Use of Cosmids as Cloning Vehicles. In *Techniques in Molecular Biology*; Walker J. M., Gaastra W., Eds.; Springer: Dordrecht, 1983.



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CHAPTER 6

Restriction Enzymes and Vectors

MOHAMMAD AMIN LONE

Department of Zoology, Government Degree College Baramulla (Autonomous), Jammu and Kashmir, India

ABSTRACT

Genetic engineering is the process of manipulating an organism's DNA to alter its characteristics. The key tools used in this field are restriction enzymes and vectors. Restriction enzymes are enzymes that cut DNA at specific sequences, also referred to as recognition sites. Because they allow for precise DNA nipulation, these enzymes are commonly used in genetic engineering techniques such as cloning and PCR. Vectors, on the contrary, are vehicles used to transfer DNA into a host organism. Plasmids and viral vectors are the most commonly used vectors in genetic engineering. Plasmids are small, circular, and extrachromosomal DNA molecules found in bacteria, while viral vectors are viruses that have been genetically modified to carry and transfer foreign DNA into cells. Both plasmids and viral vectors are generally used to introduce new genes into organisms for research or biotechnology applications. In this chapter the authors provide a detailed overview of the different types of restriction enzymes and vectors used in genetic engineering, as well as their specific applications and limitations. They also discuss the latest advancements and future directions in the field of genetic engineering using restriction enzymes and vectors. Overall, restriction enzymes and vectors are essential tools in genetic engineering, enabling scientists to manipulate the DNA of organisms with precision and accuracy.

Molecular Genetics, Structures, Mechanisms, and Functions, Volume 1: Principles of Gene Manipulation and Genomics. Tariq Ahmad Bhat, Abdul Rauf Shakoori, and Jameel M. Al-Khayri (Eds.)
© 2026 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

6.1 INTRODUCTION

Genetic engineering is the process of manipulating the genetic material of an organism to alter its characteristics. This can be done for various purposes, including medical research, crop improvement, and biotechnology. Many different tools that are used in genetic engineering, each with its unique benefits and limitations. The most widely used molecular tools of genetic engineering include restriction enzymes, ligases, polymerase chain reaction (PCR), and DNA sequencing.

6.2 HISTORICAL BACKGROUND

By the late 1960s, scientists working in the field of molecular biology had met most of the criteria needed to move forward with genetic technology. Nonetheless, they were still struggling due to the technical limitations of the research. The discovery of DNA ligase, which serves as a type of "molecular glue" by joining two strands of DNA, was isolated in 1967 by Gellert, Lehman, Richardson, and Hurwitz laboratories and was a significant breakthrough in molecular biology (Shuman, 2009). Another breakthrough came in 1970 when the first restriction enzyme was isolated, opening up possibilities for genetic engineering as it serves as a "molecular scissors" by cutting DNA at specified sequences. These enzymes can produce DNA fragments suitable for association with other fragments. Since then, researchers have continued to make advancements in the field, such as the discovery of gene cloning, which has enabled scientists to produce thousands of copies of a specific gene for further study. Additionally, in 1973, the first genetic engineering experiment was conducted, in which scientists successfully inserted a gene from one organism into another. Furthermore, in 1983, the first genetically modified organism (GMO) was created, paving the way for further developments in the field. Thus, by 1970, the basic tools needed to build recombinant DNA were available.

6.3 TOOLS OF THE TRADE

An engineer designs structures (for example, bridges, canals, and railways) and works according to an established plan. The term genetic engineering may be appropriate for someone involved in genetic manipulations. The toolbox or molecular tools of genetic engineering, which are the enzymes most commonly used in recombinant DNA experiments, are briefly described in Figure 6.1:

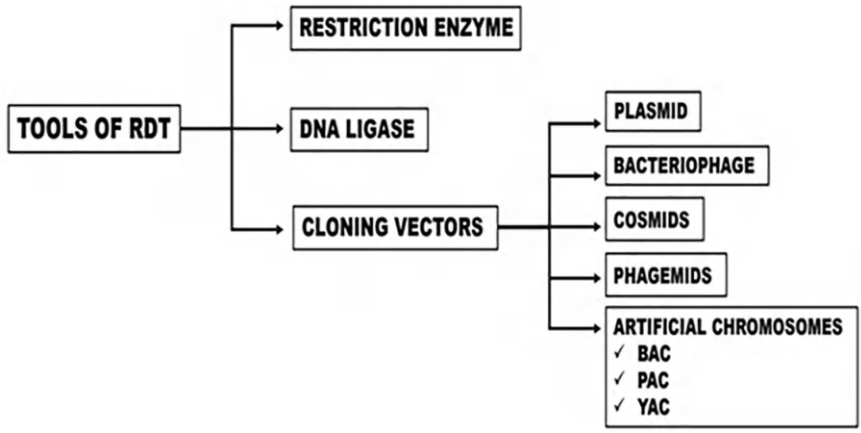


FIGURE 6.1 Tools of rDNA technology.

6.3.1 RESTRICTION ENZYMES—DEFINITION AND DISCOVERY

Restriction enzymes are a type of molecular tool that has revolutionized the field of genetic engineering. These enzymes can cut and manipulate DNA at specific sequences, allowing scientists to manipulate and modify genes with great precision. This technology has been instrumental in the development of recombinant DNA technology and has led to many breakthroughs in medicine, agriculture, and biotechnology.

The discovery of restriction enzymes was made in the 1970s by researchers at the University of Geneva. They studied the mechanisms of bacterial resistance to viral infection and discovered that certain bacteria produce enzymes that can cut and inactivate viral DNA. These enzymes were named restriction enzymes, and they quickly became one of the most important tools in the field of genetic engineering.

Scientists can use these enzymes to cut and manipulate DNA to add, delete, or replace specific genes. These enzymes were discovered in bacteria and function as a defense mechanism against invading viral DNA. Restriction enzymes are often used in recombinant DNA technology, which allows scientists to combine DNA from different sources. This is done by cutting the DNA with restriction enzymes and then joining the pieces together using an enzyme called ligase. This is the basic principle behind the creation of GMOs. For example, scientists can use restriction enzymes to cut a gene for herbicide resistance from one organism and insert it into another, creating a plant that is resistant to herbicides. The stable integration of functional genes into plant genomes has both scientific and practical applications. This

technique provides opportunities for crop improvement. By introducing functional genes that confer desirable traits, such as resistance to pests or drought, scientists can create genetically modified crops that are more resilient and productive, which can ultimately benefit agriculture and food production (Fraley, et al., 1986; Goodman et al., 1987).

6.3.1.1 RESTRICTION MODIFICATION SYSTEM

The restriction enzymes were originally discovered from bacteria. Indeed, this system has evolved in prokaryotes to protect the host bacteria from viral and pathogenic invasion. As soon as it recognizes alien DNA through its recognition site, it cleaves and destroys the pathogen, virus, or phage.

Although it can cut any type of DNA with its recognition sites, it cannot cut its own DNA due to a special mechanism—known as "methylation," which is also performed by restrictases. Bacterial DNA is protected from restriction enzymes by methylation. Another enzyme called "methyl transferase" inserts methyl groups at the site where the REase acts, preventing them from cutting their own DNA. The methyl group insertion precludes self-restriction digestion.

More than 600 distinct restriction enzymes are utilized in genetic engineering and molecular cloning operations; however, only 300 REase have been thoroughly explored.

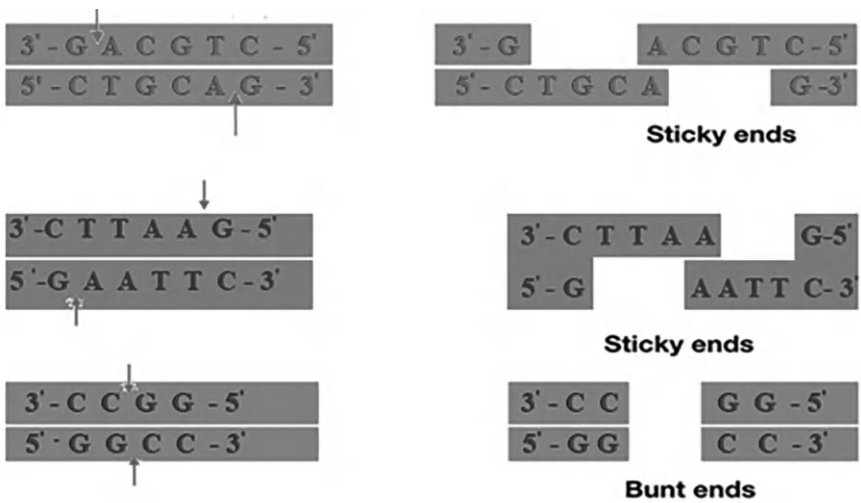


FIGURE 6.2 Recognition sequences. ⏎

6.3.1.2 RECOGNITION SEQUENCES

The sequence recognized by the restriction enzyme to cut the DNA is called the "restriction site or recognition site" (Fig. 6.2). the "recognition site" is predominantly GC-rich. It consists of 4–8 base pairs. The ability of an enzyme to recognize a particular sequence is called "sequence specificity." These short symmetric sequences of bases are termed "Palindromic sequences." In "palindromic sequences," the base composition remains the same in both strands of DNA but in the opposite direction, that is, sequences of bases are the same when read in 5'- 3' direction on both DNA strands (Fig. 6.3). It is similar to palindromic word which reads the same in both the directions, for example, M A D A M, MALYALAM.

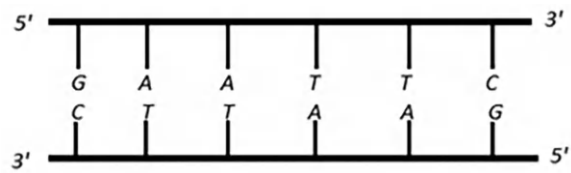


FIGURE 6.3 A palindromic sequence. ⏎

Some enzymes break down DNA after the recognition site, while others do so inside the site, outside of it, or very close to it. The enzyme-restriction endonuclease generally uses the recognition site as a "marker" to identify the cleaving site. Sticky ends and blunt ends are the two types of sequences that are produced during restriction digestion. In experiments involving genetic engineering, sticky-end-producing enzymes are highly recommended.

TABLE 6.1 Some Examples of Recognition Sites.

Restriction endonuclease	Recognition site	Source	Blunt ends or sticky ends
EcoR1	5′—G/AATTC—3′ 3′—CTTAA/G—5′	*Escherichia coli* strain RY13	Sticky ends
EcoRV	5′—GAT/ATC—3′ 3′—CTA/TAG—5′	*E. coli*	Blunt ends
PstI	5′—CTGCA/G—3′ 3′—G/ACGTC—5′	*Providencia stuartii*	Sticky ends
BamHI	5′—G/GATCC—3′ 3′—CCTAG/G—5′	*Bacillus amyloliquefaciens*	Sticky ends

TABLE 6.1 *(Continued)*

Restriction endonuclease	Recognition site	Source	Blunt ends or sticky ends
BccI	5′—CCATC(N)/—3′ 3′—GGTAG(N)/—5′	*Bacteroides caccae*	Blunt ends
HaeII	5′—GCGC/—3′ 3′—/CGCG—5′	*Haemophilus aegyptius*	Sticky ends
HaeIII	5′—GG/CC—3′ 3′—CC/GG—5′	*H. aegyptius*	Blunt ends
HindIII	5′—A/AGCTT—3′ 3′’—TTCGA/A—5′	*Haemophilus influenzae*	Sticky ends
HinfI	5′—G/ANTC—3′ 3′—CTNA/G—5′	Recombinant *E. coli* strain	Sticky ends
NIaIII	5′—CATG/—3′ 3′—/GTAC—5′	*Neisseria lactamica*	Sticky ends

6.3.1.3 TYPES OF RESTRICTION ENZYMES

The formation of particular DNA fragments is achieved by restriction enzymes, also referred to as restriction endonucleases, which cut DNA at specific locations. These enzymes serve various purposes in molecular biology, including genome analysis, DNA fingerprinting, and genetic engineering (Roberts and Kim, 2016).

There are several types of restriction enzymes, each with distinct properties and recognition sequences. The most commonly used restriction enzymes are classified into four types: types I, II, III, and IV (Sambrook and Russel, 2001).

6.3.1.3.1 Type I Restriction Enzymes

These enzymes are large, multi-subunit enzymes that recognize and cut DNA at specific sequences. These enzymes require the presence of a specific methylase enzyme to methylate the DNA before they can bind and cut it (Roberts and Kim, 2016). Examples of Type I restriction enzymes include EcoRI and EcoRV.

6.3.1.3.2 Type II Restriction Enzymes

These enzymes are the most commonly used restriction enzymes and are responsible for the majority of the research in molecular biology. These

enzymes recognize specific sequences and cut DNA at or near the recognition site (Sambrook and Russel, 2001). Examples of type-II restriction enzymes include EcoRI and BamHI.

6.3.1.3.3 Type III Restriction Enzymes

These enzymes are complex enzymes that require the presence of a specific accessory protein called a helicase to unwind the DNA before it can bind and cut it (Roberts and Kim, 2016). Examples of type-III restriction enzymes include EcoRII and EcoRIII.

6.3.1.3.4 Type IV Restriction Enzymes

These rare enzymes recognize and cut DNA at specific sequences, but they do not require the presence of a specific methylase enzyme or accessory protein (Sambrook and Russel, 2001). Examples of type-IV restriction enzymes include EcoB and EcoC.

6.3.1.4 NOMENCLATURE OF RESTRICTION ENZYMES

The organism from which they were first isolated and the particular recognition sequence that they bind to are used to give restriction enzymes their names. For example, EcoRI was first isolated from *E. coli* and recognizes the sequence "GAATTC" (Hedge, 2018). Similarly, BamHI was first isolated from *B. amyloliquefaciens* and recognizes the sequence "GGATCC" (Sambrook and Russel, 2001). Depending on the bacterial strain involved and the presence or absence of extrachromosomal components, additional descriptors might be added. Roman numeral denotes the order of discovery (Nicholl, 2008; Roberts and Murray, 1976).

A couple of examples are given below:

- EcoRI is from *E. coli*: E = *Escherichia*, co = *coli*, R = strain, and I = first.
- HindIII is from *H. influenza*: H = Haemophilus, in = influenza, d = strain, and the third endonucleases (III) to be discovered.
- Sau3A is isolated from *Staphylococcus* (S) *aureas* (au), strain 3A.

6.3.1.5 APPLICATIONS OF RESTRICTION ENZYMES

Recombinant DNA creation is one of the most significant uses of restriction enzymes. Recombinant DNA is a term used to describe DNA that has been artificially created by joining together fragments from different sources. The use of restriction enzymes allows scientists to cut DNA from different sources and then join them together using a process called ligation. This technology has led to many breakthroughs in medicine, agriculture, and biotechnology.

In medicine, restriction enzymes have been used to create GMOs that can produce life-saving drugs and vaccines. For example, scientists have used restriction enzymes to create bacteria that can produce human insulin, which is used to treat diabetes. Similarly, restriction enzymes have been used to create genetically modified plants that produce drugs and vaccines. This technology can potentially revolutionize how we produce and deliver drugs and vaccines, making them more accessible and affordable.

In agriculture, restriction enzymes have been used to create genetically modified crops more resistant to pests and diseases. This has increased crop yields and improved food security in many parts of the world. For example, scientists have used restriction enzymes to create crops resistant to pests and diseases, such as corn resistant to the European corn borer. Additionally, restriction enzymes have been used to create crops more tolerant to environmental stressors, such as drought and heat.

In biotechnology, restriction enzymes have created GMOs that can produce industrial enzymes and chemicals. For example, scientists have used restriction enzymes to create bacteria that can produce enzymes that are used to produce biofuels. Similarly, restriction enzymes have been used to create plants that can produce chemicals and materials used in producing plastics and other products.

In conclusion, restriction enzymes are a powerful tool of genetic engineering that has revolutionized the field. These enzymes can cut and manipulate DNA at specific sequences, allowing scientists to manipulate and modify genes with great precision. This technology has been instrumental in the development of recombinant DNA technology and has led to many breakthroughs in medicine, agriculture, and biotechnology. With continued research and development, restriction enzymes will play a vital role in shaping the future of genetic engineering (Smith, 1974).

6.3.2 DNA LIGASE

The DNA ligase serves as a "molecular adhesive" that helps join the ends of two double-stranded DNA molecules to generate a longer strand. However, it cannot add nucleotides to fill the gaps in the DNA. It fills the gap by forming a covalent bond between the 5'-phosphate group and the 3'-OH group in the gap. The bond formed between the 5'-phosphate and the 3'-OH groups is known as a phosphodiester bond. If one or more nucleotides are missing or the 5'-phosphate group is absent in the gap, DNA ligase can never seal the gap. This function of DNA ligase is critical for accomplishing many experiments, making it a key enzyme in genetic engineering. "DNA ligases" were first discovered by Gellert, Lehman, Richardson, and Hurwitz in 1967 (Shuman, 2009). The most popular DNA ligase derived from the T4 bacteriophage is a monomeric polypeptide with a molecular weight of 68,000 Da and is encoded by the bacteriophage gene 30. Under the right circumstances, this enzyme can successfully join DNA molecules that are cohesive and blunt ended. This enzyme is frequently used at much lower temperatures (4–15°C) to prevent thermal denaturation of the short base-paired regions that hold the cohesive ends of DNA molecules together (Shuman, 2009).

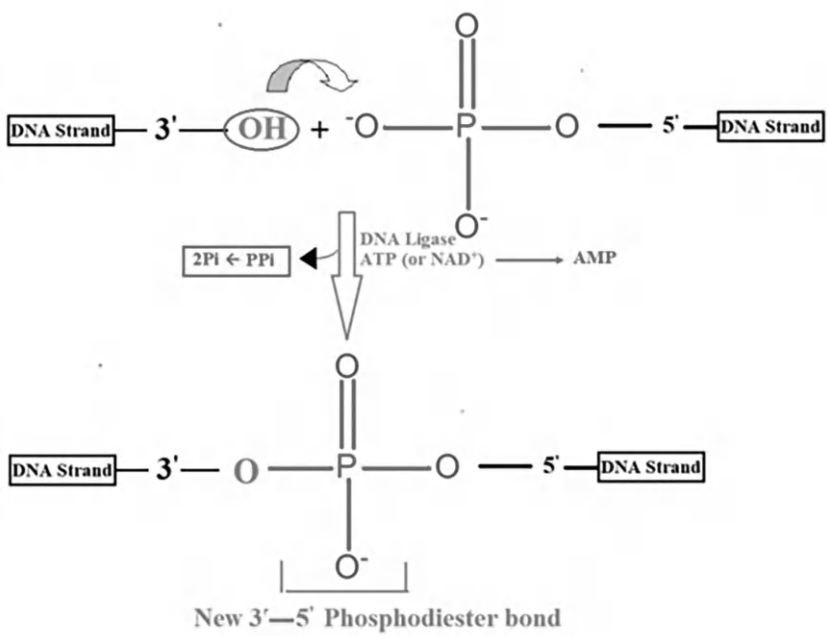


FIGURE 6.4 DNA ligase reaction.

6.3.2.1 USES OF DNA LIGASES

DNA ligase has several important uses in molecular biology and genetic engineering. Here are some of its primary uses:

Joining DNA fragments: To create a longer DNA molecule, DNA ligase is frequently used to join together two or more DNA fragments, either with cohesive or blunt ends. This is an important step in many molecular biology techniques, such as cloning, gene synthesis, and DNA sequencing.

Creating recombinant DNA: DNA ligase inserts a gene of interest into a plasmid or other vector, creating recombinant DNA that can be used for various applications, such as gene therapy or genetically modified crops.

Site-directed mutagenesis: DNA ligase can be used in site-directed mutagenesis to introduce specific mutations or changes into a DNA sequence. This is important for studying the effects of different mutations on gene expression and protein function.

Polymerase chain reaction (PCR): DNA ligase is sometimes used in PCR to "seal" the ends of PCR fragments, preventing the formation of "sticky ends" that can anneal to each other and interfere with downstream applications.

Overall, DNA ligase is a critical tool for molecular biologists and genetic engineers, enabling the manipulation and engineering of DNA molecules in a precise and controlled manner (Shuman, 2009).

6.3.3 CLONING VECTORS

Genetic engineering is the manipulation of an organism's genetic makeup using biotechnology. One of the key tools used in this process is vectors, which introduce alien DNA into the organism. The following paragraph discusses the different types of vectors used in genetic engineering, their properties, and their applications.

6.3.3.1 PROPERTIES OF AN EFFICIENT VECTOR

rDNA technology is a method of manipulating and modifying DNA molecules to produce specific desired outcomes. One important aspect of rDNA technology is using vectors, which are DNA molecules that can replicate independently and carry foreign DNA into host cells. For a vector to be efficient and ideal in rDNA technology, it must possess certain properties.

1. One important property of an efficient vector is its ability to replicate and maintain itself in the host cell. This is crucial for the stable and

long-term expression of the foreign DNA. Vectors that can replicate episomally, meaning they are not integrated into the host cell's genome, are considered more efficient as they can be easily removed from the host cell if necessary (Capecchi, 1989).

2. An efficient cloning vector must always contain the origin of replication (ORI) site, which is the location where DNA replication initiates.
3. It should have suitable marker genes. A selectable marker gene in a cloning vector is important to identify and select cells that have successfully taken up the vector during cloning. The gene confers resistance to a toxic agent, and only cells that have taken up the vector and the marker gene can survive. This simplifies the screening process in cloning experiments by enriching the cells containing the desired products.
4. A suitable cloning vector should be tiny and light in weight (<10 kb) to fit into the "target DNA" with ease.
5. Another property of an efficient vector is its ability to deliver the foreign DNA into the host cell efficiently. This is achieved through specific recognition sites, called restriction sites, for restriction enzymes. These enzymes are used to cut the vector and the foreign DNA at specific locations, allowing them to be inserted into the vector in a controlled manner (Sambrook et al., 1989).

6.3.3.2 TYPES OF CLONING VECTORS

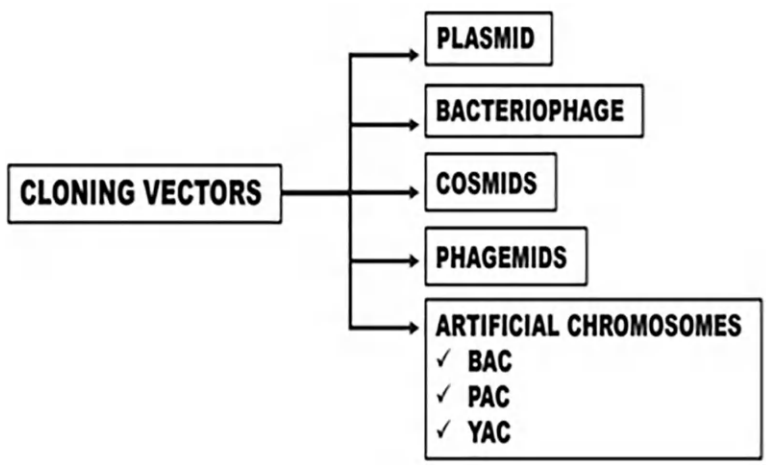


FIGURE 6.5 Types of cloning vectors

6.3.3.2.1 Plasmids

A third important tool in genetic engineering is the plasmid. A plasmid is a small circular piece of DNA that can exist independently of the chromosomes. Plasmids can be used to transfer genes between different organisms, and they are often used in genetic engineering to introduce new genes into cells. Plasmids are also used to produce large amounts of a specific protein, such as human insulin. To make a protein, a gene encoding the protein is inserted into a plasmid, which is then introduced into a cell. The cell uses the information in the gene to produce the protein, which can then be collected and used for medical or research purposes (Cohen et al., 1973).

Plasmid vectors are among the most frequently used vectors in genetic engineering. They are small, circular pieces of DNA found naturally in bacteria. They are easy to manipulate and can be easily introduced into bacteria or other organisms. Plasmids are used to introduce genes into bacteria to form recombinant proteins and other bioproducts.

One of the advantages of plasmid vectors is that they can replicate independently of the host chromosome, which means that they can exist as multiple copies in a single cell. This is useful for increasing the expression of a particular gene.

Furthermore, plasmids can be engineered to contain antibiotic resistance markers, allowing for easy identification of cells that have taken up the plasmid. Examples of plasmid vectors include pBR322 and pUC18 (Sambrook et al., 1989). These vectors have been widely used in genetic engineering and have been used to produce a variety of recombinant proteins, including enzymes, hormones, and antibodies.

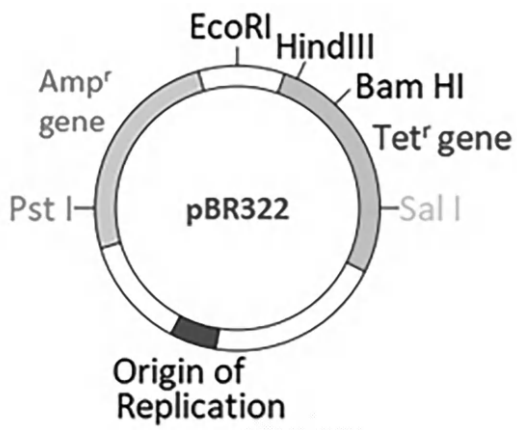


FIGURE 6.6 Plasmid pBR322.

6.3.3.2.2 Phage Vectors

Phage vectors are another type of vector used in genetic engineering. They are based on bacteriophages, which are viruses that infect bacteria. Phage vectors can introduce foreign DNA into bacteria by using the phage's natural mechanism of infection.

One of the advantages of phage vectors is that they can carry large amounts of DNA, up to 20 kb, which is much larger than plasmid vectors. Additionally, phage vectors can introduce DNA into a wide range of bacteria, including Gram-negative and Gram-positive bacteria.

Examples of phage vectors include lambda phage (Fiers et al., 1976) and M13 phage (Messing, 1982). These vectors have been used to introduce genes into bacteria to produce recombinant proteins and other bioproducts.

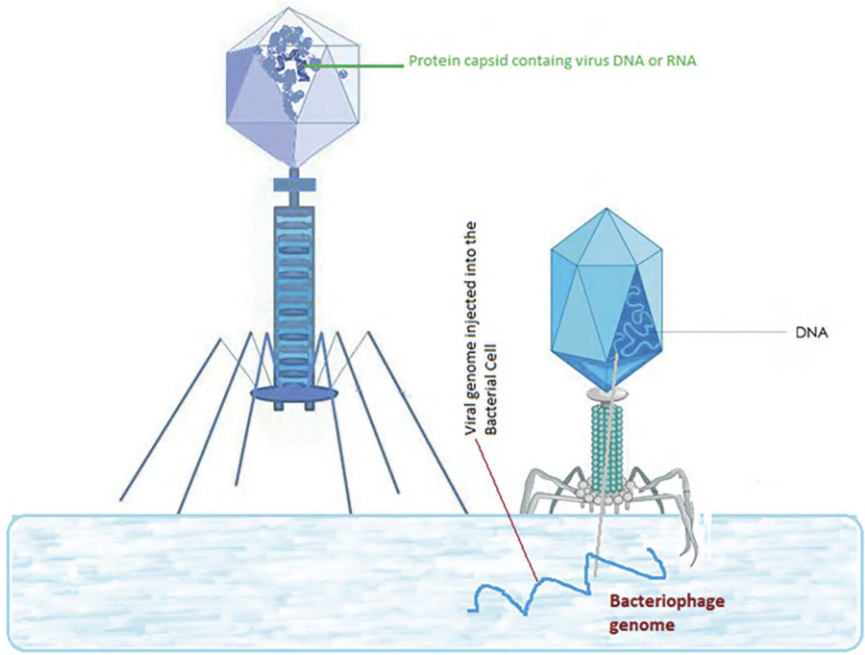


FIGURE 6.7 A bacteriophage introducing genome into bacterial cell.

Source: Modified from Graham Colm. https://creativecommons.org/licenses/by-sa/3.0

6.3.3.2.3 Cosmids

Cosmids are similar to plasmids but are larger and can carry larger pieces of DNA. They often introduce large pieces of DNA, such as entire chromosomes, into an organism. It clones large DNA fragments between 35 and 45 Kb.

6.3.3.2.4 Phagemids

Phagemids are vectors that have both plasmid and M13 elements. These provide an alternative method for obtaining single-stranded (SS) DNA. These are unique because they have two replication origins, one from a standard plasmid and the other from M13. M13 origin is required for ssDNA synthesis.

The synthesis of SS DNA by phagemids, on the contrary, necessitates the use of enzymes and coat proteins encoded by genes not found in them. Cells containing a phagemid vector must, therefore, be co-infected with a helper phage. The phagemid is converted into SS DNA molecules by this helper phage, which is then assembled into phage particles and excreted from the cell (Hoy, 2013).

6.3.3.2.5 Artificial Chromosomes

Artificial chromosomes are genetically engineered DNA molecules designed to replicate and behave like natural chromosomes within cells. They have been developed as vectors for cloning large pieces of DNA and studying chromosome structure and function.

Examples: YACs, PACs, and BACs.

6.3.3.2.5.1 Yeast artificial chromosomes (YACs)

Yeast artificial chromosomes (YACs) were developed in the 1980s to clone and manipulate large fragments of DNA (Burke et al., 1987). They consist of three essential components: a centromere, which allows the YAC to be segregated during cell division; telomeres, which protect the ends of the chromosome from degradation and fusion; and a selectable marker, which allows researchers to identify cells that have successfully taken up the YAC.

YACs have several advantages over other cloning vectors, particularly for large-scale sequencing projects. They can accommodate DNA fragments up to several hundred kilobases in size, making them ideal for sequencing large genomes. Because they are maintained as chromosomes, YACs provide a more accurate representation of genomic organization. They can be used to study complex genetic phenomena, such as gene regulation and chromatin structure.

However, there are also some challenges associated with using YACs. They are more difficult to manipulate than other cloning vectors and can

be unstable, particularly if the DNA fragment contains repetitive sequences. Additionally, YACs have lower transformation efficiency than other vectors, which can limit their usefulness for certain applications.

Overall, YACs have played an essential role in advancing our understanding of chromosome structure and function enabling sequencing several large genomes.

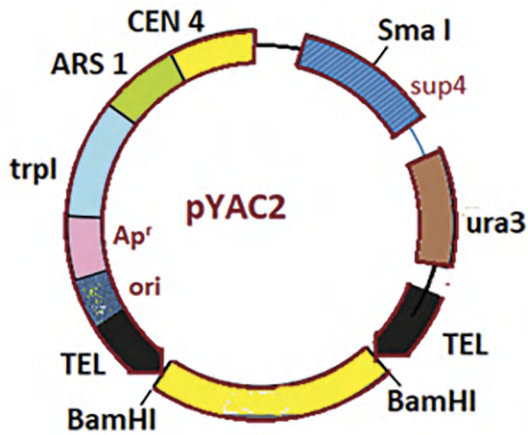


FIGURE 6.8 Yeast artificial chromosome (pYAC2).

6.3.3.2.5.2 P1-derived artificial chromosome (PAC)

P1-derived artificial chromosome (PAC) is a cloning vector created from the DNA of P1 bacteriophages and bacterial artificial chromosomes (BACs) specially designed to carry large DNA fragments of up to 300 kb (Bajpai, 2014).

The PACs have two specific sites, called loxP sites, which can be used by enzymes called phage recombinases during Cre-Lox recombination. This process causes the DNA to form a circularized plasmid that can be inserted into bacterial cells, such as *E. coli*. The transformation of the plasmids into bacterial cells is usually done by electroporation, a technique that uses electricity to facilitate the transfer of plasmids into the cells (Cohen and Sternberg, 1989).

6.3.3.2.5.3 Bacterial artificial chromosomes (BACs).

BACs, also called "backs", are cloning vectors containing the origin of replication from a natural plasmid found in *E. coli* called the F factor, a multiple cloning site (MCS), and one or more selectable markers.

BACs are large DNA molecules modified to serve as cloning vectors for DNA fragments up to several hundred kilobases in size. BACs contain the origin of replication from a natural plasmid found in *E. coli* known as F factor, a MCS, and one or more selectable markers. The selectable markers such as antibiotic resistance genes allow researchers to identify cells that have successfully taken up the BAC, while the presence of MCSs allows DNA fragments to be inserted into the vector at a specific location.

For example, pBeloBAC11 is a commonly used BAC vector similar to a plasmid vector, but with an origin of replication derived from the F factor. This origin limits the copy number of the BAC to one per *E. coli* cell, which helps to maintain the stability of the cloned DNA fragment.

This BAC vector also contains a selectable marker for chloramphenicol resistance and MCS located within a gene, allowing for easy selection of recombinant clones.

BAC vectors are widely used in molecular biology and genetics research for various applications, including genome sequencing, genetic mapping, and gene expression analysis. The large size of the DNA fragments that can be cloned in BACs makes them particularly useful for studying complex genomes, such as those of mammals and plants. BACs can accept inserts of around 300 kb, and many of the instability problems of YACs can be avoided using the bacterial version.

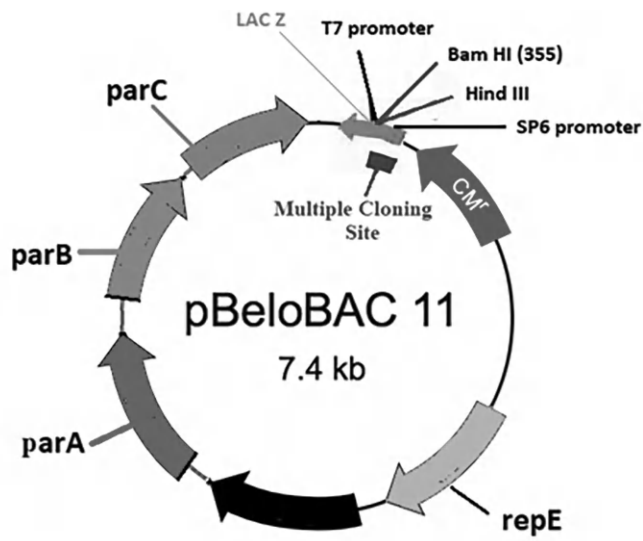


FIGURE 6.9 Bacterial artificial chromosome (pBeloBAC11).

6.4 CONCLUSION

Genetic engineering has undergone significant progress since its inception, and its associated tools have also evolved accordingly. Today, we have a wide array of powerful and precise techniques that can be used to manipulate the genetic material of living organisms. These tools have greatly transformed biology and opened up new opportunities for research, discovery, and innovation.

The future of genetic engineering is promising, given the continuous development of new tools and techniques. As we continue to improve our understanding of genetics and molecular biology, more powerful tools are expected to emerge that will further expand the prospects of genetic engineering.

To summarize, genetic engineering is a powerful tool that has transformed the field of biology, and the tools used for it have been critical in this transformation. With the continued development of novel and more effective tools, we can expect genetic engineering to advance and pave the way for exciting discoveries and innovations.

KEYWORDS

- **genetic engineering**
- **restriction enzymes**
- **vectors**
- **DNA manipulation**
- **molecular biology tools**
- **cloning**

REFERENCES

Bajpai, B. High Capacity Vectors. In *Advances in Biotechnology*; Springer: India, 2014; vol *9788132215547*, pp 1–10. https://doi.org/10.1007/978-81-322-1554-7_1

Burke, D. T.; Carle, G. F.; Olson, M. V. Cloning of Large Segments of Exogenous DNA into Yeast by Means of Artificial Chromosome Vectors. *Science* **1987,** *236* (4803), 806–812. https://doi.org/10.1126/SCIENCE.3033825

Capecchi, M. R. Altering the Genome by Homologous Recombination. *Science* **1989,** *244* (4910), 1288–1292. https://doi.org/10.1126/science.2660260

Cohen, S. N.; Chang, A. C. Y.; Boyer, H. W.; Helling, R. B. Construction of Biologically Functional Bacterial Plasmids In Vitro. *Proc. Natl. Acad. Sci. U. S. A.* **1973,** *70* (11), 3240–3244. https://doi.org/10.1073/PNAS.70.11.3240

Cohen, G.; Sternberg, N. Genetic Analysis of the Lytic Replicon of Bacteriophage P1. I. Isolation and Partial Characterization. *J. Mol. Biol.* **1989,** *207* (1), 99–109. https://doi.org/10.1016/0022-2836(89)90443-9

Complete Nucleotide Sequence of Bacteriophage MS2 RNA: Primary and Secondary Structure of the Replicase Gene - PubMed. (n.d.) [Online]. https://pubmed.ncbi.nlm.nih.gov/1264203/ (accessed Feb 28, 2023)

Fraley, R. T.; Rogers, S. G.; Horsch, R. B. Genetic Transformation in Higher Plants. *CRC Crit. Rev. Plant Sci*. **1986,** *4*, 1–46.

Goodman, R. M.; Hauptli, H.; Crossway, A.; Knauf, V. C. Gene Transfer in Crop Improvement. *Science* **1987,** *236*, 48–54.

Gossen, M.; Bujard, H. In *Tight Control of Gene Expression in Mammalian Cells by Tetracycline-Responsive Promoters*, Proceedings of the National Academy of Sciences of the United States of America, 1992; vol *89* (12). https://doi.org/10.1073/pnas.89.12.5547

Hoy, M. A. Some Additional Tools for the Molecular Biologist. In *Insect Molecular Genetics*, 2013; pp 215–249. https://doi.org/10.1016/B978-0-12-415874-0.00006-8

Messing, J.; Vieira, J. A New Pair of M13 Vectors for Selecting either DNA Strands or Double Digest Restriction Fragments. *Gene* **1982,** *19*, 269–76.

Nicholl, D. S. In *Genetic Engineering*, 3rd Ed.; 2008.

References - Scientific Research Publishing. (n.d.) [Online]. https://www.scirp.org/(S(i43dyn45teexjx455qlt3d2q))/reference/ReferencesPapers.aspx?ReferenceID=1480044 (accessed Feb 28, 2023)

Roberts, R. J.; Kim, S. H. Type I Restriction Enzymes: Properties and Applications. *Microbiol. Mol. Biol. Rev.* **2016,** *80* (2), 264–277.

Roberts, R. J.; Murray, K. Restriction Endonuclease. *CRC Crit. Rev. Biochem.* **1976,** *4* (2), 123–164.

Sambrook, J.; Fritsch, E. R.; Maniatis, T. In *Molecular Cloning A Laboratory Manual*, 2nd Ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY – (1989).

Sambrook, J.; Russell, D. W. In *Molecular Cloning: A Laboratory Manual*, 3rd Ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, N.Y. *Zool. Res.* **2001,** *1* (1).

Shuman, S. DNA Ligases: Progress and Prospects. *J. Biol. Chem.* **2009,** *284* (26), 17365–17369. https://doi.org/10.1074/jbc.R900017200

Smith, H. O. Restriction Enzymes and their Recognition Sequences. *Nature* **1974,** *250* (5468), 675–679. DOI: 10.1038/250675a0. PMID 4139092.

CHAPTER 7

Gene and Cell Division

AARADHYA BHATNAGAR, SACHCHIT PATEL, and ARPIT SHUKLA

Indrashil University, Rajpur, Gujarat, India

ABSTRACT

Genetics in recent times has been richly developed through experimentation and research. Mendel was the first to set outlines for genetics even before chromosomes or genes were known. The distribution of genes is an important aspect of the identification of coding and noncoding protein sequences. The Human Genome Project (HGP) has set a framework in the field of genetics and has been widely used in medicinal biology. There are various methods for sequencing DNA, which are at the heart of the diagnosis and treatment of multiple diseases, which are also discussed here. In this chapter, we have also described cell division, which is the last step of reproduction. Cell division is carried out in synchronized events, including different phases. Individual phases have been discussed here. During meiosis, when crossing-over occurs, both crossing-over and recombination frequencies can be generated. These frequencies are further used to determine gene mapping, which challenges Mendel's Laws of Inheritance. Different techniques are used to study phases of the cell cycle. Isotope labeling and flow cytometry are the most used analytical tools for cellular growth and development. Cellular balance is necessary for continuance. It is controlled by apoptosis and senescence. Genetics plays a vital role in both of these processes. Both of these processes should function accurately, or else the aged or damaged cells can cause harm to the hosts. In sum, there are promising new avenues opening up in the study of genes and the cell cycle. We can anticipate many more advances in our understanding of these fundamental biological processes as technology and research methods continue to evolve.

Molecular Genetics, Structures, Mechanisms, and Functions, Volume 1: Principles of Gene Manipulation and Genomics. Tariq Ahmad Bhat, Abdul Rauf Shakoori, and Jameel M. Al-Khayri (Eds.)
© 2026 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

7.1 UNDERSTANDING GENES AND CHROMOSOMES

The term "gene" is widely used these days. In the early 1860s these inheritance traits were first noticed by Gregor Mendel. He used the term "cell elements" (now known as gene) to describe the phenomenon of inheritance using the pea plant (*Pisum sativum*). The proper definition of a gene is an evolving phenomenon due to advanced discoveries that give different attributes, but for simplicity, according to National Cancer Institute (NCI) Dictionary of Genetics Terms, it is elicited as "the basic unit of heredity passed from parent to child." Each gene has a differing form present which is known as "allele." Diploid organisms like humans are called so due to the presence of two alleles. Each allele is derived from one of two parents. Alleles derived may be identical (known as Homozygous condition) or may differ (known as Heterozygous condition). In the case of unidentical or alternative forms, the allele that is able to express itself is known as the dominant allele, while the allele that is unable to express itself at the locus is known as the recessive allele. A chromosome is a thread-like structure made up of protein and a single molecule of DNA. It is present inside the nucleus of the animal or plant cell. We shall learn the difference between genes and chromosomes in Table 7.1.

Each organism contains a specific set of genes which make it a functional living system. "Genome" is the entirety of an organism's genetic makeup. The particular set of genes in an individual is termed the "genotype." The genotype of an organism determines the physical or morphological characteristics, which are referred to as "phenotype." The human genome sequencing (Lander et al., 2001) derived a great understanding of the correlation between genotype and phenotype, which concerns human health (Gonzaga-Jauregui et al., 2012).

TABLE 7.1 Difference between Gene and Chromosome. ⏎

Gene	Chromosome
• Gene is found on chromosome.	• Chromosomes are structures built with protein and DNA.
• Microscopically, genes are not visible.	• Chromosomes can be observed under a microscope.
• Gene mutations are categorized into point mutations and frameshift mutations, that is, insertion and deletion.	• Chromosome mutations lead to chromosomal disorders.

We shall talk about mutations and their types in detail in upcoming topics.

7.2 HUMAN GENOME ARRANGEMENT

Understanding the arrangement of the human genome is a critical area of study in genetics and genomics, with implications for personalized medicine and advancements in treating genetic disorders. The genome is organized into regions with different functions, including genes that code for proteins, noncoding regions that regulate gene expression, and repetitive sequences. The human genome comprises DNA molecules, which are organized into 23 pairs of chromosomes. Each chromosome contains long sequences of genetic information, which are composed of four different nucleotide bases (adenine, guanine, cytosine, and thymine). These sequences of nucleotides make up genes, which code for the proteins that carry out most of the functions in our cells. We already know that mitochondria have autonomous machinery, that is, they have their own genetic material. Therefore, the genetic information is enciphered into two genomes: Nuclear and Mitochondrial (Gonzaga-Jauregui et al., 2012). The human genome consists of genes that are not similar when it comes to size, the smallest being the testis determining factor, which is present on the Y chromosome and is essential for male character development, while the largest being dystrophin, present on the X chromosome. The gene size is a very important marker for the evolutionary aspect of human life; it is believed that genes with a longer tend to have a high intronic load and are more evolutionarily conserved (Wolf et al., 2009). On the flip side, high expression, smaller proteins, and low intronic content are all correlated with shorter gene lengths (Urrutia and Hurst, 2003). However, with regard to recent advancements, we are yet to fully understand the intricacy and operation of the human genome (Simonti and Capra, 2015).

7.2.1 HUMAN GENE NUMBER

The number of genes generated by our genome is still very much known. There were many proposed methods to estimate this number roughly. However, these methods will not give us the accurate number of genes present.

TABLE 7.2 Estimates of the Number of Genes in Human Genome based on Various Methods.

Method	Estimated number of genes
Experimental determination of protein-coding genes	19,000–20,000
Comparative genomics and gene prediction algorithms	19,000–25,000
Transcriptome sequencing and analysis	60,000–100,000
Analysis of alternative splicing and isoform diversity	90,000–200,000
Analysis of noncoding RNA transcripts	50,000–100,000

7.2.2 GENE DISTRIBUTION

The distribution of genes has been the subject of intense research for determining the location of protein-coding and noncoding sequences. There are some findings about the gene distribution in the human genome.

- Protein-coding genes: The most well-studied class of genes in the human genome are protein-coding genes, which are responsible for producing the many different proteins that carry out cellular functions. Estimates of the total number of protein-coding genes in the human genome have varied over time, but current estimates place the number at around 19,000–20,000 genes (International Human Genome Sequencing Consortium). These genes are unevenly distributed across the chromosomes, with some chromosomes (such as chromosome 19) having a higher gene density than others.["The human genome: organization and evolutionary history." *Annual Review of Genetics* vol. 29]
- Noncoding RNA genes: In addition to protein-coding genes, the human genome also contains many different types of noncoding RNA genes, which produce RNA molecules that play regulatory roles in gene expression and other cellular processes. These noncoding RNA genes are more difficult to identify and characterize than protein-coding genes, but estimates suggest that there may be tens of thousands of different noncoding RNA genes in the human genome (Carninci et al., 2005).
- Repeats and noncoding regions: The majority of the human genome is composed of noncoding DNA, including repetitive sequences and other regions that do not code for genes. Gene expression and other cellular processes are regulated by these noncoding regions, but their function remains unclear. Repetitive sequences, in particular, are thought to be involved in chromosomal structure and stability, and mutations in repetitive regions have been linked to several genetic disorders (Smit, 1996).

7.2.3 EXON CHARACTERISTIC

Exons are the type of DNA sequence present in m-RNA responsible for coding amino acid sequences. These sequences are intervened with noncoding sequences which are termed as Introns. These intronic sequences are removed by Splicing before translation occurs. Exons in the human genome vary in length from just a few base pairs to thousands, with a median length of around 145 base pairs (Sakharkar et al., 2004).

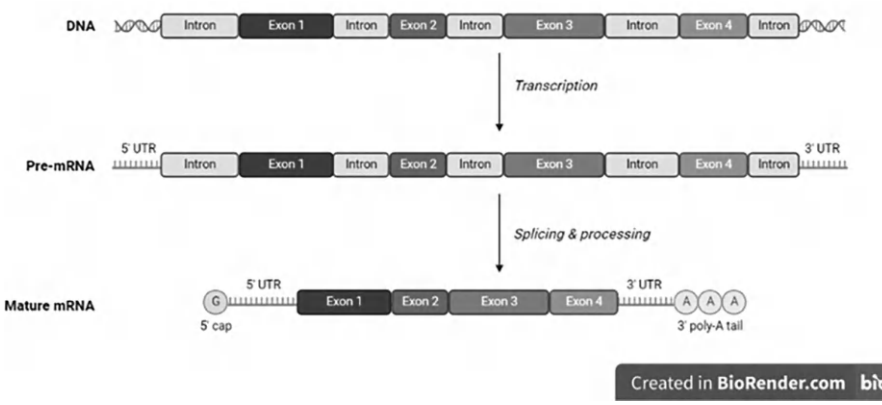


FIGURE 7.1 Eukaryotic gene expression: From DNA to mature mRNA.

7.3 HUMAN GENOME PROJECT

There is a dedicated project undertaken for sequencing of the whole human genome, which is known as The Human Genome Project (HGP), which mainly had three goals: (1) sequencing of the genome as a whole, (2) identification and lastly, and (3) gathering optimized data analysis records for the same. The project started in 1990 and was completed in 2003.

7.3.1 METHODOLOGY

The HGP was a massive undertaking that involved hundreds of researchers from around the world. The project was divided into three main phases: mapping, sequencing, and analysis. The mapping phase involved identifying the locations of genes on the chromosomes, while the sequencing phase involved determining the order of the DNA's nucleotides (A, T, C, and G). The analysis phase involved interpreting the data and understanding the functions of the genes (Lander et al., 2001).

7.3.2 RESULTS

The HGP was a resounding success. The final sequence of the human genome was published in 2003, and it revealed a wealth of information about human biology and genetics. The HGP identified approximately 20,000–25,000 protein-coding genes in the human genome, as well as thousands of noncoding genes that play important regulatory roles (McPherson et al., 2001). The project also identified many disease-causing genes and provided new insights into the genetic basis of many complex diseases, including cancer, diabetes, and heart disease.

7.3.3 IMPACT

The HGP has had a profound impact on medicine and biology. It has led to the establishment of new tests for analysis and treatments for genetic diseases, as well as new approaches to personalized medicine. The HGP has also revolutionized the field of genetics, providing researchers with the latest tools and techniques for studying the genome and understanding its functions. The HGP has spawned many follow-up projects, including the 1000 Genomes Project, which lines up to sequence the genomes of thousands of individuals from different populations, and the ENCODE project, which seeks to find all of the human genome's functional elements.

7.3.4 CONCLUSION

The Human Genome Project was a landmark achievement in genetics and has had a profound impact on medicine and biology (Green and Guyer, 2011). The project has given researchers much information about the human genome and has opened up new possibilities and advances for the future.

7.4 METHODS FOR SEQUENCING OF DNA

DNA sequence is a method used to determine the exact order of nucleotides within a DNA molecule. There are some prevalent methods for DNA sequencing:

- **Sanger sequencing:** The Sanger method is a chain-termination method that uses dideoxy nucleotides to stop DNA polymerase from adding more nucleotides to the growing chain. Using a mixture of standard and dideoxynucleosides leads to the production of nucleotide

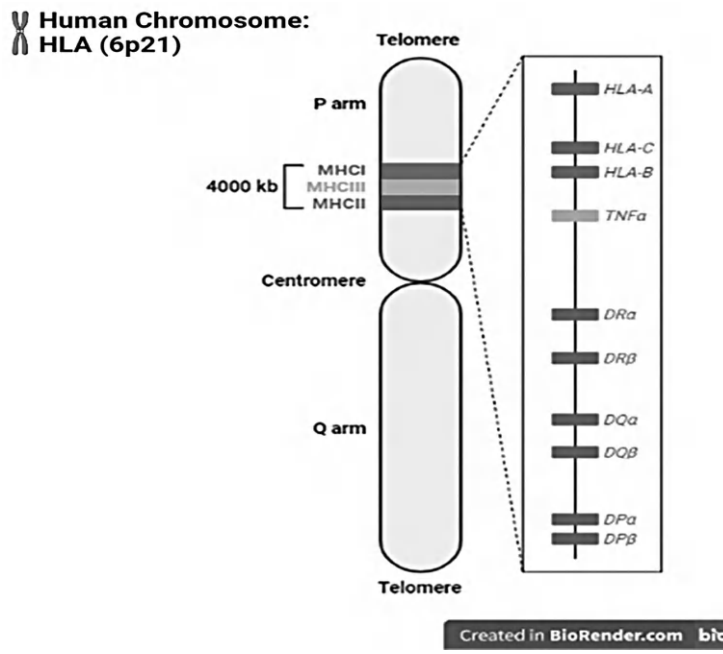


FIGURE 7.2 Human Chromosome: HLA (6p21).

sequences of different lengths. These fragments are then separated by size using gel electrophoresis and the sequence can be read off from the pattern of bands. Sanger sequencing is still used for sequencing short stretches of DNA and for confirming NGS results (Sanger et al., 1977).

- **Next-generation sequencing (NGS):** NGS technologies allow for massively parallel sequencing of many DNA fragments at once, which has greatly increased the speed and reduced the cost of DNA sequencing (Shendure and Ji, 2008). There are several different NGS platforms available, each with its own strengths and weaknesses. Some of the most widely used platforms include:

  i. Illumina: Illumina sequencing is based on the reversible terminator method, in which nucleotides are added one at a time to a growing DNA strand, and each nucleotide is labeled with a fluorescent dye. The fluorescent signal is detected and recorded by a camera, and the sequence is reconstructed by aligning the signals from each nucleotide. Illumina sequencing is highly accurate and can generate large amounts of data quickly and inexpensively (Goodwin et al., 2016).

ii. Ion Torrent: Ion Torrent sequencing is based on the observation of hydrogen ions released upon adding a nucleotide to the growing DNA strand. The release of ions is detected by a semiconductor chip, and the sequence is determined by analyzing the pattern of ion signals. Ion Torrent sequencing is fast and simple but has lower accuracy than Illumina sequencing.
iii. PacBio: PacBio sequencing uses a single-molecule real-time (SMRT) sequencing method that allows for long reads of up to tens of kilobases. The technology is based on detecting light pulses generated by incorporating nucleotides into a growing DNA strand. PacBio sequencing has high accuracy and is especially useful for sequencing complex genome regions, but it is more expensive than other NGS platforms (Rhoads A, Au KF).

7.5 MUTATION

Mutation is an alteration in the DNA sequence of any organism. Mutation can result from DNA replication errors, exposure to mutagens or carcinogens, etc. The mutation occurs in the somatic cell; it is not passed on to the next generation, while the mutation occurring in the germ cell can be passed on to the offspring.

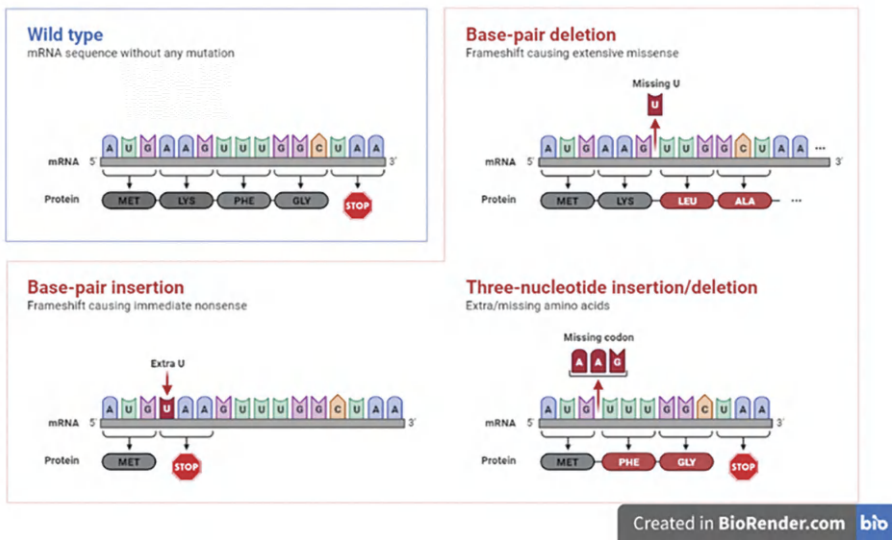


FIGURE 7.3 Effects of mutations on mRNA and protein translation.

Here are some of the most common types of mutations:

1. Point mutations: Point mutations are changes to a single nucleotide in the DNA sequence. They can be classified into three types:
    - Silent mutations: These mutations do not alter the sequence of amino acids in the protein encoded by the gene, as the mutated codon still corresponds to the identical amino acid as the initial codon.
    - Missense mutations: Due to these mutations, a single amino acid in the protein sequence is altered since the mutated codon now corresponds to a different amino acid than the original codon.
    - Nonsense mutations: These mutations introduce a premature stop codon into the protein sequence, which can result in a truncated and nonfunctional protein (Mitelman et al., 2017).

2. Insertions and deletions: These mutations involve adding or removing of one or more nucleotides from the DNA sequence. They can disrupt the reading frame of the gene, leading to a frameshift mutation that alters the amino acid sequence of the protein (Pollard et al., 2016).
3. Chromosomal rearrangements: These mutations involve changes to the structure or number of chromosomes. Examples include:
    - Inversions: When a segment of a chromosome is reversed in orientation.
    - Translocations: When a segment of one chromosome is moved to another chromosome.
    - Deletions and duplications: When a segment of a chromosome is lost or gained (Alberts et al., 2015).

7.5.1 EPIGENETIC CHANGES

Epigenetics is a study that explores how the immediate environment and personal activities can impact the functioning of genes. Epigenetic inheritance plays a major role in the adaptation of an organism in today's ever-changing environment without modifications in the DNA sequence(Peschansky, V. J., &; Wahlestedt, C. Non-coding RNAs as direct Direct and indirect Indirect modulators Modulators of epigenetic Epigenetic regulationRegulation. *Epigenetics* (2014)., *9* (1), 3–12).

Here are some examples of epigenetic changes and relevant references:

1. **DNA methylation:** This refers to a chemical alteration of DNA where a methyl group is added to the cytosine base of a CpG dinucleotide. DNA methylation can lead to gene silencing and has been implicated in a variety of biological processes, including embryonic development, X chromosome inactivation, and genomic imprinting. (Reference: Jones, P. A.; Liang, G. Rethinking how DNA Methylation Patterns are Maintained. *Nat. Rev. Genet.* 2009, *10* (11), 805–11.)
2. **Histone modification:** Histones are the proteins that package DNA into a compact structure called chromatin. Histone modification has the potential to influence chromatin structure and regulate gene expression. For instance, histone acetylation is typically linked to gene activation, whereas histone methylation can either activate or repress gene expression depending on the specific residue being modified on the histone. (Reference: Jenuwein, T.; Allis, C. D. Translating the Histone Code. *Science* 2001, *293* (5532), 1074–1080.)
3. **Noncoding RNA:** Noncoding RNA molecules, such as microRNAs and long noncoding RNAs, can regulate gene expression by binding to messenger RNA and inhibiting translation or promoting degradation. Noncoding RNAs have been implicated in a wide range of biological processes, including embryonic development, cancer, and neurodegenerative diseases. (Reference: Esteller, M. Non-Coding RNAs in Human Disease. *Nat. Rev. Genet.* 2011, *12* (12), 861–74.)
4. **Chromatin remodeling:** Chromatin remodeling refers to the process of altering the structure of chromatin to make it more or less accessible to transcriptional machinery. This can involve the action of large protein complexes that slide or remove nucleosomes or the action of ATP-dependent chromatin remodeling enzymes that use energy to move nucleosomes. Chromatin remodeling is important for regulating gene expression and has been implicated in many biological processes, including differentiation and development. (Reference: Ho, L.; Crabtree, G. R. Chromatin Remodelling during Development. *Nature* 2010, *463* (7280), 474–484.)
5. **Environmental factors:** External factors, including diet, stress, and exposure to toxins, can induce epigenetic alterations as well. For instance, the maternal diet during pregnancy can impact the DNA methylation patterns of the developing fetus, and exposure to toxins

like cigarette smoke can result in changes in DNA methylation and histone modification in adult cells. (Reference: Skinner, M. K. Environmental Epigenetics and a Unified Theory of the Molecular Aspects of Evolution: A Neo-Lamarckian Concept that Facilitates Neo-Darwinian Evolution. *Genome Biol. Evol.* 2015, *7* (5), 1296–302.)

7.6 CELL DIVISION

A cell goes through a specific sequence of events called the cell cycle before it can divide and produce copies of itself. The cell cycle is a tightly controlled process that guarantees proper cell division while keeping the cell's DNA intact.

There are the cell cycle's two primary stages, and they are known as interphase and mitosis. When a cell is not dividing, it is in a phase called interphase, during which it can expand and duplicate its DNA. The process by which DNA is replicated and then split between two new cells is called mitosis.

There are three distinct stages, G1, S, and G2, during the interphase. A cell's growth and regular metabolic activities are carried out during the G1 phase. It is during the S phase that DNA replication occurs in the cell. The cell divides into two during the G2 phase when it synthesizes the proteins and other components it will need for the process.

A mitosis progresses through four phases, namely, prophase, metaphase, anaphase, and telophase. Condensation of chromatin into visible chromosomes and disintegration of the nuclear sheath are hallmarks of prophase. During metaphase, chromosomes align along the cell's midline. When a cell enters anaphase, its sister chromatids become detached and are dragged to opposing ends. Telophase is the cell cycle's last stage, and it is when the nuclear envelope is reformed around the two sets of chromosomes and cytokinesis occurs.

Several checkpoints exist throughout the cell cycle to guarantee that the cell is properly prepared for the following phase. Cancer can originate from cells dividing incorrectly, which, if the cell cycle is not correctly controlled, may occur.

7.6.1 PHASES OF CELL CYCLE

After the conclusion of the M phase, which involves the cytoplasm's division and mitosis, the partitioning of the chromosomes and other cellular components, each cell is born. During the S phase, the chromosomal DNA is reproduced (synthetic phase). Between mitosis and the S phase, there are two further phases. The time between mitosis and the beginning of

DNA replication is known as the G1 phase (first gap phase). The interval from when DNA replication ceases to when mitosis begins is known as the G2 phase (second gap phase). There is a M phase and a S phase in every cycling cell. In certain early embryos, the G1 and G2 stages are relatively brief and vary in duration (Lodish et al., 2000).

7.6.1.1 INTERPHASE

The stage of the cell cycle known as interphase is when the cell gets ready to divide. The cell cycle's most extended phase is this one where it can be divided into three stages: G1, S, and G2. During interphase, the cell develops, copies its DNA, and gets ready for mitosis or meiosis (Becker et al., 2017).

G1 phase: During the G1 phase, the cell develops and produces new organelles and proteins. This phase also involves a series of regulatory events that ensure the cell is ready to go on to the next phase of the cell cycle. G1 is the most variable phase of interphase and can persist for many days or a few hours, depending on the cell type (Cooper and Hausman, 2019).

S phase: During the S phase, DNA replication occurs. This means that the cell's genetic material is duplicated to ensure that every daughter cell has a full complement of chromosomes. To guarantee that each chromosome is copied just once, this process is strictly controlled (National Center for Biotechnology Information, 2021).

G2 phase: During G2 phase, the cell undergoes a final period of growth and prepares for cell division. The cell checks its DNA for errors and ensures that everything is ready for mitosis or meiosis. The cell also synthesizes proteins that will be needed for cell division (Alberts et al., 2015).

Interphase is a critical part of the cell cycle, as it ensures that the cell is properly prepared for cell division. It is also a time when cells can respond to signals from their environment and adjust their behavior accordingly. For example, cells in the G1 phase can enter a state of quiescence (G0 phase) if they receive signals indicating that they should not divide (Compton, 2017).

7.6.1.2 M PHASE OR MITOSIS

M phase, also known as mitosis, is a kind of cell division that yields two daughter cells with the exact same genetic makeup. It is crucial for the

development of new tissues, tissue growth, and tissue healing of organisms during development. During mitosis, the division of a parent cell into two identical daughter cells through a series of stages that include prophase, metaphase, anaphase, and telophase. Specific cellular characteristics define each level, such as the condensation of chromosomes, the chromosomes' placement at the metaphase plate, and during anaphase, the division of sister chromatids (Hanahan and Weinberg, 2011).

The process of mitosis is tightly regulated by a variety of signaling pathways, including the checkpoint kinase pathway, which ensures that cells do not enter anaphase until all chromosomes have the mitotic spindle correctly connected (BioRender.com).

In addition to its role in normal cell division, mitosis can also play a critical role in cancer development. Abnormalities in the mitotic machinery can result in chromosomal instability and the accumulation of genetic mutations, which can ultimately lead to the development of tumors (Alberts et al., 2002).

These are the phases of mitosis:

1. Interphase: During this phase, the cell is actively performing its functions, such as growth and DNA replication, without dividing. The DNA is present in the cell as the loose, uncondensed structure known as chromatin.
2. Prophase: The first mitotic stage during which chromatin condenses into chromosomes. Under a microscope, the chromosomes can be seen as distinct, compact structures. The nuclear envelope disintegrates, and the nucleolus, which produces ribosomes, vanishes. The centrosomes move in opposite directions toward the cell's poles, where they organize the microtubules that separate the chromosomes.
3. Prometaphase: During this phase, the spindle fiber-forming microtubules affix to the kinetochores, which are chromosomal structures that aid during cell division in chromosomal alignment. The metaphase plate, the region where the chromosomes are located in the cell, starts to move toward the center.
4. Metaphase: The chromosomes, along the metaphase plate, are in alignment, which is a fictitious line in the cell's center, at this time. In order to align the chromosomes in readiness for separation, the spindle fibers attach to them at the kinetochores.
5. Anaphase: During this phase, the sister chromatids, which are identical copies of the chromosomes, are pulled apart and in opposite directions by the spindle fibers. The spindle fibers that are not

connected to the chromosomes slide past one another, further separating the poles, and the cell enlarges.

6. Telophase: At this time, the chromosomes have moved to the cell's opposite poles. A new nuclear envelope develops around each set of chromosomes as the spindle fibers disintegrate, causing the separation of two distinct nuclei. The chromosomes start to decondense and transform back into chromatin.
7. Cytokinesis: The process of dividing a cell into two daughter cells, during which the cytoplasm and organelles are divided between them. Animal cells split in half as a result of a cleavage furrow that develops and deepens over time. A plant cell's centers evolve into cell plates and eventually spread outward to divide the cell in half.

Overall, mitosis is a strictly controlled process that makes sure that genetic information is faithfully transmitted from one generation of cells to the next.

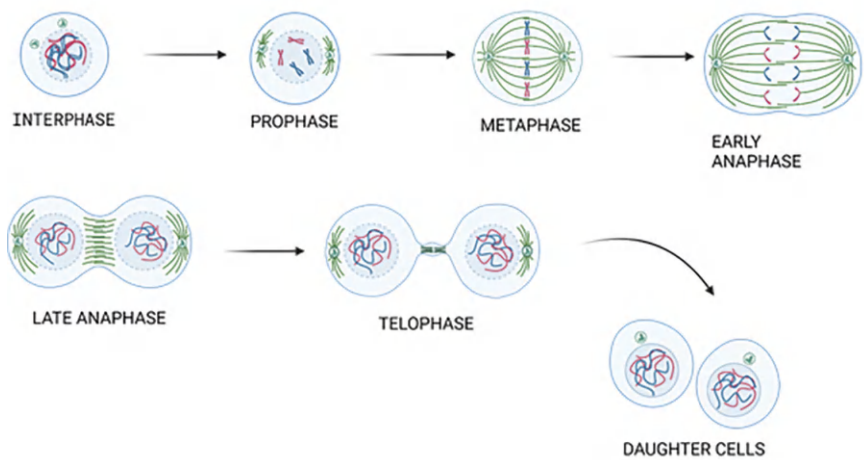


FIGURE 7.4 Stages of mitosis in a eukaryotic cell.

7.6.1.3 MEIOSIS

Meiosis is a particular kind of cell division that takes place in sexually reproducing organisms. Unlike mitosis, which results in two genetically identical daughter cells, meiosis results in four genetically diverse daughter cells. Meiosis I and Meiosis II, the two divisions that separate sister chromatids and homologous chromosomes, respectively, are used to achieve this. One pair of chromosomes from each parent begins meiosis when a diploid cell with two nuclei is divided into two. These homologous chromosomes link together and

cross across during meiosis I, exchanging DNA segments in the process. This creates new combinations of genetic information that are different from either parent. Next, the homologous chromosomes split apart and migrate to the cell's opposing ends, creating two haploid daughter cells. Each daughter cell contains one set of chromosomes, with two chromatids each (Alberts et al., 2015).

During meiosis II, the sister chromatids split apart and relocated to the opposite ends of the cell, creating four haploid daughter cells. These daughter cells are genetically diverse due to the random alignment of chromosomes during meiosis I and the separation of sister chromatids during meiosis II. Meiosis plays a crucial role in sexual reproduction and is advantageous because of the creation of genetically diverse offspring. It is also important for maintaining the correct number of chromosomes in a species, as the diploid number is restored when two haploid gametes fuse during fertilization (Campbell et al., 2017).

7.6.1.3.1 The Process of Meiosis

Cell division, also known as meiosis, takes place in sexually reproducing organisms. Two rounds of cell division are required for the process to produce four haploid cells, each of which has half as many chromosomes as the parent cell. Meiosis is a crucial process in sexual reproduction as it ensures genetic diversity among offspring. Here is a detailed overview of the process of meiosis:

Prophase I: Prophase I is the initial phase of meiosis I. In this phase, chromosomes that are homologous couple up and cross across to exchange genetic material. This exchange of genetic material helps to increase genetic diversity among offspring. Prophase I is the longest phase of meiosis I and might need many days to finish (Alberts et al., 2015; BioRender.com).

Metaphase I: During metaphase I, the homologous pairs of chromosomes align along the equator of the cell. This alignment is random, which further contributes to genetic diversity (Campbell et al., 2017; https://www.differencebetween.com/difference-between-crossover-frequency-and-recombinationfrequency).

Anaphase I: During anaphase I, the homologous pairs of chromosomes separate and are drawn to the cell's opposing ends (Alberts et al., 2015).

Telophase I: Telophase I is the last phase of meiosis I. The nuclear membrane reforms around each set of chromosomes at this point, and the chromosomes move to the cell's opposing poles. After that, the cell splits

into two daughter cells, each of which has half as many chromosomes as the parent cell (Campbell et al., 2017).

Prophase II: Meiosis II begins with prophase II, during which the chromosomes condense again, and the nuclear membrane breaks down (Alberts et al., 2015).

Metaphase II: During metaphase II, the chromosomes align along the equator of the cell, similar to metaphase I. However, unlike metaphase I, the chromosomes are not in homologous pairs (Campbell et al., 2017).

Anaphase II: During anaphase II, the sister chromatids of each chromosome separate and are pulled toward opposite ends of the cell (Alberts et al., 2015).

Telophase II: The final stage of meiosis II is telophase II. During this stage, the chromosomes reach opposite poles of the cell and the nuclear membrane reforms around each set of chromosomes. The cell then divides, resulting in four haploid daughter cells, each with a unique combination of genetic material (Campbell et al., 2017).

Overall, the process of meiosis is a complex and important process that ensures genetic diversity among offspring. By following the steps outlined above, organisms are able to produce gametes that are genetically distinct from one another, increasing the chances of survival and adaptation.

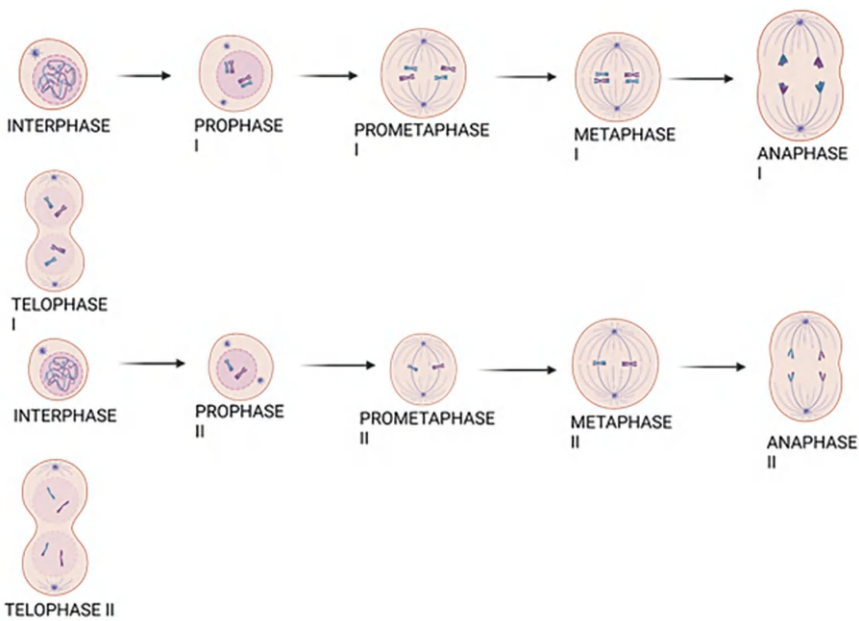


FIGURE 7.5 Stages of meiosis I and II in eukaryotic cell division

7.6.2 CROSSOVER FREQUENCY AND RECOMBINATION FREQUENCY

7.6.2.1 CROSSOVER FREQUENCY

During meiosis, genetic crossing-over occurs. This causes genetic characters to change during gamete formation. As a result, genetic variation among organisms occurs. The formation of the double-stranded breaks marks the beginning of the crossing-over. DNA topoisomerases are responsible for this process. The formation of double-stranded breaks results in the formation of crossing-over genetic material. The frequency of crossing-over is determined by how quickly the crossing-over occurs. The frequency of crossing-over is determined by the distance between the genes. A gene's genetic map is determined by the frequency of crossing-over. The frequency of crossing-over can be divided into homozygous crossing-over and heterozygous crossing-over (International Human Genome Sequencing Consortium, 2001).

7.6.2.2 RECOMBINATION FREQUENCY

Recombination is the process by which genes cross over in a heterozygous manner. The recombination frequency is the frequency or rate at which recombination occurs during crossing-over. The increase in recombination frequency suggests a pattern of offspring variation. The rate of recombination or heterozygous crossing-over is determined by the distance between the genes. As a result, the frequency of recombination increases as the distance between the genes decreases. This increases the genetic linkage between the genes (International Human Genome Sequencing Consortium, 2001).

What resemblances and contrasts exist between crossing-over and recombination frequencies?

- Both occur during meiosis.
- Crossing-over occurs when generating crossing-over and recombination frequencies.
- The frequency of crossing-over and recombination is determined by the distance between the two genes.
- Both types of frequencies are used to determine genetic linkage and gene mapping (International Human Genome Sequencing Consortium, 2001).

7.6.3 GENE MAPPING

Gene mapping is the process of determining the location and relative distances between genes on a chromosome. There are several methods for gene mapping, including linkage analysis, physical mapping, and genome-wide association studies (GWASs). Here are some references on gene mapping:

1. Linkage analysis:
   Genetic study, known as linkage, identifies the chromosomal location of genes that are responsible for a particular trait. This method involves analyzing the co-segregation of genetic markers with the trait of interest within a family. Genes may be mapped via linkage analysis is responsible for inherited disorders (Visscher et al., 2012).
2. Physical mapping:
   Physical mapping is a method that identifies the physical location of genes on a chromosome by analyzing the DNA sequence. This method involves creating the genome's physical map, which can be used to identify the location of specific genes. Physical mapping is commonly used to study complex genetic disorders (Morgan Unit(s), 2006).
3. Genome-wide association studies (GWASs):
   GWASs are a method that identifies genetic variants associated with a particular trait or disease. GWASs involve comparing the DNA of individuals with and without the trait of interest to identify genetic variations that are more common in individuals with the trait. GWASs have been used to identify genes associated with many common disorders like heart disease and diabetes (Morgan, 2007).

7.6.4 MORGAN'S UNIT

Morgan unit(s) is a unit of genetic recombination measured by crossing-over (on average, one crossing-over per meiosis). It specifies the separation between two genes. One centimorgan (cM) equals 1% recombination (Shapiro, 2003).

7.6.5 TECHNIQUES USED TO STUDY CELL CYCLE

The intricate process of the cell cycle is necessary for the growth and development of all living organisms. There are several techniques used to study

the cell cycle, including microscopy, flow cytometry, biochemical assays, and genetic analysis. Here are some of the most commonly used techniques:

1. Microscopy: Microscopy is a powerful tool for studying the cell cycle. It allows researchers to visualize the several phases of the cell cycle and the structures involved in cell division. Techniques such as fluorescence microscopy and confocal microscopy can be used to label and track specific molecules and structures within cells during the cell cycle (Alberts et al., 2002).
2. Flow cytometry: Flow cytometry is a technique that allows researchers to analyze large numbers of cells quickly and accurately. This technique involves passing cells through a flow cytometer, which measures various characteristics of the cells, such as their size, shape, and DNA content. Flow cytometry can be used to identify and analyze cells in different stages of the cell cycle (Sherr and Roberts, 1999).
3. Biochemical assays: Biochemical assays measure the activity of specific proteins and enzymes that are a part of the cell cycle. For example, CDKs assays can be used to measure the activity of CDKs, which are crucial in controlling the cell cycle (Galluzzi et al., 2018).
4. Genetic analysis: Genetic analysis can be used to study the cell cycle by manipulating the genes involved in cell division. For example, researchers can use gene knockout or RNA interference techniques to silence the expression of specific genes involved in the cell cycle and study the effects on cell division (Kalkavan and Green, 2018).

7.6.6 APOPTOSIS AND SENESCENCE

The process of apoptosis, commonly referred to as "programmed cell death," is tightly controlled by which cells undergo a series of biochemical and morphological changes that ultimately lead to their self-destruction. It plays a critical role in various biological processes, such as embryonic.

development, tissue homeostasis, and immune system function. Dysregulation of apoptosis has been implicated in a range of human diseases, including cancer, neurodegenerative disorders, and autoimmune diseases (Elmore, 2007).

The induction of apoptosis is characterized by the activation of a family of cysteine proteases called caspases, which cleave various

cellular proteins and lead to the characteristic morphological changes observed in apoptotic cells, such as cell shrinkage, chromatin condensation, and fragmentation of the nucleus into apoptotic bodies. The apoptotic process is tightly regulated by a complex network of signaling pathways that can be initiated by either extrinsic or intrinsic signals, depending on the nature of the stimulus (Evan and Vousden, 2001).

7.6.6.1 IMPORTANCE OF APOPTOSIS

The importance of apoptosis has been extensively studied in different branches of biology and medicine. For example, in embryonic development, apoptosis is crucial for sculpting and refining the shape and size of developing organs and tissues. In the immune system, apoptosis is critical for the elimination of self-reactive lymphocytes and the maintenance of immune tolerance (van Doeselaar and Burgering, 2018). In cancer, dysregulation of apoptosis can lead to uncontrolled cell proliferation and resistance to chemotherapy or radiation therapy. Apoptosis is hypothesized to contribute to the death of neurons in brain areas affected by neurodegenerative illnesses, including Alzheimer's disease and Parkinson's disease.

7.6.6.2 SENESCENCE

Senescence is a natural response of cells to prolonged stress, which involves a permanent halt to the cell cycle, preventing the accumulation of damaged cells. Different factors can trigger senescence, including telomere attrition, DNA damage, and oncogene activation. This process activates certain proteins, such as p53/p21 or p16/RB, which ultimately result in cell cycle arrest. Senescent cells are usually characterized by sustained levels of the cell cycle inhibitors p21 and/or p16.

A multitude of internal and environmental causes can cause senescence, including DNA damage, oxidative stress, telomere shortening, inflammation, and exposure to certain chemicals or toxins. Once senescence is triggered, it can spread to neighboring cells, which is known as the senescence-associated secretory phenotype (SASP).

The SASP is distinguished by the secretion of various pro-inflammatory cytokines, chemokines, and growth factors, which can have both favorable and unfavorable outcomes for surrounding tissues and organs. On the one hand, the SASP can help to eradicate damaged cells and encourage tissue

restoration, while on the other hand, excessive SASP activity can lead to chronic inflammation, tissue dysfunction, and an increased risk of age-related diseases.

Researchers are exploring various strategies to target senescence and SASP activity to promote healthier aging and extend lifespan. One approach is to create medication that removes senescent cells with precision, which is known as senolytics. Another approach is to target specific signaling pathways or molecules that are involved in senescence and SASP activity, such as the mTOR pathway, NAD+ levels, or the sirtuin proteins.

Overall, senescence is a complex biological process that is crucial for the aging process and disorders associated with aging. By better understanding the mechanisms of senescence, researchers hope to develop new interventions that can slow down or even reverse some of the negative effects of aging.

7.6.7 MAXIMUM MUTATION PROBABILITY

The probability of mutation is generally highest during DNA replication, which happens during the cell cycle's S phase (synthesis phase). During DNA replication, errors in copying the genetic material can occur, leading to mutations. These mutations can be introduced as a result of errors made by the DNA polymerase enzymes responsible for replicating the DNA or from exposure to external mutagens such as chemicals or radiation.

It is important to note that not all mutations result in harmful effects. Some mutations can be neutral or even beneficial, but others can lead to diseases or other adverse effects. The frequency of mutations can be influenced by various factors, including the age of the organism, exposure to environmental factors, and the presence of certain genetic conditions.

7.8 FUTURE PERSPECTIVE

The study of genes and the cell cycle is an important and constantly evolving field of biology, and there are many exciting developments on the horizon. Here are a few potential areas of progress in the near future:

- Gene editing: The development of CRISPR-Cas9 technology has opened up exciting possibilities for targeted gene editing. This technique allows scientists to selectively modify or remove

specific genes from an organism's DNA, which has a wide range of potential applicability to industries like medicine, agriculture, and biotechnology.

- Personalized medicine: With advances in genetics, it may soon be possible to tailor medical treatments to an individual's genetic profile. This could include developing drugs that are specifically targeted to a patient's genetic makeup or using genetic testing to diagnose and treat diseases at an early stage.
- Cell cycle regulation: The cell cycle is a complex process that controls the growth and division of cells, and understanding its regulation is key to developing treatments for diseases such as cancer. Researchers are working to identify the key molecular players involved in the cell cycle and to develop new drugs that can manipulate these processes.
- Stem cell research: Due to their capacity to develop into a variety of cell types, stem cells hold great promise as a tool for regenerative medicine. Researchers are working to understand the molecular processes involved in stem cell differentiation and to develop new techniques for growing and manipulating stem cells.

Overall, the study of genes and the cell cycle is a rapidly advancing field with many exciting possibilities for the future. As technology and research methods continue to evolve, we can expect to see many new breakthroughs in our understanding of these fundamental biological processes.

KEYWORDS

- **gene and cell division**
- **chromosomes**
- **human genome project**
- **Mendel's laws**
- **DNA sequencing**
- **gene mutations**
- **protein-coding genes**

REFERENCES

Alberts, B.; Johnson, A.; Lewis, J.; et al. Meiosis. In *Molecular Biology of the Cell*, 4th Ed.; Garland Science: New York, 2002. https://www.ncbi.nlm.nih.gov/books/NBK26823/

Alberts, B.; Johnson, A.; Lewis, J.; et al. In *Molecular Biology of the Cell*, 4th Ed.; Garland Science, 2002.

Alberts, B.; Johnson, A.; Lewis, J.; et al. In *Molecular Biology of the Cell*, 6th Ed.; Garland Science: New York, 2015.

Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. Chapter 17, Interphase and the Regulation of Cell Division. In *Molecular Biology of the Cell*, 6th Ed.; Garland Science, 2015.

Anzai, T.; Shiina, T. The Molecular Evolution of Major Histocompatibility Complex Class I Genes in Primates. *Crit. Rev. Immunol.* **2015,** *35* (2), 119–134.

Becker, W. M.; Kleinsmith, L. J.; Hardin, J.; Bertoni, G. P. Chapter 11, The Cell Cycle and Cell Division. In *The World of the Cell*, 9th Ed.; Pearson 2017.

Campbell, N. A.; Reece, J. B.; Urry, L. A.; et al. In *Biology*, 11th Ed.; Pearson: Boston, 2017.

Carninci, P.; Kasukawa, T.; Katayama, S.; et al. The Transcriptional Landscape of the Mammalian Genome. *Science* **2005,** *309* (5740), 1559–1563.

Compton, D. A. Mechanisms of Chromosome Alignment and Segregation during Mitosis. *Ann. Rev. Biochem.* **2017,** *86*, 85–108.

Cooper, G. M.; Hausman, R. E. Chapter 16, The Cell Cycle. In *The Cell: A Molecular Approach*, 8th Ed.; Sinauer Associates, 2019.

Elmore, S. Apoptosis: A Review of Programmed Cell Death. *Toxicol. Pathol.* **2007,** *35* (4), 495–516. DOI: 10.1080/01926230701320337

Evan, G. I.; Vousden, K. H. Proliferation, Cell Cycle and Apoptosis in Cancer. *Nature* **2001,** *411* (6835), 342–348. DOI: 10.1038/35077213

Galluzzi, L.; Vitale, I.; Aaronson, S. A.; et al. Molecular Mechanisms of Cell Death: Recommendations of the Nomenclature Committee on Cell Death 2018. *Cell Death Differ.* **2018,** *25* (3), 486–541. DOI: 10.1038/s41418-017-0012-4.

Gonzaga-Jauregui, C.; Lupski, J. R.; Gibbs, R. A. Human Genome Sequencing in Health and Disease. *Annu. Rev. Med.* **2012,** *63*, 35–61. 10.1146/annurev-med-051010-162644

Goodwin, S.; McPherson, J. D.; McCombie, W. R. Coming of Age: Ten Years of Next-Generation Sequencing Technologies. *Nat. Rev. Genet.* **2016,** *17* (6), 333–351.

Green, E. D.; Guyer, M. S. Charting a Course for Genomic Medicine from Base Pairs to Bedside. *Nature* **2011,** *470* (7333), 204–213.

Hanahan, D.; Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **2011,** *144* (5), 646–674.

https://www.differencebetween.com/difference-between-crossover-frequency-and-recombinationfrequency/#:~:text=Crossover%20frequency%20is%20the%20type,th%20genes%20during%20the%20crossover. (Dec. 27, 2019 Posted by Dr.Samanthi)

International Human Genome Sequencing Consortium. Initial Sequencing and Analysis of the Human Genome. *Nature* **2001,** *409* (6822), 860–921.

Kalkavan, H.; Green, D. R. MOMP, Cell Suicide as a BCL-2 Family Business. *Cell Death Differ.* **2018,** *25* (1), 46–55. DOI: 10.1038/cdd.2017.188

Katsanis, N. The Continuum of Causality in Human Genetic Disorders. *Genome Biol.* **2004,** *5* (4), 229.

Lander, E. S.; Botstein, D. Mapping Mendelian Factors Underlying Quantitative Traits using RFLP Linkage Maps. *Genetics* **1987,** *121* (1), 185–199.

Lander, E. S.; Linton L. M.; Birren, B.; Nusbaum, C.; Zody, M. C.; Baldwin, J.; et al. Initial Sequencing and Analysis of the Human Genome. *Nature* **2001,** *409,* 860–921. 10.1038/35057062

Lodish, H.; Berk, A.; Zipursky, S. L.; Matsudaira, P.; Baltimore, D.; Darnell, J. Chapter 17, The Eukaryotic Cell Cycle. In *Molecular Cell Biology*, 4th Ed.; W. H. Freeman, 2000.

Lodish, H.; Berk, A.; Zipursky, S. L.; et al. Meiosis. In *Molecular Cell Biology*, 4th Ed.; W. H. Freeman: New York, 2000. https://www.ncbi.nlm.nih.gov/books/NBK21659/

Makałowski, W. The Human Genome Structure and Organization. *Acta Biochimica Polonica* **2001,** *48* (3), 587–598.

McPherson, J. D.; Marra, M.; Hillier, L.; et al. A Physical Map of the Human Genome. *Nature* **2001,** *409* (6822), 934–941.

Mitelman, F.; Johansson, B.; Mertens, F. The Impact of Translocations and Gene Fusions on Cancer Causation. *Nat. Rev. Cancer* **2017,** *17* (4)*,* 233–245.

Morgan Unit(s). In *Encyclopedic Reference of Genomics and Proteomics in Molecular Medicine*; Springer: Berlin, Heidelberg, 2006. https://doi.org/10.1007/3-540-29623-9_7875

Morgan, D. O. In *The Cell Cycle: Principles of Control*; New Science Press, 2007.

National Center for Biotechnology Information. Chapter 17, The Cell Cycle. In *The Cell: A Molecular Approach*, 2nd Ed.; 2021.

Pollard, T. D.; Earnshaw, W. C.; Schwartz, J. L.; Johnson, G. T. In *Cell Biology,* 3rd Ed., 2016; pp 697–699.

Sakharkar, M.; Chow, V.; Kangueane, P. Distributions of Exons and Introns in the Human Genome. *In Silico Biol.* **2004,** *4* (4), 387–393.

Sanger, F.; Nicklen, S.; Coulson, A. R. In *DNA Sequencing with Chain-Terminating Inhibitors*, Proceedings of the National Academy of Sciences of the United States of America, 1977; vol *74* (12), pp 5463–5467.

Shapiro, H. M. In *Practical Flow Cytometry*; Wiley-Liss, 2003.

Shendure, J.; Ji, H. Next-Generation DNA Sequencing. *Nat. Biotechnol.* **2008,** *26* (10), 1135–1145.

Sherr, C. J.; Roberts, J. M. CDK Inhibitors: Positive and Negative Regulators of G1-phase Progression. *Genes Dev.* **1999,** *13* (12), 1501–1512. DOI: 10.1101/gad.13.12.1501

Smit, A. F. The Origin of Interspersed Repeats in the Human Genome. *Curr. Opin. Genet. Dev.* **1996,** *6* (6), 743–748.

Urrutia, A. O.; Hurst, L. D. The Signature of Selection Mediated by Expression on Human Genes. *Genome Res.* **2003,** *13*, 2260–2264.

van Doeselaar, S.; Burgering, B. M. T. FOXOs Maintaining the Equilibrium for Better or for Worse. *Curr. Top. Dev. Biol.* **2018,** *127*, 49–103. https://doi.org/10.1016/BS.CTDB.2017.10.003

Visscher, P. M.; Brown, M. A.; McCarthy, M. I.; Yang, J. Five Years of GWAS Discovery. *Am. J. Hum. Genet.* **2012,** *90* (1), 7–24.

Wolf, Y. I.; Novichkov, P. S.; Karev, G. P.; Koonin, E. V.; Lipman, D. J. The Universal Distribution of Evolutionary Rates of Genes and Distinct Characteristics of Eukaryotic Genes of Different Apparent Ages. *Proc. Natl. Acad. Sci.* **2009,** *106*, 7273–7280. 10.1073/pnas.0901808106,

CHAPTER 8

Linkage, Crossing Over, and Recombination

MAHNOOR IMRAN[1], TANZEELA REHMAN[1], MARIA BUTT[1], MUHAMMAD NAVEED SHAHID[2], and UZMA QAISAR[1]

[1]*School of Biological Sciences, University of the Punjab, Lahore, Pakistan*

[2]*Department of Botany, Division of Science and Technology, University of Education, Lahore, Pakistan*

ABSTRACT

Morphological, physiological, and biochemical traits are controlled by genes present in DNA in living organisms. These traits can be transferred to the next generation by the inheritance of genes that they receive from their parents. The genes that are present on the same chromosome located closely together are usually transferred to the next generation together, and that is why they are called "linked" while genes located far apart on the same chromosome or present on different chromosomes are unlinked in their inheritance. Physical crossing over of chromatids occurs during meiosis I, which results in the recombination of genes between homologous chromosomes. This process results in the rearrangement of the heterozygous homologous chromosomes and hence produces new combinations. The frequency of crossing over depends on how close the two genes are on the same chromosome, that is, if the genes are close together, then there will be less chance of crossing over, and the amount of non-crossing over gametes will surpass the amount of crossing over ones. If the two genes are far from each other on the same chromosome, that is, present on the two opposite ends of the chromosome, then there will be more chance of crossing over, and both crossing-over and noncrossing-over frequencies will be the same.

Molecular Genetics, Structures, Mechanisms, and Functions, Volume 1: Principles of Gene Manipulation and Genomics. Tariq Ahmad Bhat, Abdul Rauf Shakoori, and Jameel M. Al-Khayri (Eds.)
© 2026 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

8.1 INTRODUCTION

DNA is the only molecule that serves as the genetic material in living organisms and carries the information to be transferred to the next generation. Each living cell contains DNA in its nucleus, which is organized in the form of physical structures called chromosomes. Genes are present on the chromosomes and carry genetic information. The DNA that forms chromosomes comprises genes and intergenic regions. In haploid organisms, only one set of chromosomes is present, while in diploid organisms, two sets are present. Diploid organisms under alternation of generation in which diploid organisms undergo meiosis to produce haploid gametes. Haploid gametes combine during the process of fertilization and produce diploid zygotes. During gamete formation, the recombination of chromosomal fragments and genes takes place through crossing-over events that are governed by linkage.

8.2 LINKAGE

Genetic linkage is the ability of the DNA sequences present close to each other on a single chromosome to be inherited together. During crossing over, there is less chance of separation of the genetic markers that are located close to each other on the same chromosome and are inherited together, unlike the genes that are far apart. Genetic markers present on different sets of chromosomes are unlinked. In other words, genes close to each other have a lesser chance of recombination (Hirschhorn and Daly, 2005).

Mendel described the loci pair behaving independently. For example, the segregation of seed-shaped alleles was independent of the seed color alleles. This was the basis of his second law of Independent Assortment. Further research revealed that Mendel's second law does not apply to all pairs of genes. the alleles located close to each other do not behave independently and are inherited together. The exception to Mendel's second law of independent assortment is known as linkage (Lobo and Shaw, 2008). Linkage is used as a tool to discover the position of genes located on a chromosome by genetic mapping. Linkage study is important to evolution and heredity. This phenomenon was first discovered in 1905 before scientists were unfamiliar with the phenomenon behind the inheritance of certain traits. It was later discovered that the genes are separated on a chromosome by physical distance, and their distance can be measured. Centimorgan (cM) is the unit

of linkage. Two genetic markers separated once per 100 meiotic products are known to have a distance of 1 cM.

According to Mendel's second law, every trait or gene is inherited independently. However, scientists found an exception to this rule. Rebecca Saunders, William Bateson, and Punnett, in 1905, worked on pea plants to discover the inheritance pattern of traits. They crossbred pea plants to study two genes—pea flower color (P for purple and p for red) and pollen shape (L for long, l for long).

In their experiments, the pure lines of flower color and pollen shape that were PPLL (dominant) and ppll (recessive) were crossed with each other forming PpLl lines. These lines were further self-crossed. The expected phenotype should be 9:3:3:1 according to the Mendelian genetics, but they observed an increased frequency of homozygous PL and pl; and reduced pL and Pl frequency (Table 8.1).

TABLE 8.1 Punnet Experiments Showing the Deviation from Mendel's Second Law. ⏎

Phenotypic and genotypic characteristics	Experimental frequency	Predicted as of 9:3:3:1 ratio
Purple and long (P_L_)	284	216
Purple and round (P_ll)	21	72
Red and long (ppL_)	21	72
Red and round (ppll)	55	24

It was observed that P and L alleles were linked similarly to the p and l because the P and L frequency to occur together and p occurring with l was more than the Pl and pL. This indicated that the two traits worked together in masking the effect of the other two traits. This showed that some traits are related to each other if they are present close to each other on a chromosome.

8.3 GENE MAPPING

Genes that are close to each other on a chromosome are said to be linked so they will be inherited together. The data from the genetic crosses can be used to find out how tightly two genes are linked to each other and to calculate the frequency of recombination. By finding this frequency of different gene pairs, the linkage maps can be made to calculate the distance of genes relative to each other on a chromosome. Organisms have more genes than chromosomes, that is, humans have approximately 19,000 genes on 23

chromosomes, and fruit flies have 13,000 genes on 4 chromosomes, which shows that every gene is not going to get its own chromosome. The genes that are nearby on the same chromosome will stick together and their alleles are supposed to be inherited as a pair. When genes are present on different chromosomes or located apart from each other on the same chromosome, they will independently assort, that is, when they form gametes, one allele for a gene will not affect the other allele for another gene. In an organism with double heterozygous genes (AaBb), all the possible kinds of gametes will form with the same frequency (25%) shown in (Fig. 8.1.

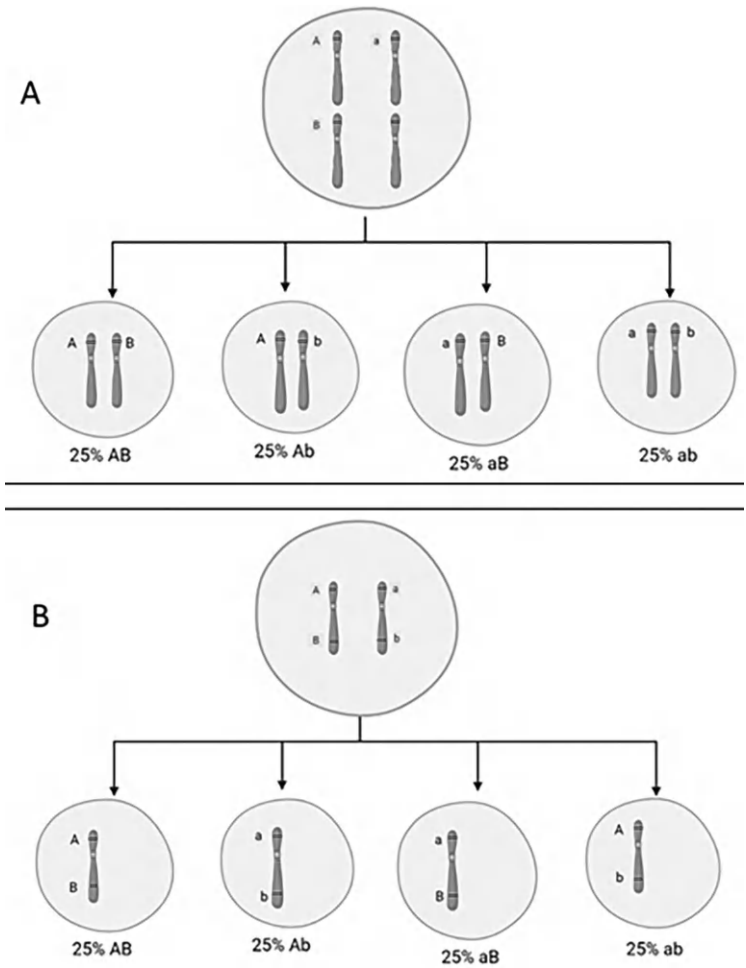


FIGURE 8.1 Genes on different chromosomes assort independently vs. genes on the same chromosomes but far apart also assort independently. ⏎

The random arrangement of the pair of homologous chromosomes in the course of meiosis is responsible for the independent assortment of the genes located on separate chromosomes. The chromosomes paired together having the same genes are known as homologous chromosomes, but the alleles for these genes may differ (Fig. 8.2). One chromosome from the pair originates from the maternal side and the other from the paternal. These pairs are separated during the first phase of meiosis and which side the maternal and paternal chromosome goes is totally random. When considering two genes, the resulting gametes will be of four types with equal frequencies.

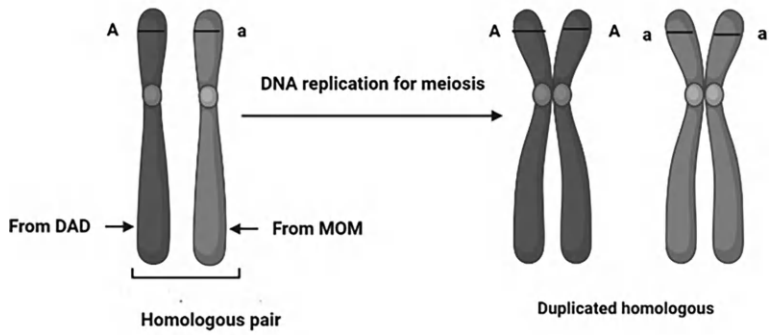


FIGURE 8.2 Homologous pair duplication for meiosis. ⏎

Crossing over occurs even when the chromosomes are present close to each other on the same chromosome, but the gametes produced in this case are of unequal proportions. This is because the alleles that are close will stick together and do not assort independently. The commonly produced gametes are the paternal ones that were already present in the parents before meiosis, and the rare kinds of gametes are also produced in lesser proportion as a result of crossing over producing recombinant configurations of alleles (Barbujani and Colonna, 2010). Crossing over between close genes is unique and does not happen commonly. During meiosis, the crossing over happens at random locations in chromosomes and depends on the distance between the genes, that is, in closely located genes, there is a lesser chance of a crossing-over event to take place as compared to genes further apart. So, the recombination frequency (linkage degree) can be used to measure the distance between two genes.

8.4 RECOMBINATION FREQUENCY AND LINKAGE MAPS

During meiosis, the frequency at which a solo crossing-over occurs between two genes is known as the recombination frequency. Recombination

frequency is used for creating genetic map and shows the linkage degree. It is denoted as θ. If recombination frequency is 1%, it means there is a 1-cM distance between the two loci. It gives the estimate of the actual genetic distance. In a double crossing over, there is no recombination because, in such circumstances, we cannot tell whether the crossing overs happened or not. If the two loci under analysis are very near a double crossing over may not occur very likely. While distances end up better, the probability of a double crossing over will increase (Ott, 1999).

Recombination frequency is not an exact degree of the physical distance between two genes on a chromosome. It only gives an estimation or rough idea about the distance between two genes. A gene pair is said to be far apart from each other if they have higher recombination frequencies while the pair is close to each other if the frequency of recombination is low.

The recombination frequency becomes maximum at 50%, that is, the genes are assorting independently and are unlinked. The maximum recombination frequency between two genes is 50%.

To map the distance between two genes, we need to include the recombination frequencies of the linked pair of genes "building up" a map of ranges between the two remote genes.

The recombination frequencies can be compared to figure out the order in which the genes are arranged on a chromosome. For example, if a chromosome has three genes, namely, A, B, and C, the arrangement of those genes on the chromosome is unknown. By calculating the recombination frequencies among all the possible combinations of genes (AB, AC, and BC), we will be able to calculate which gene lies closer to each other and which is further apart and the correct order of arrangement of the three genes. The pair having the largest recombination frequency will be the most apart (Fig. 8.3).

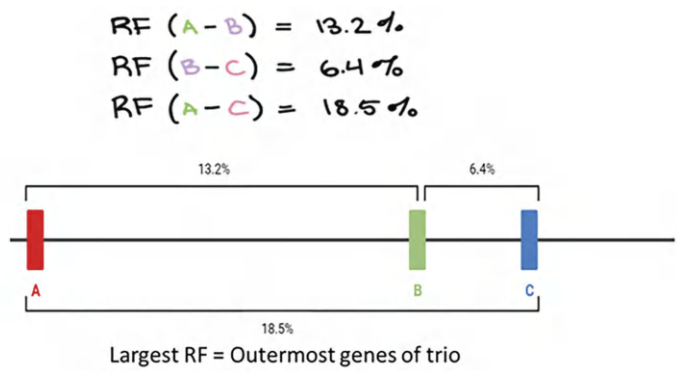


FIGURE 8.3 Recombination frequencies are based on those for fly genes v, cv, and ct. ⏎

By increasing the number of genes (adding more genes in A, B, and C, we can find out their association with A, B, and C), and in this way, we can predict the linkage map of the whole chromosome. In the case of linkage maps, the distance is measured in centimorgan.

A recombination frequency of 1% corresponds to 1 cM or 1 map unit. The immediate measure of recombination frequencies between two genes is not always the correct map distance between them. This is due to the fact that along with a single crossing over, double crossover can also take place.

Thomas Morgan expanded the knowledge on linkage through his study (Fig. 8.4). He observed that linked genes have different crossing over frequencies. It was later discovered that crossing over frequency depends on the distance between the genes, so the unit centimorgan to calculate the recombination frequency was named in his honor.

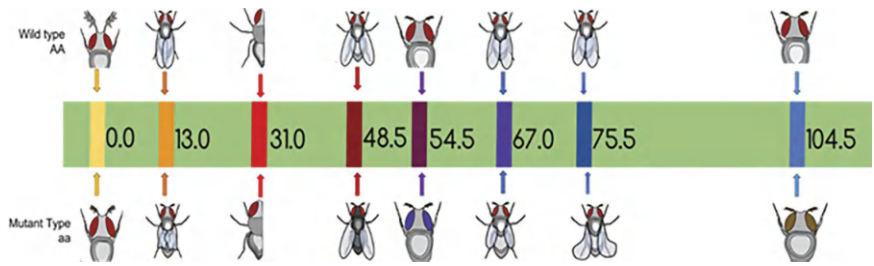


FIGURE 8.4 Genetic linkage map of *Drosophila melanogaster* by Thomas Morgan. The map provides important information about the theory of inheritance of chromosomes and was the first successful genetic map. The relative locations of alleles are presented in the map. The percentage of the crossing over that takes place between alleles is the same as the distance (centimorgan) between the genes. ⏎

Linkage maps were later discovered by Alfred Sturtevant. The recombination frequency between genes throughout the homologous chromosomes crossing over is referred to as linkage maps, that is, if there is more distance between two genetic markers, they will have higher recombination frequency (Campbell and Reece, 2005). Traditionally, the markers at the beginning used were from the coding region of DNA and had phenotypes that were detectable, such as enzyme manufacturing or eye coloration; later on, noncoding DNA sequences along with microsatellites or else the ones producing restrict fragment period polymorphisms (RFLPs) were also used. Different markers, along with other genes, can be discovered using linkage mapping for testing the linkage with the markers that are already known. In the beginning process of making a linkage map, the genes that are known to

be linked are gathered and assembled in a group. As more and more genetic markers are discovered and studied, they can be added to the group until a group covers a complete chromosome.

8.4.1 GENES AFFECTING RECOMBINATION FREQUENCY

Recombination frequency can be affected by mutations in genes that are involved in DNA processing. Mutation in bacteriophage T4 reduces the expression of DNA polymerase and increases recombination several folds in turn decreasing linkage. The increase of recombination in several folds is because of the replication errors caused by defective (mutated) DNA polymerase, for example, copy choice template switches. Mutations decreasing the expression of dCMP hydroxymethylase and DNA ligase also increase the recombination (Alberts et al., 2003). Mutation in genes encoding proteins for DNA binding proteins (gp 32) and nuclear function (gp47), reduces the recombination frequency and increases linkage. Recombination is reduced by a mutation in the bacteriophage uvsX gene, which is similar to *Escherichia coli* RecA gene, which is important for recombination.

8.5 LINKAGE EVALUATION

Linkage analysis is a technique that is used to analyze the chromosomal segments that segregate with the disorder phenotypes throughout the families. This could be used to record the quantitative qualities of genes.

Linkage study can be of two types; parametric or nonparametric. The conventional method is the parametric analysis, in which the LOD rating is used to determine the chance that a specific gene responsible for the disease is related to a genetic marker. This, in turn, evaluates the possibility that the gene for a particular disease and genetic marker segregating in a pedigree is due to linkage or by chance. Nonparametric linkage (NPL) evaluation demonstrates the possibility that an allele is identical in the affected individuals without any model assumption.

8.5.1 PARAMETRIC LINKAGE ANALYSIS

Morton developed a test for analyzing linkage in animals, humans and plants known as LOD score. It compares the likelihood of results if two loci are

linked and the likelihood if the results are by chance. Comparison of the likelihood of obtaining test results if two loci are linked or likelihood of observing that results by chance is done through LOD score. If the LOD score is positive, then there is linkage between the loci, while a negative LOD score indicates less chance of linkage. Computerized analysis of LOD score can be done to determine the complex pedigree of families to find the linkage between Mendelian traits or markers (Kruglyak et al., 1996). The steps of calculating the LOD score are creating a pedigree, estimating the frequency of recombination, then each estimate is assigned a LOD score, and the one with the highest LOD score is regarded as the best.

8.5.2 MEIOSIS INDICATORS

In whole genome sequencing with large pedigrees and dense genetic markers, it is easy to locate the recombination events. In this analysis, each part of the genome is assigned a meiosis indicator in every stage of meiosis to indicate which copy of the chromosome from a parent is involved in transmitting the gamete at the particular position (Andersen and Sekelsky, 2010), for example, if an allele from first copy of parent chromosome is transferred that meiosis will be assigned as "0" and if allele from second copy is transmitted it is assigned as "1." The two parental alleles came from the grandparents. These indicators are used to determine inheritance states which can be used to locate the genes.

8.6 GENETIC RECOMBINATION

Recombination in genetics, also recognized as genetic shuffling, is the genetic material exchange between two different organisms, producing offspring with unique traits that differ from both parents. During meiosis, recombination can lead to the production of novel traits that can pass from one generation to another. There are two types of naturally occurring genetic recombination; interchromosomal recombination and intrachromosomal recombination.

Interchromosomal recombination occurs when alleles having loci on different chromosomes assort independently, and intrachromosomal recombination occurs through crossing over (Daly and Minton, 1995). During the course of meiosis in eukaryotic cells, homologous chromosomes pair with each other for recombination. This can occur through information exchange

between chromosomes. The exchange can be nonphysical involving copying of DNA from one chromosome to another (without changing the DNA itself) or by breaking of DNA and then rejoining to produce new DNA strands.

Recombination might also arise in eukaryotes in the course of mitosis, which typically takes place between two sister chromatids that are synthesized after the replication of chromosomes. Both chromatids are identical to each other so no new combination of alleles is formed in this event. Recombination occurs between homologous chromosomes in both meiosis and mitosis. In the case of meiosis, recombination occurs between nonsister homologs. The process of recombination is commonly utilized for DNA repair.

8.6.1 GENE CONVERSION

The phenomenon of gene conversion requires the copying of DNA segments from one chromosome to another without any alterations in the DNA of the donor chromosome. It normally takes place in the fungal crosses where the meiotic products can be easily observed (Charlesworth and Campos, 2014). Recombination between sequences with no sequence homology occurs, which causes chromosomal translocation resulting in cancer.

Asexual bacteria and archaea also have recombination and DNA repair mechanisms. Recombinant DNA molecules can be produced in labs in vitro for research purposes. The adaptive immune response is a kind of recombination event allowing site-specific recombination to fight against a variety of new pathogens (Chen et al., 2007). Immunoglobulin (Ig) class switching is also a type of genetic recombination performed by the B cells, which converts the antibodies from class, that is, IgM, to the other isotype (IgG).

8.6.2 SYNAPSIS

Synapsis is a stage during meiosis involving homologous chromosome pairing, which is followed by recombination events. The process of recombination is initiated by a number of diverse enzymes. The key enzymes are the recombinases that catalyze the strand transfer in recombination. In *E. coli*, RecA is the main enzyme involved in repairing the double-stranded breaks (DSBs) in DNA. In archaea, the ortholog of RecA named RadA is present and performs the same function. In yeast two recombinases are involved in the repairing of DSBs. Mitotic and meiotic recombination is performed by Rad51 and a DMC1 protein that is particular for meiotic recombination.

8.6.3 RNA VIRUS RECOMBINATION

Genetic recombination can occur between two viral RNA genomes if they exist in the same cell. The diversity of RNA viruses due to recombination is important to survive in host cells and for immune evasion. The evolution pattern in *picornaviridae*, which is a single-stranded RNA virus, (e.g., poliovirus) was determined mainly by the RNA recombination. In the *retroviridae*, which is a single-stranded RNA virus (e.g., HIV), recombination helps in avoiding the damage in the RNA genome by strand switching during reverse transcription. Recombination also occurs in the *coronaviridae* (ssRNA) and *reoviridae*. RNA viruses' recombination seems to be involved in the protection against genome damage.

In the case of coronavirus, recombination can occur in distantly related subgenera. This is due to the formation of subgenomic mRNA that is formed because of the distinctive transcription mechanism in coronaviruses that involves template switching. During the replication of ssRNA of the poliovirus, recombination takes place with the help of RNA-dependent RNA polymerase that switches the strands during the synthesis of the negative strand. The exact mechanism of coronaviruses is not clear yet it plays a major role in the genetic diversity and capability of the coronavirus to transfer from one host to the other.

Recombination is used in genetic engineering where artificial or deliberate transfer of DNA segments to other organisms takes place to create a recombinant DNA. This recombination is used for gene targeting to add or remove a certain gene. This is mostly used by researchers to find out the role of specific genes. Recombination techniques can be used for protein engineering to produce better proteins.

8.6.4 RECOMBINATIONAL REPAIR

A variety of exogenous agents is responsible for DNA damage including X-rays, UV light, and chemical cross-linking agents. Repair of this damage can be done through homologous recombination repair (HRR). This suggests that the DNA damage caused by natural processes, for example, reactive oxygen species can be repaired by HRR. In humans, a deficiency of genes necessary for HRR during meiosis results in infertility. The deficiency of gene products for HRR, such as BRCA1 and BRCA2, increases the risk of cancer.

Transformation is the process in which genetic transfer between different bacterial cells of the same species occurs. It allows recombination by the integration of genetic material in recipient chromosomes from donor cells. This process is useful for dealing with DNA damage by repairing of recipient chromosome through HRR.

8.7 CROSSING OVER

During meiosis, the nonsister chromatids from homologous chromosomes tend to exchange their genetic material in a process known as crossing over. This results in new allelic forms of the gene in the daughter cells. Each diploid cell contains one copy of chromosome from the paternal gamete and the other copy from the maternal gamete. This pair of chromosomes is known as homologous chromosomes. When a diploid organism undergoes reproduction, haploid gametes are produced through meiosis. The frequency of crossing over is higher in meiosis as compared to mitosis in eukaryotes (Wang et al., 2015). Frans Alfons Janssens, in 1909, discovered crossing over and named it chiasmatype theory; later on, Thomas Hunt Morgan confirmed his theory and named it crossing over. Chromosomal crossing over is associated with the term chaisma. Crossing over value is the frequency of crossing over between two loci.

8.7.1 CLASS-I CROSSING OVER AND CLASS-II CROSSING OVER

In eukaryotes, crossing overs are generated by two pathways to repair DSBs that are; class-I crossing over and class-II crossing over. These two pathways are connected to each other, that is, class-I crossing over can recover the loss of class-II pathway. The class-I crossing over includes MutL homologs (MLH1 and MLH3) which repairs the majority of the double-stranded breaks. The class II pathway includes the endonuclease named MUS81, which is responsible for the remaining repair. The level of class I crossing over was higher in the mice that were MUS81 knocked out. However, the total counts of crossing over were normal at the chiasmata. The exact mechanism of this connection is still unclear.

Crossing over takes place at prophase 1 of meiosis which includes the interchange of genetic material between two nonsister chromatids of homologous chromosomes. It occurs after the replication ends at the interphase. After crossing over, half of the daughter cells will be paternal, and

half recombinant (Fig. 8.5). In 1931, McClintock and Creighton used maize plants' genetic markers to discover that during the crossing over there was a visible interchange of chromosomal markers.

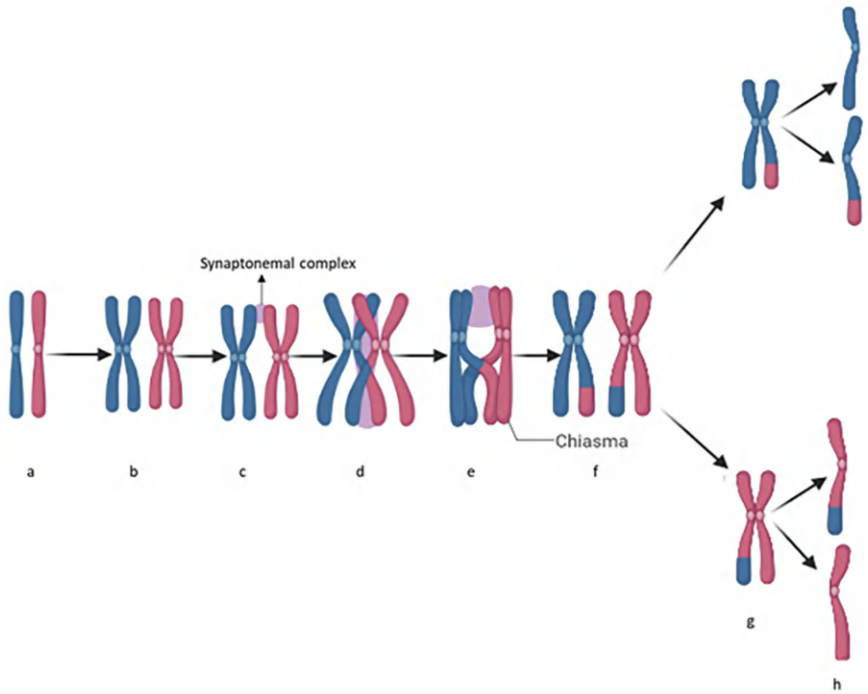


FIGURE 8.5 Crossing-over between two loci of homologous chromosomes during meiosis. (a) Before replication, (b) Premeiotic stage, (c) crossing over between nonsister chromatids of homologous chromosomes; bivalent formation, (d) Bivalent held at crossing over point by chiasma, (e) separation of chromosomes during anaphase, (f) End of meiosis resulting in four products. ⏎

In the premeiotic stage, the homologous chromosomes arrange them during the interphase. The crossing over takes place at the pachytene stage during meiosis. At this stage, the homologous chromosomes are paired with each other at a point called chiasma, and a bivalent is formed. During the diplotene stage, the homologous chromosomes separate, and the chiasma that holds the two chromosomes together becomes visible. After which the homologs are separated during the anaphase I. The involvement of diplotene chiasmata in crossing over between two chromatids can be observed in the grasshopper spermatocyte cells (Kumar et al., 2012).

If more than one crossing over occurs in the same bivalent, it is very unlikely that they occur in close vicinity to each other, that is, they usually occur far from each other. This phenomenon is known as chiasmata interference. Normally if a crossing over occurs at a particular position, it will inhibit another crossing-over in its surroundings. If two genes are close to each other on a chromosome, they will have a greater probability of being linked, and there will be less chance for a crossing-over to occur between them. So, the frequency of crossing-over between two genes can be used to estimate the distance between them and map their position on a chromosome. This is referred to as a linkage map. The appearance of chiasmata in a chromosome is mostly localized, that is, some parts of the chromosome are more likely to undergo crossing-over than others.

8.7.2 MOLECULAR MECHANISM

A DNA duplex is present in each chromatid that undergoes breaking and joining in the process of crossing-over (Muller, 1916). While the exact mechanism is still not known, it is considered that crossing-over generally involves the following steps. Nicking in one or both strands of the DNA molecule, after which a heteroduplex is formed between the nonsister chromatids, followed by the formation of chiasmata, which is resolved by breaks resulting in crossing-over.

On a molecular level, crossing-over begins with a double-strand break in one of the DNA molecules. This double-strand break can occur naturally through agents like radiation or carcinogens, or through the action of specific proteins (Krishnaprasad et al., 2015). Subsequently, exonucleases remove short stretches of nucleotides in the $5' > 3'$ orientation from both the strands of DNA. This leads to two overhanging single-stranded regions on which proteins called recombinases bind and catalyze recombination. Recombinases catalyze the insertion of complementary single-stranded regions into the strands that are suitable for base pairing. The proximity of nonsister chromatids during prophase I allows this single-stranded region to recombine with the homologous chromosome. The first invading strand behaves like a primer and synthesizes a double-stranded region utilizing its nonsister chromatid as a template. This leads to its complementary strand getting displaced and base pairing with the second single-stranded region that was initially generated by the exonuclease. This results in the formation of a structure called the Holliday junction, after the scientist who first proposed

that such a junction could explain both crossing over and gene conversion where a heterozygous gene locus becomes homozygous during cell division. Holliday junctions can also be seen microscopically as "chiasma" toward the end of prophase I, which is visible till the end of anaphase I. Holliday junctions are stabilized and resolved through proteins that modulate genomic manipulation like MSH4 and MSH5.

8.7.3 SEX CHROMOSOMES

The differentiated X and Y sex chromosomes in human males and many animals have small regions near one tip which undergo pairing and crossing over at meiotic prophase I. Electron microscopy of the pachytene XY reveals the formation of a short synaptonemal complex segment with a recombination nodule in the majority of cases; the presence of a chiasma between the X and Y at metaphase I indicates the occurrence of crossing over. An obligatory crossing over in the XY bivalent is necessary to ensure regular segregation of X and Y to opposite poles at anaphase I (Nedelcu et al., 2004). The rare occurrence of XX males in some cases is accounted for by abnormal recombination events outside the pseudo-autosomal region which have transferred the male sex-determining gene from the Y to the X chromosome.

8.8 CONCLUSION

Genetic recombination is found in all groups of living organisms, whether their genetic material is DNA or RNA. The natural sequence rearrangements observed in the genomes and in the experimental materials demonstrate that recombination plays an important role in genetic diversity and evolution. Various molecular mechanisms used in genetic recombination involve the nicking and rejoining of single or double strands. DNA viruses utilize mechanisms of homologous and nonhomologous recombination utilizing the machinery of host cells or by synthesizing their own proteins. However, further studies are required for RNA viruses to obtain more information about genetic recombination pathways. Crossing over allows limitless recombination of genes and genetic materials within a species while transmitting genetic material from parents to the offsprings, which results in variability. However, the frequency of recombination is not uniform in the genomes. It varies along the chromosomal length between sexes and species.

KEYWORDS

- **linkage**
- **inheritance**
- **crossing over**
- **meiosis I**
- **homologous chromosomes**

REFERENCES

Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. Molecular Biology of the Cell. *Scand. J. Rheumatol.* **2003,** *32*, 125–125.

Andersen, S. L.; Sekelsky, J. Meiotic Versus Mitotic Recombination: Two Different Routes for Double-Strand Break Repair: The Different Functions of Meiotic Versus Mitotic DSB Repair are Reflected in Different Pathway Usage and Different Outcomes. *Bioessays* **2010,** *32*, 1058–1066.

Barbujani, G.; Colonna, V. Human Genome Diversity: Frequently asked Questions. *Trends Genet.* **2010,** *26*, 285–295.

Campbell, N.; Reece, J. The Chromosomal Basis of Inheritance. In *Biology*, **2005;** pp 274–292.

Charlesworth, B.; Campos, J. L. The Relations between Recombination Rate and Patterns of Molecular Variation and Evolution in Drosophila. *Ann. Rev. Genet.* **2014,** *48*, 383–403.

Chen, J. M.; Cooper, D. N.; Chuzhanova, N.; Férec, C.; Patrinos, G. P. Gene Conversion: Mechanisms, Evolution and Human Disease. *Nat. Rev. Genet.* **2007,** *8,* 762–775.

Daly, M. J.; Minton, K. W. Interchromosomal Recombination in the Extremely Radioresistant Bacterium Deinococcus radiodurans. *J. Bacteriol.* **1995,** *177*, 5495–5505.

Hirschhorn, J. N.; Daly, M. J. Genome-Wide Association Studies for Common Diseases and Complex Traits. *Nat. Rev. Genet.* **2005,** *6*, 95–108.

Jones, G. H.; Franklin, F. C. H. Meiotic Crossing-Over: Obligation and Interference. *Cell* **2006,** *126*, 246–248.

Kohler, R. E. In *Lords of the Fly: Drosophila Genetics and the Experimental Life*; University of Chicago Press, 1994.

Krishnaprasad, G. N.; Anand, M. T.; Lin, G.; Tekkedil, M. M.; Steinmetz, L. M.; Nishant, K. T. Variation in Crossover Frequencies Perturb Crossover Assurance without Affecting Meiotic Chromosome Segregation in Saccharomyces cerevisiae. *Genetics* **2015,** *199*, 399–412.

Kruglyak, L.; Daly, M. J.; Reeve-Daly, M. P.; Lander, E. S. Parametric and Nonparametric Linkage Analysis: A Unified Multipoint Approach. *Am. J. Hum. Genet.* **1996,** *58*, 1347.

Kumar, A.; Bassi, F. M.; Paux, E.; Al-Azzam, O.; de Jimenez, M. M.; Denton, A. M.; Gu, Y. Q.; Huttner, E.; Kilian, A.; Kumar, S. DNA Repair and Crossing over Favor Similar Chromosome Regions as Discovered in Radiation Hybrid of Triticum. *BMC Genom.* **2012,** *13*, 1–12.

Lobo, I.; Shaw, K. Discovery and Types of Genetic Linkage. *Nat. Educ.* **2008,** *1*, 139.
Morgan, T. H. *Nobel Prize in Physiology or Medecine,* 1976, p 153.
Muller, H. J. The Mechanism of Crossing-Over. *Am. Nat.* **1916,** *50,* 193–221.
Nedelcu, A. M.; Marcu, O.; Michod, R. E. Sex as a Response to Oxidative Stress: A Twofold Increase in Cellular Reactive Oxygen Species Activates Sex Genes. *Proc. R. Soc. Lond. B Biol. Sci.* **2004,** *271*, 1591–1596.
Ott, J. In *Analysis of Human Genetic Linkage*; JHU Press, 1999.
Reece, J. B.; Meyers, N.; Urry, L. A.; Cain, M. L.; Wasserman, S. A.; Minorsky, P. V. In *Campbell Biology Australian and New Zealand Edition*; Pearson Higher Education AU, 2015.
Stapley, J.; Feulner, P. G.; Johnston, S. E.; Santure, A. W.; Smadja, C. M. Variation in Recombination Frequency and Distribution Across Eukaryotes: Patterns and Processes. *Philos. Trans. R. Soc. B Biol. Sci.* **2017,** *372*, 20160455.
Wang, S.; Zickler, D.; Kleckner, N.; Zhang, L. Meiotic Crossover Patterns: Obligatory Crossover, Interference and Homeostasis in a Single Process. *Cell Cycle* **2015,** *14*, 305–314.



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CHAPTER 9

Molecular Techniques and Methods for Mutation Detection and Screening in Plants

MAHAM AKRAM[1] and ALI RAZZAQ[2]

[1]*Centre of Excellence in Molecular Biology (CEMB), University of the Punjab, Lahore, Pakistan*

[2]*Agronomy Department, University of Florida, Gainesville, USA*

ABSTRACT

The genomic and genetic studies impart a worthy insight into the structure, inheritance, function, and organization of genes. The gained knowledge about plants by analyzing the variety of their genes is beneficial for mutation breeding. The majority of genetic changes in the genome that result in the desired crop enhancement features are caused by mutations. Over the past few decades, breeders are using induced mutation techniques compared to spontaneous mutations to enhance crop quality, yield, and stress resilience. Mutation induction can be done using mutagens, which are either chemicals, including ethyl methanesulfonate (EMS), or physical-like gamma radiations in nature. Conventionally, mutation detection was based on forward genetic approaches that were based on qualitative analysis such as plant stature, color, size, texture variation of seeds, leaves and flowers of mutants. Later on, reverse genetic approaches (genotypic analysis of mutant lines) were applied in mutation detection and screening. In this study, we comprehensively discuss the several mutation detections and screening techniques that have been extensively employed in crop research like single-strand conformational polymorphism, temperature and conformation-based

Molecular Genetics, Structures, Mechanisms, and Functions, Volume 1: Principles of Gene Manipulation and Genomics. Tariq Ahmad Bhat, Abdul Rauf Shakoori, and Jameel M. Al-Khayri (Eds.)
© 2026 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

electrophoresis, high-resolution melting analysis, dot-blot analysis, MutMap approach, MutChromSeq approach, exome capture approach, cleavage analysis, targeting-induced local lesions in genomes (TILLING), quantitative trait loci sequencing, whole-genome sequencing, and next-generation sequencing.

9.1 INTRODUCTION

Biotic and abiotic stresses cause severe impacts on the growth of plant, reproduction, and yield (Raza et al., 2019). To overcome this problem, some in vivo changes can make plants adaptable to such harsh environmental changes (Gull et al., 2019). Genetic variation is one of the solutions to this problem that makes plants adaptable to survive in a harsh environment and make them more beneficial (Govindaraj et al., 2015; Holme et al., 2019). Mutations are genetic variations that can either be spontaneous (natural) or induced (artificial), which can help to develop superior crop varieties plant (Sigurbjörnsson, 1983; Hershberg, 2015). A single-nucleotide insertion/deletions (indels), point mutations, and frameshift mutations in a genetic sequence result in a genetic transformation. Naturally induced mutations are usually in low frequency while induced mutations have a higher frequency rate that helps in plant survival, coping with harsh environmental conditions, improving crop quality and yield (Gottschalk and Wolff, 2012; Dooner et al., 2019).

Various studies are based on retrieving the genetic sequences of the mutants. The sequence availability can help to know the basic knowledge and the role of any specific gene and provide a way forward for designing new crop cultivars. Previously, mutations were detected by observing the phenotypic characters including plant stature, seed color, leaf shape, and grain size. But now, tremendous innovation has been attained in genome sequencing approaches that can assist in identifying the mutations more precisely and accurately. With the advent of genomic and molecular tools, it is easier to identify all types of mutants, either natural or induced. These new-era techniques made it possible to have a comparative phenotypic and genotypic study and overcome the disparity between phenotype and genotype (Arif et al., 2010).

Natural mutations are known as spontaneous mutations that occur without any artificial induction. A spontaneous genetic variation is observed and characterized very rarely in few organisms such as in *Drosophila*

melanogaster in which mutation occurs due to large-scale chromosomal changes. In plants, it has been characterized in maize (*wx* gene) which can be identified by endosperm traits (Dooner et al., 2019). Chemical and physical agents are employed to create artificial mutations using mutagens. These mutagens are implicated directly either on seeds or other organs of a plant to obtain heritable mutations. The practice of mutant breeding dates back more than 70 years, in order to achieve genetic diversity and to produce mutant lines with enhanced beneficiary characteristics (Stefano and Edoardo, 2004).

DNA mutations are crucial for plant breeders to manipulate the genome of any crop species. However, alterations in chromosome structure are also valuable, leading to a high recombination rate eventually causing the production of unique off-springs and breaking undesirable linkages. Physical mutagens are mostly ionizing radiations that are utilized for mutation induction. These mutagens have the capability of increasing the spontaneous mutation rate by 1000 to 1 million fold. More than 70% of mutant plant cultivars are produced through physical mutation techniques. Seeds, stem cuttings, spores, and pollens are treated generally with cobalt-60 source or irradiated in gamma fields/X-ray chambers for mutation induction (Ahloowalia amd Maluszynski, 2001; Predieri, 2001).

In chemical mutations, chemicals like sodium azide, colchicine, and ethylmethanesulfonate (EMS) are used with different concentrations for different time durations to achieve targeted mutations. After the induction of the mutation, the next step is its detection and screening (Abaza et al., 2020).

9.2 FORWARD AND REVERSE GENETIC APPROACH

The phenotypic characterization of targeted traits in a mutant line population is known as the forward genetic approach, and genetically characterization is known as a reverse genetic approach. While determining the appropriate population size for achieving a specified coverage of gene for knockdown or knockout studies, the mutation effect on gene functioning along with the effect of mutation frequency should be taken into account (Kazama et al., 2011).

In gene pool, a broad range of genetic diversity and variability is the basic necessity for genetically modified crop species. Traditional breeding is reliant on the accessibility of adequate variants, obligatory for enhancement of beneficial traits of plants (Cassells and Doyle, 2003; Chaudhary et al., 2019). Nevertheless, it is challenging to have sufficient resources for conventional

breeding and for this, a long time is required to attain the desire results. To address this problem, induced mutation techniques are applied, which ensures promising results (Oladosu et al., 2016; Prasanna and Jain, 2017).

Induced mutagenesis plays a vital role in speed-up breeding programs by manipulating the targeted genes. It contributes to the rapid improvement of plant features by supplying novel, crop-friendly alleles and a wide spectrum of genetic variation. The advantage of induced mutation over spontaneous mutation is that it can increase genetic variation frequency. These mutant alleles work as a source of developing elite crop genotypes with greater genetic variation and may help through functional genomics in identifying unique candidate genes (Kumawat et al., 2019). The mutants produced are precisely the same as the parent plant except for only a few traits that are modified for enhancing crop characteristics that are proven eco-friendly, reliable, and nonhazardous. Plant breeders are using these mutation techniques to find novel genetic sources for stress tolerance (biotic and abiotic), traits with more nutritional values, high yield ability, and resilience for herbicide tolerance (Chaudhary et al., 2019).

Induced mutagenesis has been successfully utilized to identify and map important quantitative trait loci (QTLs) and genes, as well as to investigate the role of specific genes and metabolic, biochemical, and molecular processes in the evolution of mutant phenotypes (Jankowicz-Cieslak et al., 2017). Knowing about functional genomics and regulatory mechanisms of mutated gene potentiate possibilities of crop improvement by using advanced breeding technologies. The only subjected phenotype is required in the forward genetic approach to discover the fundamental genes/mutations/QTLs using genetic mapping, positional cloning, fine mapping, and other recent genomic techniques (Pereira et al., 2020; Schneeberger, 2014). For finding fundamental mutations, a conventional technique known as positional cloning is applied, which is very expensive, time-taking, and tedious because of low recombination rates in targeted regions. The advancements in the forward and reverse genetic methods have increased the applications of mutation strategies in advanced breeding for crop development in the current era (Sahu et al., 2020).

9.3 MUTATION FREQUENCIES

Mutation frequencies in crop research have been determined usually based on observable phenotypes. For instance, color variations of leaves, flowers,

pods and seeds, changes in development, the height of plant are essential points in phenotypic mutation detection without the need of genetic information (Laskar and Khan, 2017). Induced mutations are usually done using physical (ion-beams, radiations) and chemical mutagens (EMS, sodium azide), so each mutagen category has a different impact factor in mutation frequency. A model plant *Arabidopsis thaliana* is used for studying mechanisms of DNA mutations, sequences, and frequencies of mutant alleles. Many elements, both related and unrelated, have an impact on the frequency and range of mutations (Jo and Kim, 2019).

9.3.1 LINEAR ENERGY TRANSMISSION

It has been proved from various studies that the frequency and mutation type (seen mainly in M2 and later generations) depends largely on linear energy transmission (LET) of the radiation used for mutation induction. It is described as the amount of energy that is transferred to materials when radiation travels a certain distance. Fast neutron and heavy ion beams are the most common high-linear energy transmission radiation types, although low-category X-rays and gamma rays are also included (Jo and Kim, 2019; Tanaka et al., 2010).

In the result of using high-linear energy transmissions radiations, clustered DNA damages are produced, which includes double- and single-strand breaks (DSBs and SSBs) along with assorted DNA lesions. In the case of low-energy transmission radiations, single-strand breaks are formed along with diversity in sugar or base lesions. However, DSBs are also observed in some cases that originate as the result of DNA repairing of multiple lesions. The differences in mutation induction by using high- and low-LET radiations arise mainly because of three factors that included the DNA repair efficiency, the frequency of DSBs, and mechanism of non-homologous end-joining (Ballarini et al., 2008; Jo and Kim, 2019). It was reported by Kikuchi et al. (2009) that eight-fold more chromosomal breaks are produced by 50-Gy neon beams as compared to X-rays. Hence, radiation sensitivity is also a factor in mutation induction. Another study revealed that the ratios between INDELS, single-base substitutions, and small variations also depend on LET. A wide range of phenotypic variations are observed, which showed that the phenotypic spectrum of mutations is also affected by LET. For example, carbon beam irradiation produces a much broader range of color variation of flowers as compared to gamma irradiations (Okamura et al., 2003).

9.3.2 RADIATION FREQUENCY

The first generation (M1) obtained after any mutagenic exposure is chimeric in nature and includes a variety of DNA mutations. The fertility and lethality in the first generation (M1) is considered with respect to radiation dose for generating M2 population. So, the determination of the lowest dose of radiation which produces a promising result of induced mutation is necessary (Jo and Kim, 2019). The relationship between the survival rate and the frequency of phenotypic mutations in M1 and M2 individuals of rice was examined in a study by Yamaguchi et al. (2009). The study has shown that the frequency of mutation significantly increases in the M2 plant when the survival rate decreases between 90 and 100%, and a gradual increase is seen when the survival rate falls below 90%. Together with LET, the frequency of mutation is directly correlated with the number of ion beam particles. Because there were more particles present while utilizing 30 KeV m-1 carbon beam LETs than 22.5 KeV m-1, a higher mutation frequency was seen in Arabidopsis thaliana. Choice of radiation dose is directly proportional to plant survival and a higher dose can be lethal regardless of the desired mutation results (Kazama et al., 2011).

9.3.3 CONDITION AND TYPE OF PLANT TISSUE

The mutation spectrum and frequency by any mutagen treatment also depends on the type and condition of plant tissues. The mutagenic treatment applied on a plant present in hydration state varies from a plant present in a normal dry state in case of radiation exposure as DNA damage is based on hydroxyl radicals produced from molecules of water. For example, below 14%, water content reportedly upsurges the barley seeds' sensitivity to X-rays and gamma rays. Mutation variation is also dependent on the hydration state. A study on radiation-induced mutations in rice plants at the level of the microspore gave more promising and required results as compared to radiation employed on reproductive tissues and seeds (Jo and Kim, 2019).

9.4 MUTATION DETECTION AND SCREENING TECHNIQUE

Since most of the mutations are based on single-nucleotide polymorphisms, various techniques are in use to detect the single-nucleotide mutations.



FIGURE 9.1 Overview of mutation screening and detection methodologies.

9.4.1 SINGLE-STRAND CONFORMATION POLYMORPHISM (SSCP)

Single-strand conformation polymorphism (SSCP) is an efficient method for SNPs detection. Mutation detection is built on the difference between the two alleles present at chromosome loci. Various confirmation differences arise in single-stranded DNA as nucleotides vary. For this, electrophoresis is run, and DNA is resolved on the gel (non-denaturing) to obtain visible results. Less than 500 bp point mutations are possible to be detected by SSCP (Zheng et al., 2016).

The SSCP can be coupled with polymerase chain reaction (PCR), to obtain more precise and accurate results as compared to simple SSCP. In the PCR-SSCP method, deoxyuridine triphosphates (dUTPs) with fluorescent dyes are used, which then can be analyzed easily. Another form of SSCP known as polymerase chain reaction-restriction fragment (PCR-RF-SCCP) is employed to find mutations in lengthy DNA sequences. Its foundation is a synthesis of two approaches, PCR-RFPL and SSCP analysis. Restriction enzymes digest approximately 2 kb PCR product into <500bp short DNA strands that are then electrophoresed under SSCP conditions in polyacrylamide gels. Figure 9.2 explains the basic principle behind SSCP (Beier, 1993; Hayashi, 1992; Iwahana et al., 1996; Sato and Nishio, 2003).

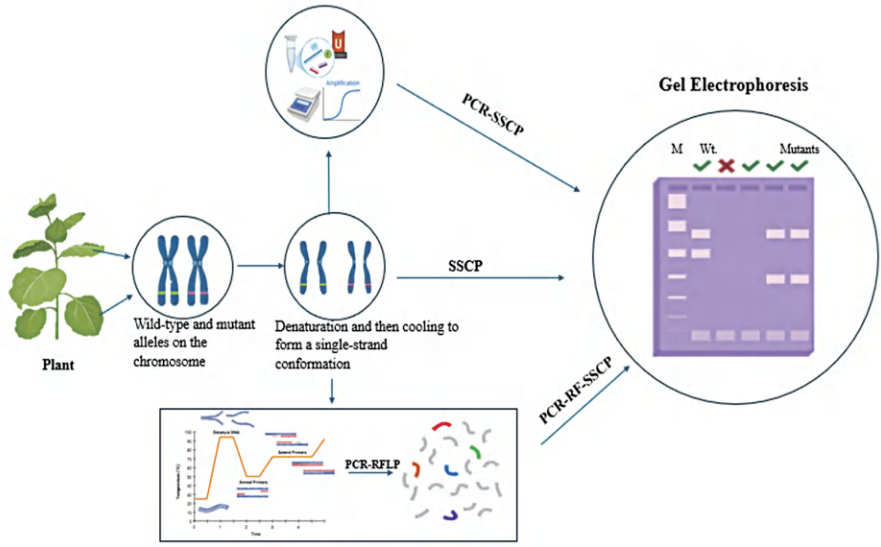


FIGURE 9.2 Schematic explanation of SSCP methodology. ⏎

9.4.2 MUTATION DETECTION THROUGH VARIOUS ELECTROPHORESIS TECHNIQUES

DNA detection is usually achieved by gel electrophoresis. It essentially works by dividing DNA molecules into different groups based on their size and conformation. There are four types of electrophoresis that are used for mutation detection.

i. Temperature gradient gel electrophoresis (TGGE)
ii. Temperature gradient capillary electrophoresis (TGCE)
iii. Conformation sensitive gel electrophoresis (CSGE)
iv. Conformation-sensitive capillary electrophoresis (CSCE)

9.4.2.1 TEMPERATURE GRADIENT GEL ELECTROPHORESIS (TGGE)

In temperature gradient gel electrophoresis (TGGE), mutations are detected by comparing the mobility of DNA molecules in the gel. For example, the mobility of wild-type DNA molecule (having no mutation) differs from the complete mutant and heteroduplex combination. The difference arises due to the melting temperatures. The temperature of heteroduplex varies; due to this, four bands are visible on the gel. One is similar to wild type, while the other is similar to a mutant, which is either above or below the wild type

band and the remaining two heterozygous mutant bands are always above the wild-type band. The heteroduplex bands are spaced approximately 3 mm apart, while the observed separation between wild type and mutant is 2 mm, according to the study.

9.4.2.2 TEMPERATURE GRADIENT CAPILLARY ELECTROPHORESIS (TGCE)

The temperature gradient capillary electrophoresis (TGCE) is just the same as TGGE with the difference of only resolving material. In TGCE, capillaries are used for resolving DNA on the basis of their melting temperatures (Peters and Robinson, 2010; Wartell et al., 1990).

9.4.2.3 CONFORMATION-SENSITIVE GEL ELECTROPHORESIS (CSGE)

CSGE (conformation-sensitive gel electrophoresis) is a differential, nonenzymatic technique for SNP detection which is based on DNA sequence. In a DNA pool, where both mutant alleles and wild-type alleles are present, the PCR results will show random annealing between amplicons that may or may not possess any kind of mutation; therefore, various duplex formations will occur. One of the homoduplex formations results from annealing of wild type (wt/wt) or mutant (mt/mt). Another one is the formation of homoduplexes, which are formed because of wild-type/mutant annealing. The heteroduplexes formed will migrate at a different pace because of mismatch formation as compared to homoduplexes, during electrophoresis (Gady et al., 2009). The mild denaturing environment in electrophoresis causes changes in the conformation of the double-helix DNA strand due to the presence or absence of single-nucleotide polymorphism. Amplicon size, mismatch location, nature of the mutation, and context of the sequence are limiting factors in CSGE, affecting mutation detection. It is challenging to differentiate between the homoduplex and heteroduplex bands and even those mismatches that are present in close vicinity of amplicon end. So, CSGE is only preferable for small indels detection, not for large ones (Ashton, 2011; Markoff et al., 1998; Körkkö et al., 1998). The method of CSGE is illustrated in Figure 9.3.

9.4.2.4 CONFORMATION-SENSITIVE CAPILLARY ELECTROPHORESIS (CSCE)

Another form of CSGE is conformation-sensitive capillary electrophoresis (CSCE). In this electrophoresis, capillaries (ABI 3130 × L capillaries filled

with CAP) are used instead of gel. In addition, tagged PCR primers are used, which help in amplicon labeling. The cellulose acetate propionate (CAP) is a semi-denaturing polymer, allowing mutation detection in targeted fragments. This makes CSCE more precise than CSGE (Ashton, 2011).

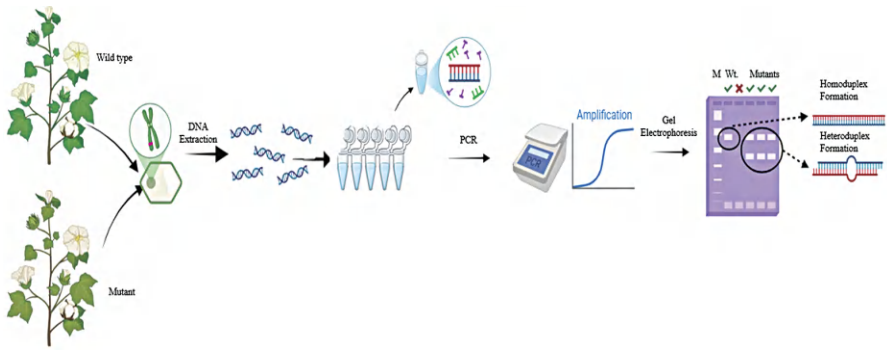


FIGURE 9.3 Schematic explanation of CSGE methodology. ⏎

9.4.3 HIGH-RESOLUTION MELTING ANALYSIS

HRM (high-resolution melting analysis) is a rapid high-throughput method for mutation detection without the need for sequencing. It is grounded on melting curve analysis of dsDNA. It is carried out after PCR in which double-stranded amplicons are subjected to denaturation. The first step in this analysis is PCR, which amplifies DNA in the availability of a fluorescent dye that binds to the DNA strand. It fluoresces more when bound to DNA and less when unbound. After performing PCR, the amplicon, usually 50–500bp long, is subjected to melting at an increasing temperature range from 0.01°C–0.2°C. During this process, the fluorescent dye is released by amplicons, and these fluoresces emission values are used for making a melting curve against gradually increasing temperature. The curve shape depends on the sequence, amplicon length, compatibility of DNA strand, and GC contents (Dujols et al., 2006; Herrmann et al., 2006; Herrmann et al., 2007; Muleo et al., 2009).

In comparison to CSCE, HRM is a faster, more cost-effective, and more sensitive method of detecting mutations. It is successfully used to identify both spontaneous and induced mutations in a wide range of plant species by evaluating genotypes independently or in combination. Moreover, HRM has been used in marker-assisted selection, barcoding, gene mapping, and DNA fingerprinting. The primary drawback of HRM is its inability to distinguish

between the identical SNP at two different loci in two separate samples (McKinney et al., 2010).

9.4.4 DOT BLOT SNP ANALYSIS

A known mutation can be detected via dot blot SNP analysis in which DNA is extra clear from the samples with the targeted sequence being amplified through a polymerase chain reaction, and dot blotted on a nylon membrane. The competitive oligonucleotides, along with unlabeled oligonucleotides, are used for hybridization with the dot blotted PCR product. Labeled oligonucleotides get hybridized with the target and give signals that eventually are detected visually. (Shirasawa et al., 2006; Widyawan et al., 2019). Figure 9.4 shows the schematic overview of dot-blot SNP analysis.

9.4.5 CLEAVAGE ANALYSIS

In molecular biology, enzymes and chemicals are utilized for cutting DNA into pieces from the specific site. These chopped pieces of DNA can be used for several objectives, i.e., cloning, gene modification, and mutation study. At present, this method is modified, and a new one is now known as TILLING (Tabone et al., 2010; Till, 2010).

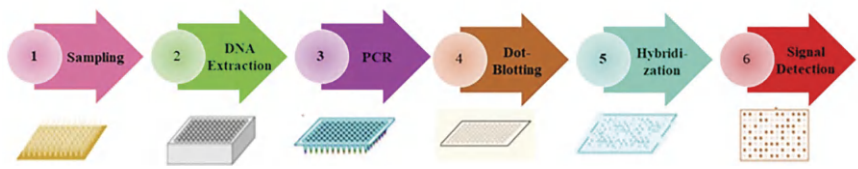


FIGURE 9.4 Genomic DNAs from samples are extracted. PCR is done to amplify specific gene sequences. The amplified product is dot-blotted on a membrane, and then hybridization is done with labeled oligonucleotides. The signals produced by fluorescent dyes are detected eventually. ⏎

9.4.6 TILLING FOR MUTANT DETECTION

Targeting induced local lesions in genomes (TILLING) is a reverse genetic approach for mutation detection. It was established first in 2000 by utilizing *Arabidopsis thaliana* as a model plant (Kurowska et al., 2011). In this method, the mutation is induced in the object artificially through chemical

mutagens, mainly EMS. After this, the DNA from the second generation is extracted, and PCR is performed. The amplicons are heated and cooled, slowly leading to the formation of heteroduplexes. The mismatches are detected and a bubble is formed at that site, which is then treated with restriction enzymes. The digested product is analyzed through various systems Li-Chor, agarose gel electrophoresis, polyacrylamide gel electrophoresis, etc. Another form of TILLING is known as Eco-TILLING in which natural mutations are studied. DEcoTILLING is a combination of both TILLING and EcoTILLING with the difference in the only detection system, which is economical (Garvin and Gharrett, 2007; Nieto et al., 2007; Rout, 2020).

Some mutagenesis techniques like fast neutron technique can result in large-scale DNA deletions. These deletions were detected previously through Deletagene high-throughput PCR. At present, this method is modified, and now it is known as De-TILLING (Kurowska et al., 2011). There are three aspects of De-Tilling. First, different mutagenetic agents are used to produce DNA deletions of various sizes. Second, DNA is pooled to save time from doing multiple PCR. Third, amplification of mutant allele through those technologies allows their amplification in DNA pools and recognizes deletions within an allele (Li et al., 2002).

When the deletion region comprises large DNA bases, i.e., 1–12 kb, there are lesser chances that wild-type (WT) allele is amplified more than the mutant allele. But when the deletion is around 1–2 kb, and the target region is about 2–3 kb, then the chances of wild-type allele amplification are very high. Specific procedures are the requirement for mutation detection as wild-type allele amplification context is more than the mutant one. So, to solve this problem, DNA pooling is executed in order to have an equal amount of DNA in each chamber, and then the process is implicated further (McCallum et al., 2000; Till et al., 2007; Uauy et al., 2009). Figure 9.5 explains the TILLING general method.

9.4.7 WHOLE GENOME SEQUENCING

Whole Genome Sequencing (WGS) is the technology through which the whole genome of an organism can be sequenced at a time along with DNA present in mitochondria and chloroplast. The whole sequenced genome is accessible, and the organism genome encyclopedia allows any kind of studies related to SNP detection, association studies, and gene functional analysis can be performed. This knowledge also provides a keen interest for a biologist to study evolutionary aspects. Therefore WGS-based

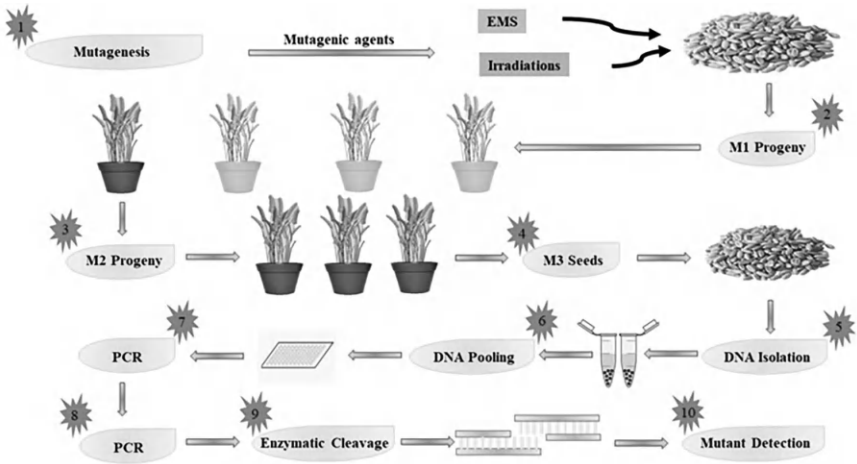


FIGURE 9.5 Firstly, EMS is used for inducing mutations. M1 obtained is self-pollinated to obtain pure mutant lines. The seed from M2 is saved, and DNA is isolated from the mutants leading to DNA pooling, which then amplified and digested enzymatically. After thus, electrophoresis is run for mutation detection. ⏎

mutation mapping can be achieved to identify various mutations. Genotype resequencing can be done quickly once an organism whole genome is sequenced; this provides a comparative study of genomes through studying variations after sequencing. Many unknown mutations can be detected through this method (Kenta et al., 2013).

9.4.8 NEXT-GENERATION SEQUENCING

In research based on genetics, DNA sequencing has a high priority for determining the sequence of discrete genes, whole chromosomes, and small, large, or entire genomic regions of any organism. Primarily, two DNA sequencing methods were used for sequencing of DNA named as Sanger sequencing and Maxam–Gilbert sequencing method, also known as chain termination and chemical degradation methods, respectively (Maxam and Gilbert, 1977; Sanger et al., 1977). Sanger sequencing method gained more popularity among biologists because of its high efficacy and various benefits. In this methodology, 1000bp read length is acquired, possessing 99.99% accuracy. Automated platforms for Sanger sequencing were developed latterly, making it more useful for constructing reference genomes of several plants like Arabidopsis, maize, rice, and soybean (Sahu et al., 2020). Some disadvantages like long processing

time, low output, and high cost limited the use of Sanger sequencing for WGS in various species, particularly having large genome size.

The advent of next-generation sequencing techniques reduces the cost stress with several folds escalation in sequencing outputs. For more than the past 12 years, many commercial NGS platforms (Roche/454, Illumina, Ion torrent, Oxford Nanopore, and Pac Bio) have developed with improvements in NGS methodologies (Maxam and Gilbert, 1977; Thudi et al., 2012). With the help of NGS, multiple samples can be sequenced at low cost and less time, and several gigabases of information can be retrieved per day, having least error rate. NGS methodologies permit resequencing of the whole genome of the species, which were sequenced previously by WGS. This makes mutation studies more convenient and identification of minor variations in sequences like Indels, SNPs by relating reference data (WGS data) and resequenced data (Sahu et al., 2020).

9.4.9 MUTATION SCREENING THROUGH ILLUMINA GENOME ANALYZER

The Illumina Genome Analyzer opens a new door for mutation discoveries and genomic variation studies. This analyzer generates high-quality sequenced DNA data, more than 35 GB within one week using >150 million templates. It offers the ability to find both micro and macro defects in intricate genomes. All genomic variations can be found and verified using a unique strategy that takes use of the enormous data throughput and the easiest methodology of any NGS technology. With a single resolution of nucleotide, SNP or multiple-nucleotide variations, small-large indels, copy number variation, inversions, complex-genomic rearrangements, and translocations can all be computed simultaneously (Lakdawalla and Schroth, 2010; Sturre et al., 2009).

9.4.10 MUTMAP APPROACH

A recently developed forward genetic technique based on NGS called MutMap technology is effective and relatively inexpensive. MutMap technology was first applied on rice plants for finding mutations, which were the reason for leaves' color variation and semi-dwarfness of the plant. A pictorial representation of this approach is demonstrated in Figure 9.6. The study done on rice includes mutagenesis by using chemical mutagens to develop a mutant rice population (Abe et al., 2012). The mutants mostly possess recessive mutations that do not alter the phenotype of the plant. To obtain

homozygosity in the mutant, the generation is subsequently self-pollinated. After this, the homozygous-recessive mutant is pollinated with the wild-type cultivar to obtain phenotypically diverse mutant lines (F_1 generation). For obtaining F_2 generation, F_1 generation plants are self-pollinated.

Moreover, the F_2 population DNA can be sequenced to get to know about indels or SNPs, which results in phenotypic modifications in plants. The MutMap method can be applied to detect genes of interest having importance in agricultural aspects because of its high-throughput application and low-cost sequencing. Recently, some modifications are also made in the MutMap approach in which equal quantity both of F_1 generation plants are self-pollinated, and DNA extracted from them is pooled which is then subjected to Illumina sequencing instead of whole-genome sequencing to find out SNP index separately and then subtract their values to obtain ΔSNP (Abe et al., 2012; Garcia et al., 2016).

9.4.11 MUTCHROMSEQ

Locating a mutation in a whole genome that is randomly induced is very difficult to detect. But finding variations in sequence gets more accessible through the advent of DNA sequencing technology. On the other hand, the variation identification with higher fidelity is only possible for those plant species having small genome sizes like rice or *Arabidopsis.* At the same time*,* it is a very extravagant and challenging approach for large genome size plants like wheat and barley. Recently, a new technique is discovered known as Mutant Chromosome Sequence (MutChromSeq) for swift recognition of random mutations. Now, novel genes in plants have been found using it. In this methodology, chromosome flow sorting is combined with classical mutagenesis to sort out the preferred region on a chromosome where the required gene can be located. The chromosome filtration is done based on marked fluorescent tags. Preferably, those chromosomes are filtered out where mutations can be suspected. Mutant and nonmutant chromosomes are sequenced then for finding genes of interest. The application of this technique can be made on large genome-possessing plants as well as polyploidy genomes (Sánchez-Martín et al., 2016).

A successful implication of MutChromSeq is made on wheat and barley for cloning *Pm2* a powdery mildew-resistant gene and *Eceriferam-q* epicuticular wax accumulation gene, respectively. Mostly, the MutChromSeq method is directed to those plants that gave positive mutagenesis response. The mutation outcomes may be seen phenotypically alongside known

chromosome loci. Therefore, it is an economical and robust methodology of cloning known large genes that are unmanageable (Chaudhary et al., 2019; Sánchez-Martín et al., 2016).

9.4.12 QUANTITATIVE TRAIT LOCI-SEQ

Several genes with negligible effects regulate the majority of agriculturally significant plant features; these are referred to as quantitative trait loci. The recognition of quantitative trait loci (QTLs) is a significant task in plant breeding (Takagi et al., 2013). In the QTL-Seq approach, whole-genome sequencing of DNA is done in order to identify and map the subjected QTLs. For sequencing, DNA from two extremities of a given phenotype in a population is selected. For minor effects QTLs, the inbred recombination lines and double haploids are suitable. The bulk of DNAs are sequenced through WGS with more than six folds of minimum genome coverage. The SNP index is calculated by aligning short reads with reference data of the genome. It is probably supposed that the bulked DNA possesses a majority part of the genome from both parents (1:1). Inequality in parental genomes having QTLs can also be identified. Short reads possessing SNPs are quantified in numbers (k) while relating reference data and sample data. The QTLs are identified in the form of valleys of peaks of the plot of SNP-index. If the complete short reads comprise same genetic fragments as possessed by parent reference data, then the SNP-index is 0. It is 1 if reads are from other parent genome, and it is 0.5 when reads have segments from both parents (Giovannoni et al., 1991; Michelmore et al., 1991).

Thus, the QTL-Seq technique is the quantitative analysis of contributed genes from both parents by using SNP-index. Identification does not necessitate genotyping or marker development. Hence, it is a more rapid and accurate approach for mutation detection. QTLs genes of cucumber, rice, tomato, chickpea, watermelon, peanut, squash, and broccoli have been mapped and identified (Sahu et al., 2020).

9.4.13 EXOME CAPTURE

The traditional mapping approach (map-based or positional cloning) has been replaced gradually with mapping by sequencing. Various methods gave tremendous results in plant models possessing a simple genomic structure like Arabidopsis thaliana and others. But these approaches have limitations in the case of complex genome studies. Studies done on barley

and wheat using these approaches do not give satisfactory outcomes as their genome possesses polyploidy and a massive amount of repetitive DNA. The complexity of these crops' genomes creates a problem in the analysis of data and probing the required gene (Mo et al., 2018).

In contrast to WGS, if only expressed part of the genome is studied, this problem can be solved. The exome is the expressed portion of a genome that includes each exon that encrypts for a protein, small RNAs, as well as some loci with well-known roles. The exome domain is sorted using biotinylated probe hybridization in exome capture technology, which eventually binds to aimed localities. The obtained sample is then amplified and utilized for NGS or any other long-range sequencing approach. This strategy eventually eradicated the genome complexity and made a mapping of specific genes more convenient. For instance, barley and wheat's actual genome size is 5 Gb and 17 Gb, respectively, while the exome size is 62 Mb and 84 Mb, respectively (Warr et al., 2015). Many genes are identified through this approach, like yellow rust-resistant gene, dwarfness gene, barley, wheat, and many other plant species (Krasileva et al., 2017; Mo et al., 2018). This approach has been used widely for forwarding genetic screening and classification of locus of mutant (Allen et al., 2013; Kaur and Gaikwad, 2017). Concurrently, an exome capture approach is helpful in reverse genetic screening. For example, over 10 million point mutations (induced through TILLING) are identified in tetraploidy and hexaploidy wheat having capture probes of 84 Mb. The average mutation rate is approximately 35–40 mutations per kb, equivalent to almost 24 anonymous mutations per interpreted wheat gene (Krasileva et al., 2017).

9.5 CONCLUSION

With the unprecedented variations in climate, there is a threat of a decrease in crop yield. So there is a need for developing agronomically significant plant varieties using advanced molecular breeding techniques, which can help cope with such harsh environmental conditions, i.e., drought stress, temperature variations, and nutritional availability. In this changing global climate, mutation breeding is a critical and highly effective tool for feeding a growing and nutritionally demanding human population. The type of mutation required to achieve the desired trait determines the type of mutagen (physical or chemical). The next step after mutation induction is detecting and screening that mutation, which is either done by phenotypic or genotypic analysis. Phenotypic analysis has now become an old method

of mutation detection. Nowaday's advanced genotypic analysis is done using high-throughput genetic analysis techniques. These techniques are utilized to detect small insertions or deletions, single-nucleotide polymorphisms, chromosomal aberrations, translocations, and transversions. These genetic variations can either be lethal to plant or may increase the survival rate. There is a high chance of getting acquired trait in induced mutations compared to spontaneous mutations. Also, mutation detection techniques are very crucial to identify the multiple mutations simultaneously produced by modern editing tools like CRISRP/Cas systems and help to speed up the breeding process (Razzaq et al., 2019b).

Every mutation detection and screening techniques mentioned above have their advantages and disadvantages. Based on mutation type, genome size, and availability of resources for screening, mutations can be detected. If the genome size is small like *Arabidopsis* and small InDeLs of SNPs have to be detected, a single-strand conformation analysis can be performed. Electrophoretic techniques like temperature-sensitive gel electrophoresis and conformation-sensitive electrophoresis can be done for mutation detection, which is based on melting and conformation of DNA, respectively. The high-throughput melting analysis is an advanced form of mutation detection which comprises melt curve analysis. Whole-genome sequencing and several NGS-based technologies are used for mutation detection in those plants, especially having large genome size (MutMap, MutChromSeq, QTL-Seq approaches). Different mutant-screening techniques can be integrated with advanced tools of metabolomics to fully characterize the type of mutations and assist the breeders to evaluate any harmful phenotype produce as a result of mutation (Razzaq et al., 2019a). Thus, mutation detection in plants is the initial step for finding and securing the improved genetic makeup, eventually leading to mutation breeding and being regarded as a safe genetically modified organism.

KEYWORDS

- **mutation frequency**
- **spontaneous mutations**
- **induced mutations**
- **mutation screening**
- **mutation detection**

REFERENCES

Abaza, G. M. S. M.; Awaad, H. A.; Attia, Z. M.; Abdel-Lateif, K. S.; Gomaa, M. A.; Abaza, S. M. S. M.; Mansour, E. Inducing Potential Mutants in Bread Wheat Using Different Doses of Certain Physical and Chemical Mutagens. *Plant Breed. Biotech.* **2020,** *8*, no. 3, 252–264.

Abe, A.; Kosugi, S.; Yoshida, K.; Natsume, S.; Takagi, H.; Kanzaki, H.; Matsumura, H.; Yoshida, K.; Mitsuoka, C.; Tamiru, M. Genome Sequencing Reveals Agronomically Important Loci in Rice Using MutMap. *Nat. Biotechnol.* **2012,** *30*, no. 2, 174–178.

Ahloowalia, B.; Maluszynski, M. Induced Mutations–A New Paradigm in Plant Breeding. *Euphytica* **2001,** *118*, no. 2, 167–173.

Allen, R. S.; Nakasugi, K.; Doran, R.; Millar, T. A.; Waterhouse, P. Facile Mutant Identification via a Single Parental Backcross Method and Application of Whole Genome Sequencing Based Mapping Pipelines. *Front. Plant Sci.* **2013,** *4*, 362.

Arif, I. A.; Bakir, M. A.; Khan, H. A.; Al Farhan, A. H.; Al Homaidan, A. A.; Bahkali, A. H.; Al Sadoon, M.; Shobrak, M. A Brief Review of Molecular Techniques to Assess Plant Diversity. *Int. J. Mol. Sci.* **2010,** *11*, no. 5, 2079–2096.

Ashton, E. J. Conformation-Sensitive Capillary Electrophoresis, In *PCR Mutation Detection Protocols*; Springer, 2011, pp 1–6.

Ballarini, F.; Alloni, D.; Facoetti, A.; Ottolenghi, A. Heavy-Ion Effects: From Track Structure to DNA and Chromosome Damage. *New J. Phy.* **2008,** *10*, no. 7, 075008.

Beier, D. R. Single-Strand Conformation Polymorphism (SSCP) Analysis as a Tool for Genetic Mapping. *Mamm. Genome* **1993,** *4*, no. 11, 627–631.

Cassells, A.; Doyle, B. Genetic Engineering and Mutation Breeding for Tolerance to Abiotic and Biotic Stresses: Science, Technology and Safety. *Bulg. J. Plant Physiol.* **2003,** 52–82.

Chaudhary, J.; Alisha, A.; Bhatt, V.; Chandanshive, S.; Kumar, N.; Mir, Z.; Kumar, A.; Yadav, S. K.; Shivaraj, S.; Sonah, H. Mutation Breeding in Tomato: Advances, Applicability and Challenges. *Plants* **2019a,** *8*, no. 5, 128.

Chaudhary, J.; Deshmukh, R.; Sonah, H.. Mutagenesis Approaches and their Role in Crop Improvement. *Plants* **2019b,** *8*, 467.

Dooner, H. K.; Wang, Q.; Huang, J. T.; Li, Y.; He, L.; Xiong, W.; Du, C. Spontaneous Mutations in Maize Pollen are Frequent in Some Lines and Arise Mainly from Retrotranspositions and Deletions. *Proc Natl Acad Sci USA* **2019,** *116*, no. 22, 10734–10743.

Dujols, V.; Kusukawa, N.; Mckinney, J. T.; Dobrowolski, S.; Wittwer, C. T. High-Resolution Melting Analysis for Scanning and Genotyping. In *Real-Time PCR*; Garland Science: New York, 2006; pp 157–171.

Gady, A. L.; Hermans, F. W.; Van De Wal, M. H.; Van Loo, E. N.; Visser, R. G.; Bachem, C. W. Implementation of Two High Through-Put Techniques in a Novel Application: Detecting Point Mutations in Large EMS Mutated Plant Populations. *Plant Methods* **2009,** *5*, no. 1, 13.

Garcia, V.; Bres, C.; Just, D.; Fernandez, L.; Tai, F. W. J.; Mauxion, J.-P.; Le Paslier, M.-C.; Bérard, A.; Brunel, D.; Aoki, K. Rapid Identification of Causal Mutations in Tomato EMS Populations via Mapping-by-Sequencing. *Nat. Protoc.* **2016,** *11*, no. 12, 2401.

Garvin, M.; Gharrett, A. DEco-TILLING: An Inexpensive Method for Single Nucleotide Polymorphism Discovery that Reduces Ascertainment Bias. *Mol. Ecol. Notes* **2007,** *7*, no. 5, 735–746.

Giovannoni, J. J.; Wing, R. A.; Ganal, M. W.; Tanksley, S. D. Isolation of Molecular Markers from Specific Chromosomal Intervals Using DNA Pools from Existing Mapping Populations. *Nucleic Acids Res.* **1991,** *19*, no. 23, 6553–6568.

Gottschalk, W.; Wolff, G.. *Induced Mutations in Plant Breeding*; Springer Science and Business Media, 2012.

Govindaraj, M.; Vetriventhan, M.; Srinivasan, M. Importance of Genetic Diversity Assessment in Crop Plants and its Recent Advances: An Overview of its Analytical Perspectives. *Genet. Res. Int.* **2015,** *2015*.

Gull, A.; Lone, A. A.; Wani, N. U. I.. Biotic and Abiotic Stresses in Plants. In *Abiotic and Biotic Stress in Plants*; IntechOpen, 2019.

Hayashi, K. PCR-SSCP: A Method for Detection of Mutations. *Genet. Anal. Biomol. Eng.* **1992,** *9*, no. 3, 73–79.

Herrmann, M. G.; Durtschi, J. D.; Bromley, L. K.; Wittwer, C. T.; Voelkerding, K. V. Amplicon DNA Melting Analysis for Mutation Scanning and Genotyping: Cross-Platform Comparison of Instruments and Dyes. *Clin. Chem.* **2006,** *52*, no. 3, 494–503.

Herrmann, M. G.; Durtschi, J. D.; Bromley, L. K.; Wittwer, C. T.; Voelkerding, K. V. Instrument Comparison For Heterozygote Scanning of Single and Double Heterozygotes: A Correction and Extension of Herrmann et al., clin chem 2006; 52: 494–503. *Clin. Chem.* **2007a,** *53*, no. 1, 150–152.

Herrmann, M. G.; Durtschi, J. D.; Wittwer, C. T.; Voelkerding, K. V. Expanded Instrument Comparison of Amplicon DNA Melting Analysis for Mutation Scanning and Genotyping. *Clin. Chem.* **2007b,** *53*, no. 8, 1544–1548.

Hershberg, R. Mutation—the Engine of Evolution: Studying Mutation and its Role in the Evolution of Bacteria. *Cold Spring Harb Perspect Biol.* **2015,** *7*, no. 9, a018077.

Holme, I. B.; Gregersen, P. L.; Brinch-Pedersen, H. Induced Genetic Variation in Crop Plants by Random or Targeted Mutagenesis: Convergence and Differences. *Front. Plant Sci.* **2019,** *10*, 1468.

Iwahana, H.; Fujimura, M.; Takahashi, Y.; Iwabuchi, T.; Yoshimoto, K.; Itakura, M. Multiple Fluorescence-Based PCR-SSCP Analysis Using Internal Fluorescent Labeling of PCR Products. *Biotechniques* **1996,** *21*, no. 3, 510–519.

Jankowicz-Cieslak, J.; Mba, C.; Till, B. J. Mutagenesis for Crop Breeding and Functional Genomics. In *Biotechnologies for Plant Mutation Breeding*; Springer: Cham, 2017, pp 3–18.

Jo, Y. D.; Kim, J.-B. Frequency and Spectrum of Radiation-Induced Mutations Revealed by Whole-Genome Sequencing Analyses of Plants. *Quantum Beam Sci.* **2019,** *3*, no. 2, 7.

Kaur, P.; Gaikwad, K. From Genomes to GENE-omes: Exome Sequencing Concept and Applications in Crop Improvement. *Front. Plant Sci.* **2017,** *8*, 2164.

Kazama, Y.; Hirano, T.; Saito, H.; Liu, Y.; Ohbu, S.; Hayashi, Y.; Abe, T. Characterization of Highly Efficient Heavy-Ion Mutagenesis in Arabidopsis Thaliana. *BMC Plant Biol.* **2011,** *11*, no. 1, 161.

Kikuchi, S.; Saito, Y.; Ryuto, H.; Fukunishi, N.; Abe, T.; Tanaka, H.; Tsujimoto, H. Effects of Heavy-Ion Beams on Chromosomes of Common Wheat, *Triticum Aestivum*. *Mutat. Res. Fund. Mol. Mech. Mutagen.* **2009,** *669*, nos. 1/2, 63–66.

Körkkö, J.; Annunen, S.; Pihlajamaa, T.; Prockop, D. J.; Ala-Kokko, L. Conformation Sensitive Gel Electrophoresis for Simple and Accurate Detection of Mutations: Comparison with Denaturing Gradient Gel Electrophoresis and Nucleotide Sequencing. *Proc Natl Acad Sci USA* **1998,** *95*, no. 4, 1681–1685.

Krasileva, K. V.; Vasquez-Gross, H. A.; Howell, T.; Bailey, P.; Paraiso, F.; Clissold, L.; Simmonds, J.; Ramirez-Gonzalez, R. H.; Wang, X.; Borrill, P. Uncovering Hidden Variation in Polyploid Wheat. *Proc Natl Acad Sci USA* **2017,** *114*, no. 6, E913–E921.

Kumawat, S.; Rana, N.; Bansal, R.; Vishwakarma, G.; Mehetre, S. T.; Das, B. K.; Kumar, M.; Yadav, S. K.; Sonah, H.; Sharma, T. R. Expanding Avenue of Fast Neutron Mediated Mutagenesis for Crop Improvement. *Plants* **2019,** *8*, no. 6, 164.

Kurowska, M.; Daszkowska-Golec, A.; Gruszka, D.; Marzec, M.; Szurman, M.; Szarejko, I.; Maluszynski, M. TILLING-a Shortcut in Functional Genomics. *J. Appl. Genet.* **2011,** *52*, no. 4, 371.

Lakdawalla, A.; Schroth, G. P. Mutation Discovery with the Illumina® Genome Analyzer. In *The Handbook of Plant Mutation Screening: Mining of Natural and Induced Alleles*; 2010; pp 103–120.

Laskar, R.; Khan, S. Mutagenic effectiveness and efficiency of gamma rays and HZ with phenotyping of induced mutations in lentil cultivars. *Int. Lett. Nat. Sci.* **2017,** *64*, 17–31.

Li, X.; Lassner, M.; Zhang, Y. Deleteagene: A Fast Neutron Deletion Mutagenesis-Based Gene Knockout System for Plants. *Comp. Funct. Genome* **2002,** *3*, no. 2, 158–160.

Markoff, A.; Sormbroen, H.; Bogdanova, N.; Preisler-Adams, S.; Ganev, V.; Dworniczak, B.; Horst, J. Comparison of Conformation-Sensitive Gel Electrophoresis and Single-Strand Conformation Polymorphism Analysis for Detection of Mutations in the BRCA1 Gene Using Optimized Conformation Analysis Protocols. *Eur. J. Hum.* **1998,** *6*, no. 2, 145–150.

Maxam, A. M.; Gilbert, W. A New Method for Sequencing DNA. *Proc Natl Acad Sci USA* **1977,** *74*, no. 2, 560–564.

Mccallum, C. M.; Comai, L.; Greene, E. A.; Henikoff, S. Targeting Induced Locallesions in Genomes (TILLING) for Plant Functional Genomics. *Plant Physiol.* **2000,** *123*, no. 2, 439–442.

Mckinney, J. T.; Nay, L. M.; Koeyer, D. D.; Reed, G. H.; Wall, M.; Palais, R. A.; Jarret, R. L.; Wittwer, C. T. Mutation Scanning and Genotyping in Plants by High-Resolution DNA Melting. In *The Handbook of Plant Mutation Screening: Mining of Natural and Induced Alleles*; 2010; pp 149–165.

Michelmore, R. W.; Paran, I.; Kesseli, R. Identification of Markers Linked to Disease-Resistance Genes by Bulked Segregant Analysis: A Rapid Method to Detect Markers in Specific Genomic Regions by Using Segregating Populations. *Proc Natl Acad Sci USA* **1991,** *88*, no. 21, 9828–9832.

Mo, Y.; Howell, T.; Vasquez-Gross, H.; De Haro, L. A.; Dubcovsky, J.; Pearce, S. Mapping Causal Mutations by Exome Sequencing in a Wheat TILLING Population: A Tall Mutant Case Study. *Mol. Genet. Genome* **2018,** *293*, no. 2, 463–477.

Muleo, R.; Colao, M. C.; Miano, D.; Cirilli, M.; Intrieri, M. C.; Baldoni, L.; Rugini, E. Mutation Scanning and Genotyping by High-Resolution DNA Melting Analysis in Olive Germplasm. *Genome* **2009,** *52*, no. 3, 252–260.

Nieto, C.; Piron, F.; Dalmais, M.; Marco, C. F.; Moriones, E.; Gómez-Guillamón, M. L.; Truniger, V.; Gómez, P.; Garcia-Mas, J.; Aranda, M. A. EcoTILLING for the Identification of Allelic Variants of Melon eIF4E, a Factor that Controls Virus Susceptibility. *BMC Plant Biol.* **2007,** *7*, no. 1, 34.

Okamura, M.; Yasuno, N.; Ohtsuka, M.; Tanaka, A.; Shikazono, N.; Hase, Y. Wide Variety of Flower-Color and-Shape Mutants Regenerated from Leaf Cultures Irradiated with Ion Beams. *Nucl. Instrum. Meth. B.* **2003,** *206*, 574–578.

Oladosu, Y.; Rafii, M. Y.; Abdullah, N.; Hussin, G.; Ramli, A.; Rahim, H. A.; Miah, G.; Usman, M. Principle and Application of Plant Mutagenesis in Crop Improvement: A Review. *Biotechnol. Equip.* **2016,** *30*, no. 1, 1–16.

Pereira, R.; Oliveira, J.; Sousa, M. Bioinformatics and Computational Tools for Next-Generation Sequencing Analysis in Clinical Genetics. *J. Clin. Med.* **2020,** *9*, no. 1, 132.

Peters, H.; Robinson, P. Temperature and Denaturing Gradient Gel Electrophoresis. In *Molecular Diagnostics*; Patrinos, G. P.; Ansorge, W.; Elsevier, 2010.

Prasanna, S.; Jain, S. M. Mutant Resources and Mutagenomics in Crop Plants. *Emir. J. Food Agri.* **2017**, *29*, 651–657.

Predieri, S. Mutation Induction and Tissue Culture in Improving Fruits. *PCTOC* **2001,** *64*, nos. 2/3, 185–210.

Raza, A.; Razzaq, A.; Mehmood, S. S.; Zou, X.; Zhang, X.; Lv, Y.; Xu, J. Impact of Climate Change on Crops Adaptation and Strategies to Tackle Its Outcome: A Review. *Plants* **2019,** *8*, 34.

Razzaq, A.; Sadia, B.; Raza, A.; Khalid Hameed, M.; Saleem, F. Metabolomics: A Way Forward for Crop Improvement. *Metabolites* **2019a,** *9*, 303.

Razzaq, A.; Saleem, F.; Kanwal, M.; Mustafa, G.; Yousaf, S.; Imran Arshad, H. M.; Hameed, M. K.; Khan, M. S.; Joyia, F. A. Modern Trends in Plant Genome Editing: An Inclusive Review of the CRISPR/Cas9 Toolbox. *Int. J. Mol. Sci.* **2019b**, *20*, 4045.

Rout, B. M. TILLING: A Novel Reverse Genetics Technique for Screening of Gene Mutants in Plants. *Biotica Res. Today* **2020,** *2*, no. 5, 252–254.

Sahu, P. K.; Sao, R.; Mondal, S.; Vishwakarma, G.; Gupta, S. K.; Kumar, V.; Singh, S.; Sharma, D.; Das, B. K. Next Generation Sequencing Based Forward Genetic Approaches for Identification and Mapping of Causal Mutations in Crop Plants: A Comprehensive Review. *Plants* **2020,** *9*, no. 10, 1355.

Sánchez-Martín, J.; Steuernagel, B.; Ghosh, S.; Herren, G.; Hurni, S.; Adamski, N.; Vrána, J.; Kubaláková, M.; Krattinger, S. G.; Wicker, T. Rapid Gene Isolation in Barley and Wheat by Mutant Chromosome Sequencing. *Genome Biol.* **2016,** *17*, no. 1, 221.

Sanger, F.; Nicklen, S.; Coulson, A. R. DNA Sequencing with Chain-Terminating Inhibitors. *Proc Natl Acad Sci USA* **1977,** *74*, no. 12, 5463–5467.

Sato, Y.; Nishio, T. Mutation Detection in Rice Waxy Mutants by PCR-RF-SSCP. *Theor. Appl. Genet.* **2003,** *107*, no. 3, 560–567.

Schneeberger, K. Using Next-Generation Sequencing to Isolate Mutant Genes from Forward Genetic Screens. *Nat. Rev. Genet.* **2014,** *15*, no. 10, 662–676.

Shirasawa, K.; Fukuoka, H.; Matsunaga, H.; Kobayashi, Y.; Kobayashi, I.; Hirakawa, H.; Isobe, S.; Tabata, S. Genome-Wide Association Studies Using Single Nucleotide Polymorphism Markers Developed by Re-Sequencing of the Genomes of Cultivated Tomato. *DNA Res.* **2013,** *20*, no. 6, 593–603.

Shirasawa, K.; Shiokai, S.; Yamaguchi, M.; Kishitani, S.; Nishio, T. Dot-blot-SNP Analysis for Practical Plant Breeding and Cultivar Identification in Rice. *Theor. Appl. Genet.* **2006,** *113*, no. 1, 147–155.

Sigurbjörnsson, B. Induced Mutations. In *Crop Breeding*; 1983; pp 153–176.

Stefano, P.; Edoardo, G. In vitro Techniques and Physical Mutagens for the Improvement of Fruit Crops. In *In Vitro Application in Crop Improvement*; CRC Press, 2004; pp 37–52.

Sturre, M. J.; Shirzadian-Khorramabad, R.; Schippers, J. H.; Chin-a-Woeng, T. F.; Hille, J.; Dijkwel, P. P. Method for the Identification of Single Mutations in Large Genomic Regions Using Massive Parallel Sequencing. *Mol. Breed.* **2009,** *23*, no. 1, 51–59.

Tabone, T.; Sallmann, G.; Cotton, R. G. Chemical Methods for Mutation Detection: The Chemical Cleavage of Mismatch Method. In *The Handbook of Plant Mutation Screening: Mining of Natural and Induced Alleles*; 2010; pp 121–130.

Takagi, H.; Abe, A.; Yoshida, K.; Kosugi, S.; Natsume, S.; Mitsuoka, C.; Uemura, A.; Utsushi, H.; Tamiru, M.; Takuno, S. QTL-seq: Rapid Mapping of Quantitative Trait Loci in Rice by Whole Genome Resequencing of DNA from Two Bulked Populations. *Plant J.* **2013,** *74*, no. 1, 174–183.

Tanaka, A.; Shikazono, N.; Hase, Y. Studies on Biological Effects of Ion Beams on Lethality, Molecular Nature of Mutation, Mutation Rate, and Spectrum of Mutation Phenotype for Mutation Breeding in Higher Plants. *J. Radiat. Res.* **2010,** *53*, no. 3, 223–233.

Thudi, M.; Li, Y.; Jackson, S. A.; May, G. D.; Varshney, R. K. Current State-of-Art of Sequencing Technologies for Plant Genomics Research. *Brief. Funct. Genome* **2012,** *11*, no. 1, 3–11.

Till, B. J. Mutation Detection in Plants by Enzymatic Mismatch Cleavage. In *The Handbook of Plant Mutation Screening: Mining of Natural and Induced Alleles*; 2010; pp 131–148.

Till, B. J.; Cooper, J.; Tai, T. H.; Colowit, P.; Greene, E. A.; Henikoff, S.; Comai, L. Discovery of Chemically Induced Mutations in Rice by TILLING. *BMC Plant Biol.* **2007,** *7*, no. 1, 19.

Uauy, C.; Paraiso, F.; Colasuonno, P.; Tran, R. K.; Tsai, H.; Berardi, S.; Comai, L.; Dubcovsky, J. A Modified TILLING Approach to Detect Induced Mutations in Tetraploid and Hexaploid Wheat. *BMC Plant Biol.* **2009,** *9*, no. 1, 115.

Warr, A.; Robert, C.; Hume, D.; Archibald, A.; Deeb, N.; Watson, M. Exome Sequencing: Current and Future Perspectives. *Genes Genomes Genet.* **2015,** *5*, no. 8, 1543–1550.

Wartell, R.; Hosseini, S. H.; Mohan Jr, C. P. Detecting Base Pair Substitutions in DNA Fragments by Temperature-Gradient Gel Electrophoresis. *Nucleic Acids Res.* **1990,** *18*, no. 9, 2699–2705.

Widyawan, M.; Khumaida, N.; Kitashiba, H.; Nishio, T. Optimization of Dot-Blot SNP Analysis for the Detection of Drought or Salinity Stress Associated Marker in Foxtail Millet (*Setaria italica* L.). *SABRAO J. Breed. Genet.* **2019,** *50*, no. 1, 72–84.

Yamaguchi, H.; Hase, Y.; Tanaka, A.; Shikazono, N.; Degi, K.; Shimizu, A.; Morishita, T. Mutagenic Effects of Ion Beam Irradiation on Rice. *Breed. Sci.* **2009,** *59*, no. 2, 169–177.

Zheng, X.; Yang, S.; Zhang, D.; Zhong, Z.; Tang, X.; Deng, K.; Zhou, J.; Qi, Y.; Zhang, Y. Effective Screen of CRISPR/Cas9-Induced Mutants in Rice by Single-Strand Conformation Polymorphism. *Plant Cell Rep.* **2016,** *35*, no. 7, 1545–1554.



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

CHAPTER 10

Mutations and Mechanisms

MUHAMMAD TARIQ ZAHID[1], HABIBA RAMZAN[1], GHULAM MUSTAFA[1], FARAH RAUF SHAKOORI[2], and TARIQ AHMAD BHAT[3]

[1]*Department of Zoology, Government College University Lahore, Pakistan*

[2]*Institute of Zoology, University of the Punjab, Lahore, Pakistan*

[3]*Department of Education, Government of Jammu and Kashmir, Jammu and Kashmir, India*

ABSTRACT

Mutations are the alterations in the nucleotide sequence of DNA in somatic and germ cells arising spontaneously or induced by external factors, leading to nonheritable and heritable consequences, such as various types of cancers, color blindness, hemophilia, diabetes, and sickle cell anemia. Mutations occur due to errors in DNA replication, where the replication machinery misreads the template, often under the influence of different environmental factors called mutagens. These mutagens, including radiation and various chemicals, either modify the nucleotides or mask them to be misread. Mutations typically arise in one of the two strands, synthesizing 50% mutated DNA. The mutation process may involve **base substitutions**, **deletions**, and **insertions.** Base substitutions involve the replacement of one nucleotide with another, potentially leading to protein malfunctions, while deletions and insertions cause shifts in the reading frame, resulting in truncated or nonfunctional proteins. Based on their phenotypic effects, mutations are categorized as silent, missense, and nonsense mutations. Silent mutations do not affect protein sequences but may influence protein function or gene regulation, while missense and nonsense mutations can significantly alter protein structure or lead to premature termination of protein synthesis. This

Molecular Genetics, Structures, Mechanisms, and Functions, Volume 1: Principles of Gene Manipulation and Genomics. Tariq Ahmad Bhat, Abdul Rauf Shakoori, and Jameel M. Al-Khayri (Eds.)
© 2026 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

chapter provides a detailed discussion of different mutagens and their role in mutation induction.

10.1 INTRODUCTION

In 1953, the DNA replication model by James Watson and Francis Crick proposed that daughter DNA molecules are exact copies of parental DNA. This concept still prevails except for replication errors. However, the DNA of a cell may also undergo some changes due to environmental factors that may transfer to daughter cells after cell division. These changes are called mutations. The mutation is any change in the nucleotide sequence of genetic material. This change may be of a single nucleotide or a fragment of nucleotides. The term mutation was first coined by Hugo de Vries in 1886 to describe the inheritable changes in the genetic material of an organism. Although mutations are notoriously famous for their deleterious effects, they are beneficial as well; others are neutral. In fact, majority of the mutations are silent with no phenotypic effects. Mutations provide the raw material for genetic variability and evolutionary processes. The diversity and versatility of living organisms are mainly because of mutations. It is also a fact that the majority of genetic diseases are associated with gene mutations (Zlotogora, 2010; Mezzi et al., 2021).

10.2 METHODS OF GENE MUTATIONS

DNA is stable enough to work as heredity material, yet it undergoes some variations due to process of mutagenesis. Mutations may arise in the germ cells (gametes) or somatic cells. Germ line mutations are transferred to offspring and are a major cause of genetic diseases. Although somatic mutations are not inheritable, they may result in different types of cancers in that individual. Mutations occur due to error in DNA replication, mutagen exposure, and viral infection. There are three methods by which gene mutations are incorporated:

1. Base substitutions/base analogs
2. Deletions
3. Insertions

10.2.1 BASE SUBSTITUTIONS

Base substitution is a mutagenesis process in which a nucleotide (e.g., Adenosine) in DNA molecule is replaced with another nucleotide. The base substitutions are categorized as transitions or transversions. The phenomenon of a purine – purine (A↔G) or a pyrimidine – pyrimidine (C↔T) replacement is called **transition**. But a purine – pyrimidine (R↔Y) replacement is called **transversion**. All the base substitutions result in point mutations. A famous example of base substitution is M1V mutation where AUG encoding methionine is mutated to valine encoding GUG codon. This results in an abnormal protein expression. M1V mutation has been reported in several diseases like phenylketonuria (Lyonnet et al., 1992), adrenocorticular disease (Pereira et al., 2010), and deafness (Thönnissen et al., 2002). Base substitutions may be induced by viruses, toxins, chemicals, or radiations.

10.2.2 DELETIONS

Deletion is the loss of a part of DNA molecule which may be of single nucleotide or a group of nucleotides. Normally, a deletion mutation in a DNA molecule disrupts the complete sequence of downstream codons (frame shift mutation), resulting in nonfunctional or truncated proteins. Deletion mutations have been reported in many types of cancers and other diseases for example prostate cancer (Chen et al., 2003), primary bladder cancer (Cairns et al., 1998), lung cancer (Yu et al., 2020), and glycogen storage disease (Choi et al., 2016). Cystic fibrosis is another example where phenylalanine at 508 position of CFTR protein is lost due to deletion mutation (Bartoszewski et al., 2010). This disrupts the chloride ion channels, preventing them from transportation of the chloride ions across the cell membrane. This results in thick and sticky mucous that blocks the tubular structures of the body.

10.2.3 INSERTIONS

When one or more nucleotides are added to a DNA molecule, the phenomenon is called insertion. Insertion mutations may be due to transposition, crossing over, or DNA replication errors. Insertion mutations have been reported in lung cancer (Sasaki et al., 2007), Cohn's diseases (Hampe et al., 2001), and Huntington disease (HD). In HD genes, less than 27 CAG repeats are normal. However, the presence of 40 and more repeats causes

Huntington disease (Myers, 2004). An insertion of one or more complete codons will not change the nature of downstream amino acids.

10.3 SOURCES OF MUTATIONS

10.3.1 SPONTANEOUS AND INDUCED MUTATIONS

It is believed that mutations arise continuously, with or without the presence of any mutagen. Mutations occurring naturally without the involvement of any environmental factor or mutagen are called spontaneous or random mutations. Spontaneous mutations may arise due to enzymatic errors during DNA replication, transcription, recombination, and DNA repair processes (Wright, 2000). During the DNA replication process, incorporation of a wrong nucleotide and error in proof reading activity of DNA polymerase introduce new mutations. Similarly, during the transcription process, the strand of melted DNA may undergo deamination, which results in C to T mutation (Beletskii and Bhagwat, 1996). Spontaneous mutations affect the genome at a nearly constant rate. All other mutations are induced mutations as they are caused by some mutation-causing agents called mutagens. Mutagens are classified into three major categories, namely physical, chemical, and biological mutagens.

10.3.2 PHYSICAL MUTAGENS

Ultraviolet radiation, X-rays, and γ-rays are among the best-known physical mutagens. Physical mutagens cause mutations by breaking the hydrogen and covalent bonds in DNA molecule. UV radiations particularly UVB ($\lambda_{290–320\ nm}$) present in sunlight directly damage DNA molecule. The pyrimidine bases (Thymine, cytosine, and 5-methyl cytosine) have greater absorption of UVB from sunlight while adenine and guanine are least affected (Ravanat et al., 2001). The UV radiations break the hydrogen bonds between purine (A/G) and pyrimidine (C/T) in complementary strand, resulting in the formation of hydrogen bonds with the adjacent pyrimidine nucleotide. This results in the formation of cyclobutene pyrimidine dimers, pyrimidine–pyrimidone photoproduct, and Dewar valence isomers (Sinha and Häder, 2002; Douki and Sage, 2016; Rastogi et al., 2010). These dimers are more frequently formed when there is a T-T or T-C nucleotide sequence (Courdavault et al., 2005). If the dimers remain intact, they halt DNA replication machinery, or a

random nucleotide is added at that position. Deamination and further repair mechanism may convert a cytidine nucleotide into thymidine. The mutations due to UV radiations result in various types of skin cancers (Patra et al., 2019; Narayanan et al., 2010; Pfeifer et al., 2005).

X-rays and γ-rays work as ionizing radiation. These radiations ionize molecules especially water and produce free radicals (having unpaired electron), which may modify the bases or make lesion due to unrepairable damage. They form free hydroxyl radical inducing single-strand breaks (SSB). There may be multiple lesions of ionization within short area of DNA molecule causing double-strand breaks (DSB) in the DNA molecule (Sage and Shikazono, 2017). Due to this property, they are called clastogens. Gamma radiations are even more powerful ionizing agents causing clustered damage (Sutherland et al., 2002).

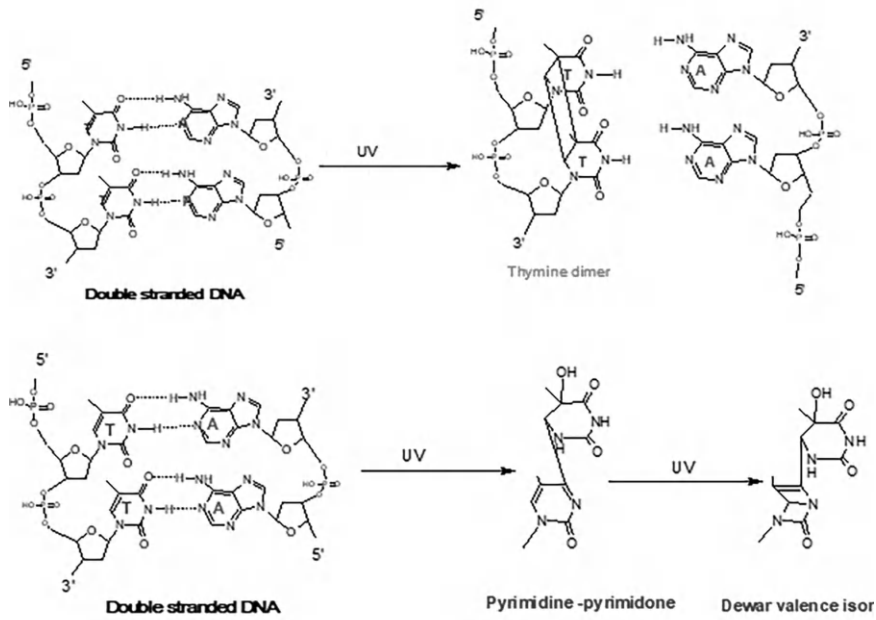


FIGURE 10.1 Mutagenic effect of ultraviolet radiation on the double-stranded DNA molecule.

10.3.3 CHEMICAL MUTAGENS

Different chemicals interfere with DNA, produce mutations, and result in sever abnormalities. Important chemical mutagens include nitrous acid,

ethyl methane sulfonate, 5-bromouracil, and hydroxylamine. They may be of natural or synthetic origin. Chemical mutagens can be divided into three main categories:

10.3.3.1 BASE ANALOGUES

Base analogues are the chemical molecules showing resemblance with nitrogen bases. They substitute the original nitrogen bases and alter DNA structure and gene expression. At the time of DNA replication, they often incorporate a wrong nucleotide resulting in point mutation (Mathews, 2006). They have been very useful to understand the mutagenesis process (Negishi et al., 2002). A base analogue 5-bromouracil 5BrU in normal keto form resembles with thymine, but its less stable enolic form resembles with cytosine. When a thymine is substituted with 5-BrU in a DNA molecule, that pairs with adenine like that of thymine; however at the time of replication, tautomeric shift (keto to enol form) works and it incorporated guanine in daughter strand instead of adenine (Hu et al., 2004).

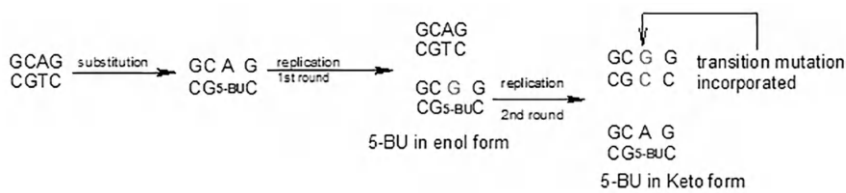


FIGURE 10.2 Mechanism of 5-bromouracil-induced substitutions.

Similarly, 2-amino purine is an analogue of adenine. In its normal amino form, it will pair with thymine like that of adenine; however, due to tautomeric shift it appears in imino form and behaves like guanine resulting in pairing with cytosine at the time of replication (Pitsikas et al., 2004).

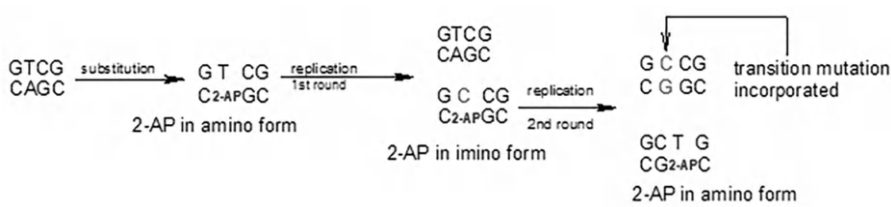


FIGURE 11.3 Mechanism of 2-amino purine-induced substitutions.

10.3.3.2 BASE MODIFYING AGENTS

Certain chemicals such as nitrous acid (HNO_2), hydroxylamine (NH_2OH), and ethyl methanesulfonate are not incorporated into DNA, but they modify the nitrogen bases. HNO_2 is a powerful deaminating agent that removes amino group of cytosine, converting it to uracil. At the time of DNA replication, this uracil will pair with adenine resulting in GC to AT transition. Similarly, deamination reaction of HNO_2 convert adenine to hypoxanthine results in the TA to CG transition. These mutations have been reported in a large number of cancer patients (Parry, 1997; Yonekura et al., 2009).

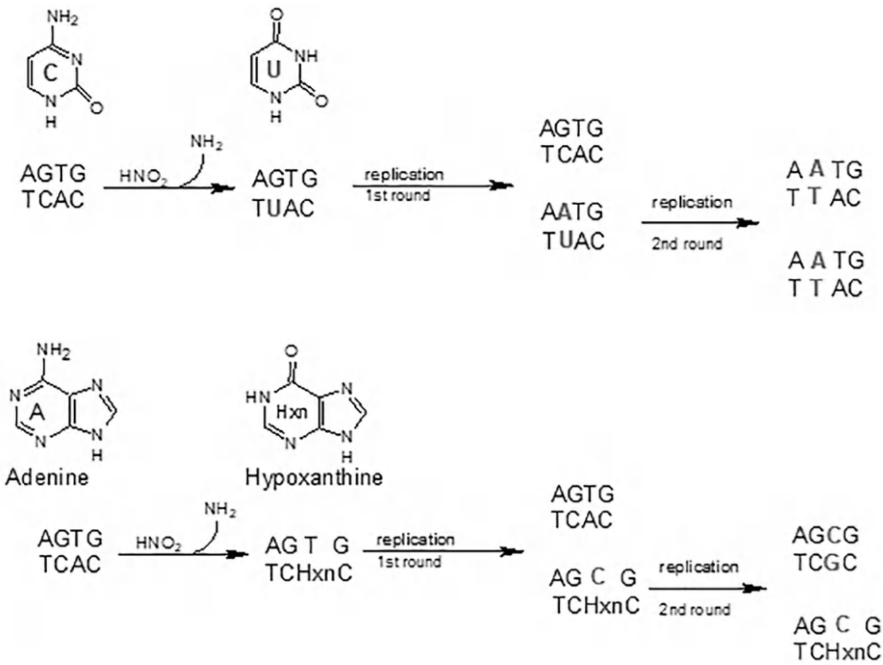


FIGURE 10.4 Substitution mutations under the influence of HNO_2.

Hydroxylamine adds OH group to cytosine converting it into hydroxycytosine which results in AT to GC transition mutation. Similarly, alkylation of guanine to O-6-ethylguanine by ethylmethane sulfonate results in GC to AT transition mutation (Maloy, 2007).

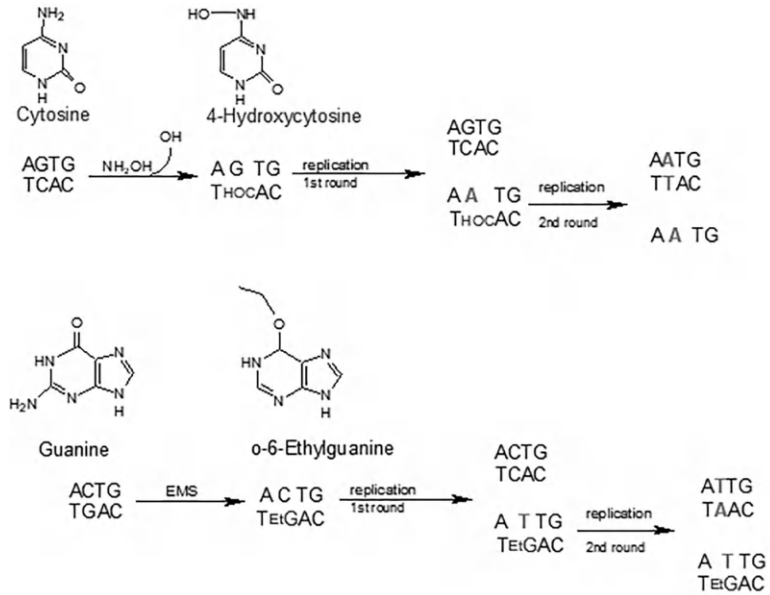


FIGURE 10.5 Mechanism of substitution mutations by hydroxylamine and ethylmethane sulfonate.

10.3.3.3 INTERCALATING AGENTS

Certain chemicals like acridine orange, ethidium bromide, and proflavine insert themselves within the double helix of DNA. This increases the space between nitrogen bases which results in insertions, deletions, and frame shift mutations (Kapuscinski and Darzynkiewicz, 1984; Arshad et al., 2006).

10.3.4 TYPES OF GENE MUTATIONS

Based on its phenotypic effects, there may be silent, frameshift, nonsense, or missense mutations.

10.3.4.1 SILENT OR SYNONYMOUS MUTATIONS

These are the mutations in the coding region of a gene which do not affect the amino acids sequence of the protein molecule. This is because majority of the amino acids have more than one codon due to the degeneracy of genetic code. For example, in the given nucleotide sequence AAG encodes lysine. If in the fourth codon, G is mutated to A, the resulting codon (AAA) will also encode same amino acid (lysine).



FIGURE 10.6 Effect of silent mutations; a substitution in the nucleotide sequence results in a different codon encoding for the same amino acid.

Although silent mutations do not change the amino acid sequence in polypeptide chain, they affect the mRNA splicing and protein function. The silent mutations have been reported to cause different types of genetic disorders such as β-thalassemia, Phenylketonuria, and Colorectal cancer. In familial adenomatous polyposis cancer, a silent mutation in APC gene results in truncated protein leading to disease (Aretz et al., 2004). Silent mutations also affect the folding rate and insertion of P-glycoproteins into cell membrane, resulting in altered substrate specificity (Kimchi-Sarfaty et al., 2007).

10.3.4.2 FRAMESHIFT MUTATIONS

An addition or deletion mutation in the DNA that changes the reading sequence of nucleotides, resulting in the formation of a polypeptide with different amino acids, is called frameshift mutation. This is because when a nucleotide is added or deleted, it changes the nucleotide sequence of the mutated and all the downstream codons. The resulting polypeptide will have entirely different amino acid sequence. Frameshift mutations have been reported in different types of cancers, cystic fibrosis, and Tay-Sachs disease and several other illnesses.

TABLE 10.1 Effect of Frameshift Mutation on the Amino Acids Sequence of a Protein Molecule.

Normal sequence	DNA:	atg agc aaa **gta** tta gtt ctg aaa tcc agt att ctg gca ggg tac tca caa tct ggt cag
	Protein:	M S K V L V L K S S I L A G Y S Q S G Q
Frame-shift mutation	DNA:	atg agc aaa **tat** tag ttc tga aat cca gta ttc tgg cag ggt act cac aat ctg gtc ag
	Protein:	M S K Y - F - N P V F W Q G T H N L V

10.3.4.3 MISSENSE MUTATIONS

The mutations in DNA molecule which result in a different amino acid at a specific position of the polypeptide chain are called missense

mutations. This substitution of amino acids may alter the stability and flexibility of proteins. If a missense mutation occurs in the binding area of protein, it will affect the affinity/binding of protein with others. These mutations may change the subcellular location of a protein. For example, a missense mutation in FANCA protein (DNA repair protein) results in its unavailability into the nucleus to repair DNA double-strand damage (Castella et al., 2011). Missense mutations also cause an altered gene expression.

TABLE 10.2 Missense Mutations Change a Single Amino Acid, Leading to a Change in the Amino Acid Properties of Protein Molecules.

Normal sequence	DNA:	atg agc aaa **gta** tta gtt ctg aaa tcc agt att ctg gca ggg tac tca caa tct ggt cag
	Protein:	M S K V L V L K S S I L A G Y S Q S G Q
Missense mutation	DNA:	atg agc aaa **cta** tta gtt ctg aaa tcc agt att ctg gca ggg tac tca caa tct ggt cag
	Protein:	M S K L L V L K S S I L A G Y S Q S G Q

10.3.4.4 NONSENSE MUTATIONS

It is a mutation when an amino acid-specific codon is converted to a stop codon. This premature stop codon results in a nonfunctional, truncated protein. Approximately, 12% of the patients with genetic diseases have nonsense mutations (Krawczak et al., 2000).

TABLE 10.3 Nonsense Mutations Change an Amino Acid Encoding Codon (Sense Codon) into a Stop Codon (Nonsense), Resulting in Truncated Protein Molecules.

Normal sequence	DNA:	atg agc aaa gta tta gtt ctg **aaa** tcc agt att ctg gca ggg tac tca caa tct ggt cag
	Protein:	M S K V L V L **K** S S I L A G Y S Q S G Q
Frame-shift mutation	DNA:	atg agc aaa gta tta gtt ctg **taa** tcc agt att ctg gca ggg tac tca caa tct ggt cag
	Protein:	M S K V L V L - S S I L A G Y S Q S G Q

10.4 EFFECTS OF MUTATIONS

Mutations affect the survival and fitness of an organism. Sometimes, mutations make the respective gene or protein completely nonfunctional, and organism may not be viable. Such mutations are called lethal mutations as they cause the death of an organism. For example, the mutations in RB genes make the developing embryo nonviable. However, in some

cases the gene or its protein is partially active, and the mutation is called sublethal. For example, a mutation in 3' end of *BRCA2* gene makes the mice viable, although life expectancy is highly reduced (Houghtaling et al., 2003). Many mutations in genetic material do not affect the survival of organisms; such mutations are termed vital mutations. Sometimes, the mutations can increase the fitness of an organism, such mutations are termed adaptive mutations. The concept of adaptive mutations focuses that some nonlethal mutations relieve the selective pressure on organisms and make them best fit in the environment (Foster, 2000). For example, mutations in spike protein of Omicron virus stabilize spike-up RBD conformation and increase its fitness (Zhao et al., 2022). The process of evolution is also due to continuous generation of adaptive mutations (Bataillon, 2000; Perfeito et al., 2007). However, in some cases, they may reduce the fitness of an organism. For example, vital mutations in human *EGFR*, *ALK*, *kras* and many other genes have been reported to play a crucial role in the progression of diseases like non-small cell lung cancer (Cardarella et al., 2013).

10.4.1 SINGLE-NUCLEOTIDE POLYMORPHISM

It is the variation of a single nucleotide in DNA sequence. SNPs are present in a considerable portion of population, normally more than 1%. They may be present in the coding or noncoding region of a gene. Normally, SNPs do not cause any disorder; however, some are associated with certain diseases such as Alzheimer's disease, diabetes, and cancer (Cruchaga et al., 2010; Dayeh et al., 2013; Allemailem et al., 2021).

10.4.2 MUTATION FREQUENCY

Mutation frequency is the proportion of mutations in a population. Frequency of different genetic mutations in a population is not the same. The mutation frequency largely depends upon the sequence of nucleotides in DNA of the mutated regions. The nucleotides with high mutation frequency are called mutation hotspots. Cancer genes are the major candidates of hotspot mutations (Tuupanen et al., 2014). Mutation rates are very important with an evolutionary view point (Berdan et al., 2021).

KEYWORDS

- **mutations**
- **mutagens**
- **nucleotide modification**
- **DNA replication**
- **proteins malfunctioning**

REFERENCES

Allemailem, K. S.; et al. Single Nucleotide Polymorphisms (SNPs) in Prostate Cancer: Its Implications in Diagnostics and Therapeutics. *Am. J. Transl. Res.* **2021,** *13* (4), 3868.

Aretz, S.; et al. Frequency and Parental Origin of De Novo APC Mutations in Familial Adenomatous Polyposis. *Eur. J. Hum. Genet.* **2004,** *12* (1), 52–58.

Arshad, R.; et al. Mutagenic Effect of Acridine Orange on the Expression of Penicillin G Acylase and β-lactamase in Escherichia coli. *Lett. Appl. Microbiol.* **2006,** *42* (2), 94–101.

Bartoszewski, R. A.; et al. A Synonymous Single Nucleotide Polymorphism in ΔF508 CFTR Alters the Secondary Structure of the mRNA and the Expression of the Mutant Protein. *J. Biol. Chem.* **2010,** *285* (37), 28741–28748.

Bataillon, T. Estimation of Spontaneous Genome-Wide Mutation Rate Parameters: Whither Beneficial Mutations? *Heredity* **2000,** *84* (5), 497–501.

Beletskii, A.; Bhagwat, A. S. Transcription-Induced Mutations: Increase in C to T Mutations in the Nontranscribed Strand during Transcription in Escherichia coli. *Proc. Natl.. Acad. Sci.* **1996,** *93* (24), 13919–13924.

Berdan, E. L.; et al. Unboxing Mutations: Connecting Mutation Types with Evolutionary Consequences. *Mol. Ecol.* **2021,** *30* (12), 2710–2723.

Cairns, P.; et al. Point Mutation and Homozygous Deletion of PTEN/MMAC1 in Primary Bladder Cancers. *Oncogene* **1998,** *16* (24), 3215–3218.

Cardarella, S.; et al. Clinical, Pathologic, and Biologic Features Associated with BRAF Mutations in Non–Small Cell Lung CancerClinical and Biologic Characteristics of BRAF-Mutant NSCLC. *Clin. Cancer Res.* **2013,** *19* (16), 4532–4540.

Castella, M.; et al. Origin, Functional Role, and Clinical Impact of Fanconi Anemia FANCA Mutations. *Blood* **2011,** *117* (14), 3759–3769.

Chen, C.; et al. Deletion, Mutation, and Loss of Expression of KLF6 in Human Prostate Cancer. *Am. J. Pathol.* **2003,** *162* (4), 1349–1354.

Choi, R.; et al. PHKA2 Mutation Spectrum in Korean Patients with Glycogen Storage Disease Type IX: Prevalence of Deletion Mutations. *BMC Med. Genet.* **2016,** *17*, 1–9.

Courdavault, S.; et al. Repair of the Three Main Types of Bipyrimidine DNA Photoproducts in Human Keratinocytes Exposed to UVB and UVA Radiations. *DNA Rep.* **2005,** *4* (7), 836–844.

Cruchaga, C.; et al. SNPs Associated with Cerebrospinal Fluid Phospho-Tau Levels Influence Rate of Decline in Alzheimer's Disease. *PLoS Genet.* **2010,** *6* (9), e1001101.

Dayeh, T. A.; et al. Identification of CpG-SNPs Associated with Type 2 Diabetes and Differential DNA Methylation in Human Pancreatic Islets. *Diabetologia* **2013,** *56*, 1036–1046.

Douki, T.;Sage, E. Dewar Valence Isomers, the Third Type of Environmentally Relevant DNA Photoproducts Induced by Solar Radiation. *Photochem. Photobiol. Sci.* **2016,** *15* (1), 24–30.

Foster, P. L. Adaptive Mutation: Implications for Evolution. *Bioessays* **2000,** *22* (12), 1067–1074.

Hampe, J.; et al. Association between Insertion Mutation in NOD2 Gene and Crohn's Disease in German and British Populations. *Lancet* **2001,** *357* (9272), 1925–1928.

Houghtaling, S.; et al. Epithelial Cancer in Fanconi Anemia Complementation Group D2 (Fancd2) Knockout Mice. *Genes Dev.* **2003,** *17* (16), 2021–2035.

Hu, X.; et al. Theoretical Study of the Proton Transfer of Uracil and (Water) n (n= 0− 4): Water Stabilization and Mutagenicity for Uracil. *J. Phys. Chem. B* **2004,** *108* (34), 12999–13007.

Kapuscinski, J.; Darzynkiewicz, Z. Condensation of Nucleic Acids by Intercalating Aromatic Cations. *Proc. Natl. Acad. Sci.* **1984,** *81* (23), 7368–7372.

Kimchi-Sarfaty, C.; et al. A "Silent" Polymorphism in the MDR 1 Gene Changes Substrate Specificity. *Science* **2007,** *315* (5811), 525–528.

Krawczak, M.; et al. Human Gene Mutation Database—A Biomedical Information and Research Resource. *Hum. Mutat.* **2000,** *15* (1), 45–51.

Lyonnet, S.; et al. Time and Space Clusters of the French-Canadian MIV Phenylketonuria Mutation in France. *Am. J. Hum. Genet.* **1992,** *51* (1), 191.

Maloy, S. R. Localized Mutagenesis. *Methods Enzymol.* **2007,** *421,* 42–50.

Mathews, C. K. DNA Precursor Metabolism and Genomic Stability. *FASEB J.* **2006,** *20* (9), 1300–1314.

Mezzi, N.; et al. Spectrum of Genetic Diseases in Tunisia: Current Situation and Main Milestones Achieved. *Genes* **2021,** *12* (11), 1820.

Myers, R. H. Huntington's Disease Genetics. *NeuroRx* **2004,** *1*, 255–262.

Narayanan, D. L.; Saladi, R. N.; Fox, J. L. Ultraviolet Radiation and Skin Cancer. *Int. J. Dermatol.* **2010,** *49* (9), 978–986.

Negishi, K.; Loakes, D.;Schaaper, R. M. Saturation of DNA Mismatch Repair and Error Catastrophe by a Base Analogue in Escherichia coli. *Genetics* **2002,** *161* (4), 1363–1371.

Parry, T. The Non-Random Distribution of Point Mutations in Leukaemia and Myelodysplasia—A Possible Pointer to their Aetiology. *Leuk. Res.* **1997,** *21* (6), 559–574.

Patra, V.; et al. Skin Microbiome Modulates the Effect of Ultraviolet Radiation on Cellular Response and Immune Function. *Iscience* **2019,** *15*, 211–222.

Pereira, A. M.; et al. Association of the M1V PRKAR1A Mutation with Primary Pigmented Nodular Adrenocortical Disease in Two Large Families. *J. Clin. Endocrinol. Metab.* **2010,** *95* (1), 338–342.

Perfeito, L.; et al. Adaptive Mutations in Bacteria: High Rate and Small Effects. *Science* **2007,** *317* (5839), 813–815.

Pfeifer, G. P.; You, Y. H.; Besaratinia, A. Mutations Induced by Ultraviolet Light. *Mutat. Res.* **2005,** *571* (1-2), 19–31.

Pitsikas, P.; Patapas, J. M.; Cupples, C. G. Mechanism of 2-Aminopurine-Stimulated Mutagenesis in Escherichia coli. *Mutat. Res.* **2004,** *550* (1-2), 25–32.

Rastogi, R. P.; et al. Molecular Mechanisms of Ultraviolet Radiation-Induced DNA Damage and Repair. *J. Nucleic Acids* **2010,** *2010*.

Ravanat, J. L.; Douki, T.; Cadet, J. Direct and Indirect Effects of UV Radiation on DNA and its Components. *J. Photochem. Photobiol. B Biol.* **2001,** *63* (1-3), 88–102.

Sage, E.; Shikazono, N. Radiation-Induced Clustered DNA Lesions: Repair and Mutagenesis. *Free Radic. Biol. Med.* **2017,** 107, 125–135.

Sasaki, H.; et al. EGFR Exon 20 Insertion Mutation in Japanese Lung Cancer. *Lung Cancer* **2007,** *58* (3), 324–328.

Sinha, R. P.; Häder, D. P. UV-Induced DNA Damage and Repair: A Review. *Photochem. Photobiol. Sci.* **2002,** *1* (4), 225–236.

Sutherland, B. M.; et al. Clustered DNA Damages Induced by X Rays in Human Cells. *Radiat. Res.* **2002,** *157* (6), 611–616.

Thönnissen, E.; et al. Human Connexin26 (GJB2) Deafness Mutations Affect the Function of Gap Junction Channels at Different Levels of Protein Expression. *Hum. Genet.* **2002,** *111*, 190–197.

Tuupanen, S.; et al. Identification of 33 Candidate Oncogenes by Screening for Base-Specific Mutations. *Brit. J. Cancer* **2014,** *111* (8), 1657–1662.

Wright, B. E. A Biochemical Mechanism for Nonrandom Mutations and Evolution. *J. Bacteriol.* **2000,** *182* (11), 2993–3001.

Yonekura, S. I.; et al. Generation, Biological Consequences and Repair Mechanisms of Cytosine Deamination in DNA. *J. Radiat. Res.* **2009,** 50 (1), 19–26.

Yu, S.; et al. EGFR E746-A750 Deletion in Lung Cancer Represses Antitumor Immunity through the Exosome-Mediated Inhibition of Dendritic Cells. *Oncogene* **2020,** *39* (13), 2643–2657.

Zhao, Z.; et al. Omicron SARS-CoV-2 Mutations Stabilize Spike Up-RBD Conformation and Lead to a Non-RBM-Binding Monoclonal Antibody Escape. *Nat. Commun.* **2022,** *13* (1), 4958.

Zlotogora, J. The Molecular Basis of Autosomal Recessive Diseases Among the Arabs and Druze in Israel. *Hum. Genet.* **2010,** *128*, 473–479.

CHAPTER 11

Population Genetics

SHAGUFTA PERVEEN, FAIZA AKHTAR, NADIA YASMEEN, KHADIJA HAFEEZ, ABIDA PARVEEN, and SARA ZAFAR

Government College University, Faisalabad, Pakistan

ABSTRACT

Modern theories of inter-organism variation integrate genetic theory established on Mendel's ideas with theory of evolution based on Darwin and Wallace's work. Variation, inheritance, and evolutionary theory are all part of the biological field known as population genetics. It deals with the prevalence of various alleles within populations and how the distribution of a given allele changes over time. It also includes the gene pool (the total number of genes and alleles present in the population), allele frequency (the number of organisms in the population bearing a particular gene), genotype frequency, and the Hardy–Weinberg equation and its implications. The Hardy–Weinberg equation states that, under certain circumstances, dominant and recessive alleles frequency within a population will remain constant from generation to generation. One of its main uses in population genetics is to calculate allele and genotype frequencies. It provides a straightforward mathematical formula of how genetic equilibria can be sustained in a gene pool. Nonrandom mating, mutation, migration, genetic drift (changes in gene frequency within populations due to chance), and natural selection are some of the factors that induce population changes. The evolution of boundaries to flow of gene among populations is known as speciation**.** Speciation genes enable us to pinpoint the selective forces that led to the evolution of reproductive isolation while also emphasizing how shaky the connection is between the genetics of observable traits deviation and innate postzygotic reproductive barriers**.**

Molecular Genetics, Structures, Mechanisms, and Functions, Volume 1: Principles of Gene Manipulation and Genomics. Tariq Ahmad Bhat, Abdul Rauf Shakoori, and Jameel M. Al-Khayri (Eds.)
© 2026 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

11.1 INTRODUCTION

11.2 POPULATION

The total population of a species in a particular area at the same time and can interbreed in nature. For the interbreeding process to occur and produce fertile offspring individuals must be able to mate with each other and other members of a population, for example, Pakistan population, and animal population include Griffey plant and crop plants population.

11.3 GENETICS

The study of inheritance is called genetics. The gene is a DNA nucleotides sequence that directs the synthesis of RNA as well as protein and regulates the creation of further copies of the gene, serves as the chemical basis for genetic inheritance (James, 2003).

11.4 POPULATION GENETICS

Population genetics is a subarea of genetics that characterizes the shape of breeding populations. The forces of mutation, migration, choice, and genetic float will adjust the shape of populations (Merrick et al., 2023). Population genetics investigates genetic variations within and between populations of different organisms. Phenomena including speciation, population structure, and adaptation in this branch of biology are examined (Chaudhary). In the development of modern evolutionary theory, a key component was population genetics. Its primary founders who also laid the foundations of related disciplines of quantitative genetics were Sewall Green Wrigas an American Geneticist, J.B.S Haldane w British-Indian Scientist, and Sir Ronald Fisher was native Biologist, Geneticist. Modernized population genetics, which was once primarily a mathematical work, now includes theoretical, lab, and field study. Using population genetics models, from DNA sequence both statistical interference and concept proof/disproof are performed (Servedio et al., 2014).

11.4.1 ORIGIN OF SPECIES

History of biology is necessary to comprehend when genetic variation occurs. Charles Darwin published his book in 1859 “Origin of Species” that made two major assertions: first, that all living species today are derived from a single common ancestor, and second, natural selection is the primary

force driving evolution. The scientific community quickly accepted the first hypothesis but not the second. Many people had problems understanding that biological evolution could serve to improve understanding that Darwin's theory of evolution required of it. Through the early 20th century, people continued to agree that change had occurred but questioned Darwin's theory of how it had happened (Bowler, 1985).

11.4.1.1 DARWIN THEORY OF POPULATION GENETICS

Darwin's theory was resisted, which was understandable considering that Darwin's theory lacks a crucial explanation for how heredity works despite being compelling. Darwin's theory lacked a key explanation for how inheritance works. Parents and offspring must resemble one another for evolution through natural selection to take place; otherwise, fitness-enhancing features will not spread throughout a population. While acknowledging that he did not know why this was the case, Darwin based his thesis in the Origin on the observation that children indeed tend to resemble their parents, which he called "the strong principle of inheritance." But it was eventually determined that Darwin's stated theory of heredity, which was predicated on imaginary "gemmules," would have no real-world basis. Being unable to adequately address a significant challenge to his theory, Darwin was worried by his incomplete knowledge of the inheritance mechanism. If all the creatures in a community are the same, selection cannot take place, and a population cannot evolve via natural selection. So, Darwin's theory of evolution requires a constant supply of variety for selection to gradually alter a population over time. The accessible variety, according to Fleeming Jenkin, would be consumed too quickly (Jenkin, 1867). His justification was founded on the inheritance theory's "blending" idea, which states that an offspring's phenotypic traits "mix" into those of its ancestors. (For instance, if two creatures who are tall and small mate.) According to Jenkin's idea, phenotypic homogeneity will be reached in a species that reproduces sexually in a couple of generations, which is significantly quicker than the generations needed for natural processes to build complex adaptations.

However, Jenkin was mistaken about how inheritance works, supporting Darwin's theory. In contrast to "blending," the sort of inheritance we refer to as "Mendelian"—named after Gregor Mendel—is "particulate" because children inherit discrete genetic components (genes) from their parents. This indicates that sexual reproduction has no effect on the population's heritable variance. But for two reasons, it took a long time for this insight to arrive.

Firstly, the scientific community ignored Mendel's findings for 40 years. Second, even when Mendel's research was rediscovered at the turn of the 20th century, many people thought Darwinian evolution and Mendelian inheritance could not coexist.

11.4.1.2 MENDEL THEORY OF INHERITANCE

Mendel's theory of inheritance is based on simple concepts. Mendel noticed an interesting phenomenon while conducting his research on pea plants. Starting with two "pure breeding" genes, he produced plants with spherical seeds on one and plants with wrinkled seeds on the other. The first generation of daughters was then produced by his crossing them. The (w) wrinkled characteristic had been eliminated from the population, and all the F1 plants had spherical seeds (W). To create the F2 generation, Mendel crossed the F1 plants. Surprisingly, 25% or so of the F2 plants had wrinkled seeds. In other words, skipping a generation, the wrinkly trait had returned; Mendel offered the following explanations for these and related observations.

Darwinian gradualism was primarily interested in continuously changing features like height as opposed to "discrete" traits like seed shape that Mendel investigated. The Mendelians, led by William Bateson, were in opposition to the biometricians. They emphasized discontinuous variation and hypothesis rather than by cumulative natural selection in the way of Darwin, substantial adaptive change might be achieved by single mutational steps. The Mendelians and the biometricians got into a heated argument. Mendelian inheritance hence became linked to a non-Darwinian theory of evolution. In part, the desire to reconcile Mendel and Darwin led to the development of population genetics. This need grew more pressing as more and more empirical data supported Mendelian inheritance. A significant step in reconciling Darwin and Mendel was establishing that the distribution of traits with such continuous variation was compatible with Mendelism. This was due to the widespread belief that traits with such features would respond best to the Darwinian process. With the efforts of Fisher, Haldane, and Wright's mathematical work, complete reconciliation was reached in the 1920s and early 1930s (Fisher, 1930; Haldane, 1930–1932).

The dominant methodology of population genetics today still involves developing formal models to provide insight into the evolutionary process (Hartl, 2020). Both Fisher and Haldane were passionate toward Darwinists who held that natural selection was the primary influence on the genetic makeup of a population.

11.4.2 ALLELE

The frequency of alleles can be used to study and quantify genetic variation in a population. Population genetics depends upon two essential calculations: both genotype and allele frequencies.

11.4.2.1 ALLELE FREQUENCY

Allele frequency is a percentage or number that represents the average frequency of an allele (genetic variant) at a specific locus in a subpopulation (Gillespie, 2004). The slow change in allelic within a population is known as microevolution.

How to calculate allele frequency?

- **Monoploid organisms**

  The frequency of an allele (p) is a function product of the number of copies of that gene (A) in the population or sample size I (N).

$$p = i/N$$

- **Diploid organism**

  By counting alleles in a population, it is possible to determine the frequency of 3 different genotypes at a loci with two alleles, f(AA), f(BB), and f(AB), as well as the f allele A frequency of the allele A and the B-allele frequency of the B-allele (Aptroot et al., 2004).

$$p = f(AA) + 1/2\ f(AB) = A \text{ frequency of } A$$

$$q = f(BB) + 1/2\ f(AB) = B \text{ frequency of } B$$

  because frequencies of p and q of two alleles at a locus must come to 1. To check this

$$p + q = f(AA) + f(BB) + f(AB) = 1$$

$$q = 1\text{-}p \text{ and } p = 1\text{-}q$$

The population's number of copies is divided by the total number of copies of the ABO gene, or all the alleles, to get the frequency of an allele, like IA in the ABO biological system. Allele frequencies always = 1, or 100%, of the population, whether they are stated as a % or a decimal figure. For instance, the blood-type protein on the surface of red blood cells is determined by three alleles (IA, IB, or i) in the ABO blood type system in humans. The surface of a person's red blood cells will exhibit A-type protein (antigens) if they have the type of IA allele. In a sample population of humans, the frequency of the IA allele can be 0.26, meaning that 26% of the population's chromosomes carry the IA gene. We may determine the frequency of the I allele by using the formula below if we also know that this population's frequency of the IB allele is 0.14: 1 – 0.26 – 0.14 = 0.6. If the frequencies of any of these alleles changed over time, population evolution would take place. This frequency is calculated by deducting all known allele frequencies from 1. Those with type A blood have the genotype IAIA or IAi, type B have the genotype IBIB or IBi, type AB have the genotype IAIB, and type O have the genotype ii.

11.4.2.2 GENE FREQUENCY

The genotype frequency of a population is computed by dividing the proportion of individuals with a given genotype by the population's overall number of members. In population genetics, the genotype frequency is the frequency or ratio (i.e., $0< f <1$) of genotypes in a population.

Example: A population of 100 four o'clock (*Mirabilis jalapa*) plants with the relevant genotypes is used as an example. Of total, there are 49 plants with red blooms genotype (AA), 42 plants with pink blooms genotype (Aa) and 9 plants with white blossoms genotype (aa). In this illustration, each of the 9 white-flowered homozygotes has two copies of an allele a, as do each of the 42 pink-flowered heterozygotes. According to this finding, the allele frequency of aa is 0.3. In other words, an allele makes up 30% of the population's alleles for this gene. Calculate the genotype frequency of aa homozygotes to compare genotype frequencies white-flowered plant (9%). The sum of the allele and genotype frequencies is always one (100%).

Because of the accessibility of extensive genomic data and the requirement to analyze ever-more complicated evolutionary situations, population genetics is evolving into a data-driven field. Machine learning, and particularly deep learning algorithms, are becoming increasingly popular methods for making population genetic predictions as likelihood and Bayesian methods become either infeasible or computationally impractical. These

methods rely on algorithms that discover nonlinear correlations between input data and model parameters through representation learning on training sets of data. Deep learning techniques now used in science include exclusionary and generating networks having densely integrated, convolution, or recurring levels (Korfmann et al., 2023).

11.4.3 HISTORY OF POPULATION GENETICS

At the beginning of the 20th century, a conceptual framework called population genetics emerged to reconcile two divergent theories of evolution. As a result, it created a rich body of theory that transformed into an enormous collection of probabilistic models to create advanced statistical techniques whenever molecular data became available. The complexity of this corpus of theory has increased considering more plausible evolutionary and genetic situations as well as more effective computing techniques. As a result, statistical methods based on models have dominated the study of population genetics. One might even assert that most population geneticists would accept the idea of slightly altering George E.P. Box's axiom to read: "In our discipline, all models are inaccurate, but many are beneficial" (Provine et al., 2001).

The dominance of design statistics may help to explain why our profession has been slower to adopt deep learning techniques than other life science disciplines. The black-box aspect of deep learning is undoubtedly a major barrier to applications in the field of population genetics, whose fundamental goal is to identify the evolutionary and genetic factors underlying the variety of life on our planet. The apparent distinction in focus between the domains of machine learning and statistics is another barrier. In contrast to machine learning, which focuses on prediction using all-purpose algorithms to identify patterns in complex and massive data sets, statistics is concentrated on interpretation through the development and fit of a probability model (Bzdok et al., 2018). However as is evident from the broad interest in drawing conclusions about the demography history of species and the interest in identifying signs of natural selection or classifying individuals into groups, population geneticists are equally interested in inference and prediction. Despite this, many genetic clustering techniques and hence genomic scanning of selections are built on probabilistic models, sometimes mechanistic. For instance, Shape (Pritchard et al., 2000), Bayescan (Foll and Gaggiotti, 2008), DAPC (Jombart et al., 2010), and LFMM (Frichot et al., 2013) are two examples of phenomenological approaches.

The enormous data sets produced by next-generation sequencing technology have made the emphasis on design statistics in population genetics difficult (Levy and Myers, 2016). This is especially true for greatest and Bayesian procedures, which are implemented utilizing pricey computing techniques like Expectation-Maximization and Monte Carlo Markov chain. Approximate Bayesian computation (ABC), which uses summary statistics to extract the information inherent in raw populations' genetic data, can theoretically be used to minimize the computational burden of generating the conditional probability of exceedingly complex models (Bertorelle et al., 2010). Without calculating a probability function, ABC approximates the posterior probability of the parameter(s) that needs to be estimated. Instead, the set of simulation summary statistics that match the observed values yields a model fit (Beaumont et al., 2002). A popular and effective method for drawing conclusions about population genetics is ABC (Lopes and Beaumont, 2010). The "curse of dimensionality" refers to the fact that as the amount of summary statistics rises, the inaccuracy in concluding since collecting enough information necessitates a large number of many summary statistics (Prangle, 2015). Due to this issue, research in macrolearning strategies is rising (Schrider and Kern, 2018). The fact that genetics has been evolving over the past few years from a principal science into a data-driven area with access to huge amounts of genome and metadata is a significant driving force behind the change, nevertheless. For instance, researchers working on human population genetics have access to more than 150,000 rising entire sequencing data from the UK Biobank (Halldorsson et al., 2022), access to hundreds of genomic data from old samples, which come from more than 3000 people spread out over the globe (Byrska-Bishop et al., 2022).

11.4.4 GENETICS AND EVOLUTION OF PLANTS POPULATION

Recent years have seen a sharp rise in recent research focused on plant populations of genetics and evolution. While the motives for this activity are a little blurry, two things stick out. First, there is a growing understanding among plant scientists of population genetics' theoretical and conceptual approaches. The second is the understanding that plants are frequently quite receptive to combined ecological and genetic investigations. As a result, a lot of phenotype biologists, including some who did not receive official training in botany, have used plants as research subjects (Hamrick, 1982).

Now that several single-gene features are available to plant scientists, they can examine the evolutionary processes of populations. Furthermore,

the accessibility of novel tools has enabled physiological ecologists to explore plant adaptation at a functional level that was previously impossible. Over the past 15 years, several studies have found levels of genetic diversity among or between species of plant and animal species. In this research, natural populations were analyzed using biochemical techniques, notably electrophoresis (Hamrick, 1982).

Although the quality of these studies differs greatly between species and investigators regarding the number of populations they examined or the number of loci they studied, some fascinating results have been drawn. For instance, there is notably less variance in vertebrates. According to Nevo (1978), the average heterozygosity per person is 0.036, while it is 0.112 in insect populations and 0.141 in plant populations. Hamrick et al. (1979) may exhibit even higher degrees of variance. There is also significant heterogeneity in levels of intrapopulation variability across organisms within major taxonomic families (Powell, 1975; Nevo, 1978; Hamrick et al., 1979; Brown, 1979; Gottlieb, 1981). Hamrick et al. (1979) demonstrated that for plants, species with specific combinations of these characteristics, such as those that were widespread, long-lived, primarily outcrossed by wind, had opulent life fecundities, and were representative of the later phases of succession, maintained higher levels of intrapopulation genetic variation. These findings are impressive but require further testing through comparative studies among closely related species, with various dating, longevity, collaborators, etc. We are unable to comprehend the cause-and-effect link between an especially ecological characteristics and its allozyme variation until such research is made available. The plant's ready access to electrophoresis may be its greatest long-term benefit. Scientists of several genetic markers used to explore the principles of plant population evolution.

11.4.5 FLOW OF GENES

The mobility of genes among populations has proved challenging to assess and has generated a lot of debate. Previously, it was believed that gene flow between populations happened swiftly enough to "maintain the diversity of species." Recently, this viewpoint has come under fire (Ehrlich and Raven, 1969; Levin and Kerster, 1974). According to Levin and Kerster (1974), gene flow must typically be less than 1% between populations or population divisions that are a few hundred meters apart and at least two orders of magnitude less between populations that are 1.5 km or more apart. For instance, plants that are pollinated by the wind seem to transfer pollen far

more than the majority of species that are insect pollinated. According to statistics, fir trees and the grass *Agrostis tenuis* may be significantly affected by long-distance pollen migrations. Yet, most of the pollen produced by Phlox and Viola is distributed over distances of less than a meter, contrary to studies by Bateman (1947), which suggested that Brassica and Raphanus may have movement of pollen in appreciable volumes beyond 25 m. Indirect evidence for the differences between species that are pollinated by wind and those that are pollinated by insects comes from research by Brown (1979) and Hamricket et al. (1979), who have shown that wind-pollinated species had better variety within the populations and smaller variance across populations than an insect- or self-pollinated species.

11.4.6 MATING PROCESS IN THE ORGANISM

The genetic makeup of a species' population is greatly influenced by its mating system characteristics. For example, populations with a predominance of self-fertilized individuals in organisms are homozygous at many of their loci. Because of this, populations of selfing organisms have much fewer genotypes, and genotype frequency significantly deviates from predictions based on random number assumptions. Moreover, Brown (1979) has demonstrated that selfing and outbreeding populations have quite different patterns of genetic variation distribution crossover of species. In outcrossing species, polymorphic loci (H) have an average genetic variability per individual of 0.29, with a 20 W population-to-population variability ratio. For selfing species, H is equivalent to 0.13, and the ratio is 106%. In population subdivisions when H in outcrosses was 0.26 and the ratio was 3%, a similar pattern was seen. These numbers for self-starters were 0.24 and 32%. In plant research, calculations of the mating system are rather uncommon, even though they are crucial (Clegg, 1980). Important for numerous population genetic analyses, it is necessary to use a relatively common single-gene marker to perform quantitative calculations.

11.4.7 POPULATION GENETICS OF MAMMALS LOCATED AT ARBOREAL HABITAT IN A FRAGMENTED FORM LANDSCAPE

When patches are isolated from one another and individuals are unable to disperse, fragmentation can have an impact on the population and genetic diversity structure of arboreal animal populations (Lancaster et al., 2016; Lino et al., 2019; Knipler et al., 2022). This decrease in the connection may

result in a loss of allelic diversity and heterozygosity as well as promote inbreeding (Lacy, 1993; Weeks et al., 2016). They are particularly vulnerable to bottlenecks and the consequences of gene flow where habitat patches result in small, isolated populations (Willi et al., 2007). This can thus raise the possibility of isolated destruction (Reed, 2004). Identifying the genetic make-up of populations is the fundamental method for characterizing patterns and polymorphism variations between populations. The genotype structure and allele structure (i.e., the number of alleles present at each location or set of loci) together describe this genetic structure (A locus' genotype frequency distribution). Technically, a population is only in genetic equilibrium if certain conditions are met, including a sizable population that mates randomly, the absence of substantial mutation, equality in the fitness to procreate, and the absence of migration. In texts on population genetics, these groups in genetic equilibrium are frequently referred to Mendelian populations. According to the Hardy–Weinberg theorem (HWT), genotype frequencies will be stable in Mendelian populations and will have a multinomial dispersion (binomial for a model with one locus and two alleles). In other words, there are little long-term changes in the genotypic and allelic architecture. Mendelian populations will not undergo biological evolution as a result.

These presumptions hardly ever come true in actual life.

1. **No population** is endlessly huge, and even if one postulates a very huge and rising population, panmixia will not be achievable beyond a finite period since no natural system could maintain the energy required to ensure entirely random mating between more distant individuals (Valenzuela, 1973). Assortative mating has been empirically shown to exist even now in cosmopolitan communities, and it is also influenced by socioeconomic position, the shorter distance among places of residence or of birth, and other factors (Zei et al., 1983; Valenzuela et al., 1977). Unless the area being studied is very tiny, this effects create geographical effects (Chisholm et al., 1995). Moreover, religious traditions have a significant impact on assortative mating and may result in patterns of genetic isolation that are not limited by geography, as is the case with Jewish communities (Carmelli et al., 1979).

2. **Mutation**

   A large amount of variety can be produced quickly by the mutation rate at several loci (Kimura, 1991). According to Kimura's neutral hypothesis, the majority of the change in population genetic makeup

at the molecular level is caused by genetic drift acting on basically harmless loci exposed to recurrent mutation (Kimura, 1968).

3. **Selection**

   Strong evidence from populations with different diseases implies selection, defined as the difference in the ability to produce fertile descendants, has been active at numerous loci, as shown for alleles in the globin and Duffy loci connected to malaria protection, among other genes (Schaner et al., 2001).

4. **Natural population**

   Due to natural population growth and the ensuing desire for spatial expansion, populations move in intricate patterns (Barbujani, 2000). In the past, small migration to close towns was simpler than long migration. Because of this, at close distances between two populations, their two alleles will be comparable. Historically, the group of people in human populations has only involved a small number of people shifting to a new location that was primarily determined by the availability of food (Barbujani, 1997). Geographical constraints like mountains, deserts, and seas as well as cultural deterrents like linguistic and spiritual disparities, cause these newly established groups to expand quickly and separate from the others in these new enclaves. Genetic segregation, also known as genetic sampling, is the result of repeated inbreeding and gametic sampling mistakes in linguistic population genetics going to lead to micro differentiation or population subdivision, is the main effect when this process of separation persists over several generations. According to the idea of evolution, since all members of a species are descended from distant common ancestors, they are all somewhat inbred (Robert, 1972). Yet, a similar ancestor must not be too far away, or the notion loses its significance.

11.5 HARDY–WEINBERG LAW

A German physician, Wilhelm Weinberg, and an English mathematician, Godfrey Harold, independently proposed the Hardy–Weinberg law in 1908.

"Genotype and allele frequencies will predictably remain constant throughout generation in the absence of additional evolutionary factors" (Edwards, 2008).

Hardy–Weinberg principle (HWP), which is fundamental to the study of population genetics, works depending on a mathematical relationship that related genotype to allele frequencies and is an adaptation of the binomial distribution theorem. The Hardy–Weinberg equilibrium (HWE), which may identify both sexual and natural selection operating on a characteristic, and the historical significance of the Hardy–Weinberg principle (HWP) in explaining how interpretation is conserved during evolution are the reasons why the HWP is a mainstay of the introductory science bachelor program in the U.S. (Driver and McRae, 2021).

The terms genotypic frequency and gene frequency are defined. Along with the law's justification, the Hardy–Weinberg equilibrium law is introduced. HWE is examined rising the chi-square test. HWE is used in two applications: calculating carrier probability and determining recessive allele frequency (Xu, 2022).

When the population is enormous, mating occurred randomly and without selection, mutation, or migration reason for allele and gene frequency constant in a population (Graffelman et al., 2017). Human populations do not quite fit the criteria for HWE because of their change in allele frequencies from generation to generation.

11.5.1 HARDY–WEINBERG EQUATION

According to Graffelman et al. (2017, 2022), Merrick et al. (2023) the basic formula of HWE is:

- Two alleles of one gene "A" and "a" on the same locus p are presented as "A" frequency of allele and q is presented as "a" allele frequency.

Then the next generation will have

- P^2 represented homozygote frequency "AA" alleles and q^2 represented homozygote frequency "aa" alleles.
- 2pq represented heterozygote frequency "Aa" alleles.

Hardy–Weinberg equations based on binomial distribution:

$$p^2 + q^2 + 2pq = 1$$

In population recessive and dominant frequency of genes can be determined using this formula:

- p represented dominant gene frequency "A" and q represented the recessive gene frequency "a."
- p^2 represented homozygous dominant gene frequency "AA" and q^2 represented homozygous recessive gene frequency "aa."
- 2pq represented heterozygous gene frequency "Aa."

Example: 100 pea plants in a garden, 86 plants may produce yellow peas, whereas 16 may produce green peas. The number of (Yy) homozygous dominant or (Yy) heterozygous individuals is unknown, but 16 are homozygous recessive (Yy) (yy). The frequency of that genotype is 0.16. If homozygous recessive plants are represented by q^2, then q^2 must equal 0.16. Hence, q equals 0.4. The frequency of that genotype is 0.16. If q^2 represents homozygous recessive plants, then q^2 has to be equal to 0.16. Thus, q is equal to 0.4. Consequently, if p + q = 1, p is in an equal framework. The frequency of homozygous dominant plants is p 0.36. The total number of individuals for Y are homozygous plants is 36. The frequency of heterozygous plants is 0.48.

11.5.2 ASSUMPTIONS OF HWE

Population genetics is studied based on the Hardy–Weinberg law. It offers a mathematical framework for understanding genetics and evolution.

In the population, often known as the HW balance nonevolutionary model, the frequency of genes is constant across many generations.

The evolution rate is zero when the population is in genetic equilibrium. In other words, a population that abides by the Hardy–Weinberg law will not evolve. Hence, evolution can only take place when the Hardy–Weinberg equilibrium is disturbed.

The H–W equations determined the allele and genotype frequencies relationship based on five conditions (Merrick et al., 2023).

1. Allele changes cannot happen because there can be no gene mutations.
2. There can better be inflowing or nonflowing population migration.
3. Random mating must occur.
4. Genetic drift, a haphazard shift in allele frequency, is not possible.
5. Natural selection, a shift in allele frequency brought on by the environment, is not permitted.

11.5.3 RELATIONSHIP BETWEEN GENOTYPE AND GENE FREQUENCIES

The Hardy–Weinberg law explains the characteristics of a population to a single autosomal locus (Merrick et al., 2023) as follows:

- Gene and genotype frequencies are stable in a huge random mating population without genetic variation.
- Genotypes frequency between the parents alone dictates the frequencies of genotypes in the offspring produced by random mating between the parents.
- Homozygotes frequency = relevant gene frequency square.
- Heterozygote frequency = relevant gene frequencies square.
- So, the squares of the gametic frequencies correspond to the genotype frequencies of the homozygotes.
- It is said that such populations exhibit genetic equilibrium. The Hardy Weinberg frequencies of genotype for a single autosomal locus are due to generation of random mating, nevertheless the genotype frequencies between parents.
- Two alleles with a single locus: maximum heterozygous frequency will be 0.5. Hence $p = q = 0.5$.

To determine homozygous and heterozygous variant carriers in populations that are not evolving, HWE uses allele frequency. To find genotyping mistakes that might result in significant heterozygote excess (HetExc), large population databases have been utilized to measure deviations from HWE. HetExc may also be the consequence of natural selection, as advantageous recessive disease variants may reach high allele frequencies in a heterozygous state despite being less prevalent in a homozygous form in the population (Abramovs et al., 2020).

11.6 APPLICATIONS OF HW LAW

Applications of the HW law include:

- Calculating the population frequencies of a recessive and dominant gene.
- Calculating the population's frequencies of heterozygotes.
- Hardy–Weinberg equilibrium to test agreement with a population.

11.6.1 CALCULATING THE POPULATION FREQUENCIES OF RECESSIVE AND DOMINANT GENES

- It is possible to estimate gene frequencies from their genotype frequencies, but doing so requires knowledge of the frequencies of all three genotypes.
- The expression q is equal to Q + 12 H which cannot be used because complete dominance at a locus makes it impossible to classify genotypes.
- Because it cannot be distinguished phenotypically from the dominant homozygotes, the genotype frequency of heterozygotes cannot be estimated.
- For example: Garden pea plant stem length: tall stems are dominant (TT, Tt), while dwarf stems are recessive (tt).
- The frequencies of all genotypes are not necessary, though, if they are all in Hardy–Weinberg equilibrium.
- The genotype frequencies are not necessary, though, if they are all in Hardy–Weinberg equilibrium:
    - p represented dominant gene frequency "T" and q represented recessive gene frequency "t," q^2 has presented the homozygote frequency "tt," and homozygote frequency is the square root of the recessive allele frequency.
- Recessive homozygotes should not be selectively eliminated for this estimation to be accurate.
- Given that p + q = 1.
- p (dominant allele frequency) = 1 – q.

11.6.2 CALCULATING THE POPULATION'S FREQUENCIES OF CARRIERS OR HETEROZYGOTES

- Knowing the prevalence of heterozygotes, carriers of recessive genes, or individuals with recessive defects is frequently interesting. If the frequency of the gene is known, this can be determined.
- The formula 2pq = 2q (1-q) can be used to estimate the frequency of heterozygous in the population if HW equilibrium is assumed.
    - For instance, if $q^2 = 0.04$, q is the gene frequency = (0.04) 1/2 = 0.2.
    - The frequency of a heterozygote is 0.32.
    - Carriers or heterozygotes make up 32% of the population.

11.6.3 HARDY–WEINBERG EQUILIBRIUM TO TEST FOR AGREEMENT WITH A POPULATION

- Information is given for a locus of every genotype is distinguishable; the observed genotype frequencies utilized for the HWE test.
- The genotype frequencies of offspring are calculated using the gene frequencies of their parents by the Hardy–Weinberg law.
- The frequency of offspring and parents is equal in a population that is in H–W equilibrium.
- The assumed genotype frequencies are determined from the gene frequency. The predicted numbers are obtained from these. Using the Chi-square test, it is determined whether the observed and expected numbers arrive.

11.7 HARDY–WEINBERG LAW-MULTIPLE ALLELES

- In multiple allelic systems, gene frequency calculation is easy if dominance is absent.
- Simple gene counting provides the most accurate estimate of gene frequency.
- Three transferrin alleles, A, D, and E, for instance, are found in European cow breeds.
- As there is no dominance, six genotypes AA, AD, AE, DD, DE, and EE can be identified.
- If we use these symbols to represent the numbers of each genotype and N to represent the overall number, the following formulas to find out the frequencies of gene:

$$f(A) = p = \frac{2AA + AD + AE}{2N}$$

$$f(D) = q = \frac{2DD + DA + DE}{2N}$$

$$f(E) = r = \frac{2EE + EA + ED}{2N}$$

- The equilibrium genotype frequencies for a three-allelic system can be mathematically represented as

$$p + q + r = 1$$

$$(p + q + r)^2 = p^2 + q^2 + r^2 + 2pq + 2pr + 2qr = 1.$$

- Three multiple allelic genes, A, B, and O alleles, together determine the ABO blood types in man.
- If dominance exists, the estimate is a little more difficult and requires the use of approximations.

11.7.1 MORE THAN ONE LOCUS

- All autosomal loci taken individually reach an equilibrium in genotype frequencies after one generation of random mating.
- However, when all genotypes are taken into account, it is not true.
- Think about a population that is equally composed of AABB and aabb people of both sexes. In both loci, the gene frequency is 0.5.
- The following genotypes may result from the random mating of the individuals.
- Only two of the original homozygotes (AABB and aabb) and the double heterozygote (AaBb), or three out of nine genotypes, would show up in the progeny in the following generation.
- The genotype AAbb would not exist even though its frequency in the equilibrium population would be 1/16.

11.7.2 LINKED LOCI

- When random mating occurs, connected loci reach equilibrium more slowly than isolated loci do.
- The closer the relationship, the more slowly the equilibrium approaches. When equilibrium is attained, both coupling and repulsion phases occur frequently:
  - Heterozygotes paired.
  - AB/ab are the repulsion heterozygous: Ab/aB.

11.8 FACTORS AFFECTING HWE

Following evolutionary forces deviate this equilibrium:

- Nonrandom mating.
- Migration and gene flow.
- Mutation.
- Genetic drift.
- Natural selection.

11.8.1 NONRANDOM MATING

Nonrandom mating occurs when two non-similar individuals mate for all possible pairs of individuals. Nonrandom mating occurs when individuals are closer together (inbreeding) or unrelated partners more often than would be expected by chance population. Nonrandom mating from a sure area of interest is probably certainly considered one among the reasons in transferring Hardy–Weinberg equilibrium. This is because of heterozygous scarcity in splendid populace at various origins (Graffelman, 2017).

The mutual selection of phenotypic or behavioral types results in assortative mating, which is not random mating. Assortative mating can result from social rank-based competition for mating in polygynandrous animals. Social relationships, however, have an impact on assortative mating prospects even though they are not an individual attribute (Ricardo et al., 2023). Many conditions, including phages, can lead to nonrandom mating among bacteria. This can be explained by the fact that uncommon transduction events will naturally lead to some bacterial genotype pairs occurring while others do not (Abedon, 2022).

Nonrandom mating is of two types:

11.8.1.1 ASSORTATIVE MATING

Assortative mating happens when two members of the same population are more alike than would be predicted by chance (Xu, 2022; Lewontin et al., 1968; Burley, 1983; Crespi, 1989).

Because of the strong correlation between the phenotypes of the paired individuals, assortative mating frequently takes place. For additional traits, such as morphology and behavior, as well as for an increasing number of

taxa, including humans, plants, non-human animals, and fungi, similarities between mating partners are increasingly being documented (Buss, 1985; Jiang et al., 2013; Trendily and Leuchtmann, 2019; Weis et al., 2005).

Several factors about this correlation may be brought about by several factors, including (1) phenotype-based mate selection, (2) fine-scale either spatial or temporal stationarity of phenotypes, including circumstances in which space or time is viewed as "traits," (3) phenotypic changes over time, increasing phenotypic correlations between mates, or (4) selection patterns on the pool of potential breeders (Jiang et al., 2013; Kopp et al., 2018; Rios Moura et al., 2021; Sørdalen et al., 2018).

11.8.1.2 DISASSORTATIVE MATING

"If the mated pairs are of different phenotypes, it is known as disassortative mating. In these types of mates are genetically different."

Disassortative mating is an unusual type of mate preference that encourages polymorphism's persistence. Hence, heterozygote advantage should encourage disassortative mating, which increases the number of heterozygous offspring. Disassortative mating pairings are more likely to result in fitter offspring, which encourages this strange mating behavior when heterozygous offspring gain from increased fitness (Maisonneuve, 2021).

The self-incompatibility in families of plant, such as Brassicaceae, where the Inhibits S-locus fertilization among indicate of individual of identical allele, serves as an illustrative example (Durand et al., 2020). Behavior of disassortative mating in animals is also documented, albeit it is typically not very rigid (Jiang et al., 2013). In humans (Wedekind et al., 1995) and mice (Potts et al., 1991), it has been shown that the locus of immunity-related MHC (histocompatibility complex) mediating particular observation of peptides is associated with disassortative mate choice, with females favoring bodily excretions linked with MHC alleles. Nonetheless, a recent meta-analysis found that primates have a relatively low tendency disassortative mating related to MHC (Winternitz et al., 2017).

Classified Mating:

- Positive: mating between mates who have same phenotype.
- Negative: mating between individuals has different phenotypes inbreeding breeding between relatives. "Negative-assortative mating for color in wolves" from Yellowstone National Park was described in Evolution, 70, 757–766, as the "first recorded occurrence of major negative mating in animals" (Muir, 2023).

11.8.1.3 INBREEDING COEFFICIENT

Wildlife populations' conservation may depend on how inbreeding depression is identified. Genome-wide marker data can be used to calculate the coefficients of personal inbreeding. It is unclear to what extent substructure of population genetics affects predictions of inbreeding despair in wildlife populations, which affects the sensitivity of inbreeding coefficients (Zilko et al., 2020). Males and females may be adversely affected by inbreeding depression, which is the decreased fitness of a linked individual's progeny (Vega-Trejo et al., 2022).

Compare the actual proportions of heterozygous genotype with expected during random mating.

$$F = (H_O - H)/ H_O$$

F is coefficient of inbreeding, decreased heterozygosity.

H = heterozygotes frequency
H_O = heterozygote number in random matings

What happens in selfing?
Selfing powerful inbreeding system allows you to reach your highest potential. Several generations of homozygosity at the same time, heterozygosity is reduced.

11.9 FORCES CHANGING GENE AND GENOTYPE FREQUENCY

11.9.1 SYSTEMATIC PROCESS

- They alter the gene frequencies in an anticipated manner in both direction and amount.
- They act both in large and small population.

Three systematic processes include:

- Migration
- Mutation
- Selection

11.9.2 DISPERSIVE PROCESS

They develop in tiny populations because of sampling effects. While the quantity can be predicted, the direction cannot.

Due to the impact of sampling, they behave in tiny populations**.**

11.9.2.1 MIGRATION

- It is individual movement, or introduction of genes within different populations.
- Migration is becoming more diverse, with rates because it is large, the frequency changes significantly.
- The change in gene frequency is frequency difference between beneficiary population and average donor population.

Many insect species, especially pests of herbivorous crops, frequently migrate. Spodoptera exigua, a polyphagous armyworm, migrates over great distances, although little is known about its underlying environmental factors. In this study, we used population genetics to identify the movement patterns of *S. exigua* across several Asian countries (Wang et al., 2022).

Immigration is the transfer of people into one population from another permanently.

Emigration is the transferred to people from one population to another. The size of the gene pool is diminished.

It is typical for migration to act as an evolutionary pressure, causing groups to become more genetically similar. Under predictable conditions herbal choice will be working with inside the contrary direction, and its miles genuinely of hobby to peer to what extent one pressure will be triumphant over the other.

For example, salmon start life in rivers and migrate to sea to get food and grow. After spending up to seven years in ocean, they migrate to rivers where they were born so they can spawn.

From a genetic point of view, migration does not just mean individual migration however, this move may lead to new alleles (gene flow). Movement causes changes in gene frequency because more copies of an existing allele are introduced, or this is because new alleles are added.

Several factors affect migration of crop species:

- Breeding system
- Compassion for wild and/or weedy relatives
- Pollinators
- Seed dispersal

A direct effect of migration increases the population of genetic diversity, thus helping to increase the resilience of this population to tolerate environmental changes. Migration also helps mix and prevent population deviation.

Consider a population where the native's frequency is 1-m and there are m new immigrants every generation. Let q0 stand for the frequency of the gene in the native population and qm for the frequency of the gene in the immigrant population. In the mixed population, the frequency will then be:

$$q1 = qm + (1-m)\, q0$$

$$q1 = m(qm-q0) + q0$$

Migration causes a change in gene frequency that can be measured by alleles, the frequency of allele before and after migration.

$$\Delta q = q1-q0$$

$$\Delta q = m(qm-q0)$$

As a result, both the velocity of migration and the differential in allele frequency of allele among the immigrants and native populations affect the change caused by migration. A population may acquire new alleles at a rate and a greater number of loci than would be anticipated from mutation due to migration or gene flow (Merrick et al., 2023).

The genome-wide frequency of allele data is increasingly being utilized to calculate the estimated amount of migration of processes, or edges (m), on population trees using the software Treemix. It is unclear, though, how many edges should be included. The optimum of the m value can be deduced from the 2nd rate of change in probability (△m) over successive values of △m. Populations of m is stimulated so that m performs just as well as the current Treemix recommendations when used for its intended purpose, which was to estimate the number of population groups (△K) represented by the structure of software (Fitak et al., 2021).

11.9.2.2 MUTATION

It is crucial to recognize that Monroe et al. (2022) made several false claims concerning the mutation and selection of population For instance, they claimed that "the widespread fact that genetic variants are discovered

less frequently in functionally limited parts of the genome is assumed to be owing exclusively to selection after random mutation."

Mutation is a change in DNA sequence. It may be caused by mistakes during cell divisions or exposure of DNA-damaging agents in the environment. It may lead to the occurrence of new alleles, so it changes the gene pool of population.

The concept that mutations occur randomly regarding their impact has dominated evolutionary theory since the first half of the 20th century. Here, we utilize massive de novo mutation surveys in Arabidopsis thaliana plant to test this presumption. In parts of the genome that are functionally constrained, we discover that the frequency of mutations is reduced by two-thirds in important genes and by half inside gene bodies, in contrast to expectations. We demonstrate that epigenomic and physical factors account for over 90% of the variability in the genome-wide pattern of mutation bias surrounding genes using different genomic mutation datasets, including those from the largest Arabidopsis mutation accumulation experiment.

Researching the reaction of mutation rate to population environments aids in our understanding population of destiny and dynamics in novel contexts since mutation is the basis of evolutionary material and because organisms constantly meet environmental changes. Mutation rates can change and adapt. For instance, it has been discovered that several stressful situations, including extreme hunger, antibiotic exposure, and high population densities, increase the plastically enhanced mutation rates of bacteria. Moreover, hypermutator allele that boost spontaneous mutation rates—can become fixed in microbial populations that have evolved in challenging conditions (Galhardo et al., 2007; Liu and Zang, 2019; Long et al., 2016; Summaa et al., 2006; Weilgoss et al., 2013), for example, at a faster mutation rate during about 2400 generations in glucose-limited medium (Good et al., 2017).

Our understanding of a long-lasting genome is that changing in genomes is malleable and adaptable to environmental modifications. A growing body of research demonstrates that a range of environmental pressures cause genetic instability in human cancer cells, yeast, and bacteria, leading to the sporadic appearance of mutants that are more fit and possibly hastening adaptive evolution (Wielgoss et al., 2013).

A distinct purpose for mutation exists. Since it causes arbitrary changes to the genetic code, it is in a way a harmful force. Such alterations are overwhelmingly likely to be harmful in any specially engineered organism. Making arbitrary alterations to a well-crafted watch or making arbitrary changes are the standard analogies we use in such situations. While random

changes occasionally improve the accuracy of the watch or the impact of the poem, given more chances, one is more likely to make things worse. In a manner that is comparable, migration can also cause organisms to be relocated into settings that are not suitable for them.

11.9.2.3 SELECTION

11.9.2.3.1 Natural Selection

Both a limited and comprehensive view of natural choice may be taken. When seen narrowly, it is unmistakably one of the elegances of defying the Hardy–Weinberg Laws, specifically when viability or fertility depends on genotype. According to a broad definition, it is the primary force that drives evolution's inventive and forward-thinking component, adaptation. Ecological processes, which determine the diversity of circumstances and the fitness of phenotypes, developmental processes, which determine the influence of genotype on phenotype, and population genetics, which inform us of changes in genetic makeup in a population while the fitness's of the genotypes are known, could all be combined to form a complete theory of evolution, one that no longer exists**.**

- Alleles that are favored by selection increase frequency and vice versa.
- The consequences of selection on variety are:
    - Directivity, so variety is reduced.
    - Balanced for diversity. Heterozygotes have the highest suitability such that selection favors maintenance of multiple alleles.

When goats were domesticated, artificial selection for desirable features was found, which tended to result in some alleles wanting specific properties. This shaped genetic variation among populations. Due to this, numerous breeds that were generated and raised for a variety of purposes were established and given birth. Despite the strength of artificial selection, natural selection continues, possibly leaving a more subtle impact over time that may be difficult to discern. In order to examine selection footprints in response to environmental change, we genotyped 993 goats from 4 transboundary goat species (Angora, Boer, Nubian, and Saanen) using the SNP chip 50 K. We used outlier detection, runs of homozygosity, and haplotype-based detection strategies to analyze the whole herd (Manunza et al., 2023).

11.9.2.3.2 Artificial Selection

Artificial selection is the steps of modifying individuals in a population accidently or on purpose by human intervention, resulting in the phenotypic evolution of both animals and plants. Urbanization of food crops derived from populations of hectic ancestors and advancement via the breeding of crops were important changes for plants in the rise of human civilization. As a result, crop varieties underwent intense selection at the genes governing key agronomic properties. As a result, it is reasonable to be supportive that genes identified as the objects of artificial selection are crucial genetic regulators of agronomic features (Vigouroux et al., 2002b; Yamasaki et al., 2005).

Over the past 50 years, dairy animal breeds have supported various artificial selections to improve features related to milk production. To determine the important physiological routes and areas that existed impacted by this selection, in this investigation we ran a total genomic analysis for variation using 42,486 SNPs in the 3 dominant French dairy breeds of cattle (Holstein, Normande, and Montbéliarde) (Flori et al., 2009).

Goldfish breeding for aesthetic reasons can be traced back to Chinese writings from the Song dynasty, making it the most ancient fish species to be subject to artificial selection. During the Ming dynasty, many breeds emerged that are morphologically divergent, indicating that artificial selection has had a significant impact on the development of goldfish morphology. Goldfish are a wonderful model for studying the evolution of fish shape because of their extensive diversity and obvious history of artificial selection (Ota and Abe, 2016).

It is unknown if evolutionary adaptation will organize sharply enough to assure the persistence of many species in the current atmosphere of global warming. In particular, we lack the understanding regarding the capacity for adaptation of the various populations about the many limitations that can impede specific adaptations. There is evidence that plants in temperate zones often benefited from early flowering as an adaptive mechanism in response to global warming. As a result, the goal of this study was to evaluate the ability of *Lupinus angustifolius L*. to adapt to advance the onset of flowering (Fabaceae). Four populations in Spain from 2 various latitudes were collected, and their seeds were then grown in a public garden. Three different early flowering selection lines were established over three generations, using both self-crosses and outcrosses, by selecting 25% of the individuals who flowered earlier in the first generation. In the northernmost populations, all artificial selection lines significantly advanced their flowering compared to the control line, but not in the southern populations (Sacristan-Bajo et al., 2023).

Complete selection against dominant gene

- The coefficient of selection is 1 or fitness is 0.
- One generation of selection is enough to eliminate the dominant genes, provided there is complete penetrance.
- In the next generation, all the individuals will be of recessive homozygotes and the frequency of a recessive allele (q) will be one.

11.9.3 GENETIC DRIFT

Genetic drift is the shift in population frequency of a pre-existing gene variant brought on by random chance. Gene variations may vanish due to genetic drift, hence reducing genetic variation. Additionally, it might lead to the substantially greater frequency and possibly fixation of previously rare alleles. Founder effect and bottleneck are two terms referring to various processes by which a small population develops reproductive isolation, the two basic types of genetic drift.

Linkage disequilibrium is a secondary result of genetic change. Multiple haplotypes make it impossible to leave the population precisely in a state of linkage stability since genetic drift will fluctuate the frequencies of each one. It does not typically couple certain alleles more frequently, of course. There is the potential for both good and bad associations for any two sets of alleles, one at each locus.

The most popular breed of dog both internationally and in Brazil is the German Shepherd. Understanding the history of the Brazilian German Shepherd dog breed needs a knowledge of the population structure through pedigree records. The Brazilian German Shepherd dog breed's genetic variety and population network was evaluated. A total of 77,938 animals born between 1970 and 2014 were featured in the pedigree. In this population, there were 3.91 generations on average each year. 2,183 founders were revealed when the reference population was factored in. During 2010 and 2014, the native population lost nearly 3% of its genetic diversity, mainly on account of genetic drift (Bignardi and Santana, 2023).

In the breeding of plants and animals, it is important to predict the genetic gain that will arise from artificial selection of a trait. Lush's classical breeder's equation (BE), which could also be exploited to evolution under the natural selection, is usually used for this purpose. Breeders may be able

to detect both the additive and nonadditive genetic impacts at the individual gene level owing to the recent use of high-throughput sequencing techniques (He et al., 2023).

Research has indicated that the scale and rate at which this is currently happening is greatly decreasing the adaptive possibility of the greater part of species that inhabit human-impacted landscapes. Genetic significance of population subdivision is regarded essential to long-term evolution. Here, we present a rare and, to our knowledge, the first study of this process in real time to demonstrate how to evaluate the efficiently mitigating actions like specie crossings. We accomplished this by using a huge genetic dataset accumulated from a koalas population that exist and understanding the full while a linear transport system was being built, which caused the population to be separated. First, we observed that the effective population size and allelic richness both decreased over the time and have shown that after just 10 generations, the two separated populations could face a genetic diversity loss of between 12% and 69% (Frere et al., 2023).

11.10 SPECIATION

The evolution of boundaries to flow of gene among populations is known as speciation Understanding the process of natural selection is mainly responsible for special showed an important theory of speciation yet needs to consider the forces of selection as well as the genetic and genomic processes that induce divergence. In the beginning, we explain how finding "speciation" genes enables us to pinpoint the selective forces that led to the evolution of reproductive isolation while also emphasizing how shaky the connection is between the genetics of observable traits deviation and innate postzygotic reproductive barriers. Second, we explore the crucial role that recombination suppressors play in promoting speciation with gene transfer, but we also note that it is still unclear how and when reproductive obstacles become connected to recombination cold spots. Third, we establish the significance of ancient genomic variation in speciation; however, we contend that conclusions drawn from studies of speciation that concentrate on relatively recent groups may be biased in favor of old variation rather than recent changes (Schluter and Rosenberg, 2022).

KEYWORDS

- **population genetics**
- **gene pool**
- **gene frequency**
- **allele frequency**
- **Hardy–Weinberg equation**
- **on-random mating**
- **mutation**
- **migration**

REFERENCES

Abedon, S. T. Phage Impact on Non-random Mating Non-random Mating Among Bacteria Bacteria Non-random Mating. In *Bacteriophages as Drivers of Evolution: An Evolutionary Ecological Perspective*; Springer International Publishing: Cham, 2020; pp 95–102.

Abramovs, N.; Brass, A.; Tassabehji, M. Hardy-Weinberg Equilibrium in the Large Scale Genomic Sequencing era. *Front. Genet.* **2020,** *11*, 210.

Aptroot, A.; Czarnota, P.; JÜ, I.; Kocourkova, J.; Esslinger, T. L. Department of Biological Sciences, North Dakota State University, Fargo. *Bryologist* **2004,** *107* (2), 250–261

Barbujani, G. DNA Variation and Language Affinities. *Am. J. Hum. Genet.* **1997**, *61* (5), 1011–1014.

Bateman, A. J. Contamination of Seed Crops: I. Insect Pollination. *J. Genet.* **1947,** *48*, 257–275.

Beaumont, M. A.; Zhang, W.; Balding, D. J. Approximate Bayesian Computation in Population Genetics. *Genetics* **2002,** *162* (4), 2025–2035.

Bertorelle, G.; Benazzo, A.; Mona, S. ABC as a Flexible Framework to Estimate Demography Over Space and Time: Some Cons, Many Pros. *Mol. Ecol.* **2010,** *19* (13), 2609–2625.

Bignardi, A. B.; Santana Júnior, M. L. Genetic Diversity of Shepherd Dog using Pedigree analysis. *Ciência Rural* **2023,** *53*.

Bowler, P. J. In *Evolution*: *The History of an Idea*, 1985.

Brown, A. H. D. Enzyme Polymorphism in Plant Populations. *Theor. Popul. Biol.* **1979,** *15* (1), 1–42.

Burley, N. The Meaning of Assortative Mating. *Ethol. Sociobiol.* **1983,** *4*, 191–203.

Buss, D. M. Human Mate Selection. *Am. Sci.* **1985,** *73* (1), 47–51.

Byrska-Bishop, M.; Evani, U. S.; Zhao, X.; Basile, A. O.; Abel, H. J.; Regier, A. A.; Corvelo, A.; Clarke, W. E.; Musunuri, R.; Nagulapalli, K.; Fairley, S. High-Coverage Whole-Genome Sequencing of the Expanded 1000 Genomes Project Cohort Including 602 Trios. *Cell* **2022,** *185* (18), 3426–3440.

Bzdok, D.; Altman, N.; Krzywinski, M. Statistics Versus Machine Learning. *Nat. Publ. Gr*. **2018,** *15*, 233–234.

Carmelli, D.; Cavalli-Sforza, L. L. The Genetic Origin of the Jews: A Multivariate Approach. *Hum. Biol.* **1979,** 41–61.

Chaudhary, S. Breeding of Disease Resistance and Pathogen Population Genetics in Plants**.**

Chisholm, B. In *The History and Geography of Human Genes*; Cavalli-Sforza, L. L., Menozzi, P., Piazza, A., Ed.; Princeton; *J. Asian Stud.* **1995,** *54* (2), 490–492.

Clegg, M. T. Measuring Plant Mating Systems. *Bioscience* **1980,** *30* (12), 814–818.

Crespi, B. J. Causes of Assortative Mating in Arthropods. *Anim. Behav*. **1989,** *38*, 980–1000.

Driver, R. J.; McRae, S. B. Considering Instructional Approach & Question Design with the Hardy-Weinberg Principle. *Am. Biol. Teacher* **2021,** *83* (3), 191–194.

Edwards, A. W. F. GH Hardy (1908) and Hardy–Weinberg equilibrium. *Genetics* **2008,** *179* (3), 1143–1150.

Ehrlich, P. R.; Raven, P. H. Differentiation of Populations: Gene Flow Seems to be Less Important in Speciation than the Neo-Darwinians Thought. *Science* **1969,** *165* (3899), 1228–1232.

Fisher, R. A. XV-The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Earth Environ. Sci.* **1930,** *52* (2), 399–433.

Fitak, R. R. OptM: Estimating the Optimal Number of Migration Edges on Population Trees using Treemix. *Biol. Methods Protoc.* **2021,** *6* (1).

Flori, L.; Fritz, S.; Jaffrézic, F.; Boussaha, M.; Gut, I.; Heath, S.; ... Gautier, M. The Genome Response to Artificial Selection: A Case Study in Dairy Cattle. *PloS One* **2009,** *4* (8), e6595.

Foll, M.; Gaggiotti, O. A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective. *Genetics* **2008,** *180* (2), 977–993.

Frère, C. H.; O'Reilly, G. D.; Strickland, K.; Schultz, A.; Hohwieler, K.; Hanger, J.; ... Sherwin, W. Evaluating the Genetic Consequences of Population Subdivision as it Unfolds and How to Best Mitigate Them: A Rare Story about Koalas. *Mol. Ecol.* **2023**.

Frichot, E.; Schoville, S. D.; Bouchard, G.; François, O. Testing for Associations between Loci and Environmental Gradients using Latent Factor Mixed Models. *Mol. Biol. Evol.* **2013,** *30* (7), 1687–1699.

Galhardo, R. S.; Hastings, P. J.; Rosenberg, S. M. Mutation as a Stress Response and the Regulation of Evolvability. *Crit. Rev. Biochem. Mol. Biol.* **2007,** *42* (5), 399–435.

Gillespie, J. H. In *Population Genetics*: *A Concise Guide*, 2004.

Good, B. H.; McDonald, M. J.; Barrick, J. E.; Lenski, R. E.; Desai, M. M. The Dynamics of Molecular Evolution over 60,000 Generations. *Nature* **2017,** *551* (7678), 45–50.

Gottlieb, L. D. Electrophoretic Evidence and Plant Populations. *Prog. Phytochem.* **1981,** *7*, 1–46.

Graffelman, J.; Jain, D.; Weir, B. A Genome-Wide Study of Hardy–Weinberg Equilibrium with Next Generation Sequence Data. *Hum. Genet.* **2017,** *136*, 727–741.

Graffelman, J.; Weir, B. S. The Transitivity of the Hardy–Weinberg Law. *Forensic Sci. Int. Genet.* **2022,** *58*, 102680.

Haldane, J. B. S. In *A Mathematical Theory of Natural and Artificial Selection*, Mathematical Proceedings of the Cambridge Philosophical Society, 1930–1932; vol (23), pp 607–615.

Halldorsson, B. V.; Eggertsson, H. P.; Moore, K. H.; Hauswedell, H.; Eiriksson, O.; Ulfarsson, M. O.; Stefansson, K. The Sequences of 150,119 Genomes in the UK Biobank. *Nature* **2022,** *607* (7920), 732–740.

Hamrick, J. L. Plant Population Genetics and Evolution. *Am. J. Bot.* **1982,** *69* (10), 1685–1693.
Hamrick, J. L.; Holden, L. R. Influence of Microhabitat Heterogeneity on Gene Frequency Distribution and Gametic Phase Disequilibrium in Avena barbata. *Evolution* **1979,** 521–533.
Hartl, D. L. In *A Primer of Population Genetics and Genomics*; Oxford University Press, 2020.
He, Z. H.; Xiao, Y.; Lv, Y. W.; Yeh, F. C.; Wang, X.; Hu, X. S. Prediction of Genetic Gains from Selection in Tree Breeding. *Forests* **2023,** *14* (3).
Mauseth, J. D. In *Botany: An Introduction to Plant Biology*, 3rd Ed.; 2003; p 470.
Jenkin, F. The Origin of Species. *N. Brit. Rev.* **1867,** *46* (92), 277–318.
Jiang, Y.; Bolnick, D. I.; Kirkpatrick, M. Assortative Mating in Animals. *Am. Nat.* **2013,** *181* (6), E125–E138.
Jombart, T.; Devillard, S.; Balloux, F. F. Discriminant Analysis of Principal Components: A New Method for the Analysis of Genetically Structured Populations. *BMC Genet.* **2010,** *11*, 94.
Kimura, M. Genetic Variability Maintained in a Finite Population Due to Mutational Production of Neutral and Nearly Neutral Iso Alleles. *Genet. Res.* **1968,** *11* (3), 247–270.
Kimura, M. The Neutral Theory of Molecular Evolution: A Review of Recent Evidence. *Jap. J. Genet.* **1991,** *66* (4), 367–386.
Knipler, M. L.; Dowton, M.; Clulow, J.; Meyer, N.; Mikac, K. M. Genome-Wide SNPs Detect Fine-Scale Genetic Structure in Threatened Populations of Squirrel Glider Petaurus Norfolcensis. *Conserv. Genet.* **2022,** *23* (3), 541–558.
Kopp, M.; Servedio, M. R.; Mendelson, T. C.; Safran, R. J.; Rodríguez, R. L.; Hauber, M. E.; Scordato, E. C.; Symes, L. B.; Balakrishnan, C. N.; Zonana, D. M.; van Doorn, G. S. Mechanisms of Assortative Mating in Speciation with Gene Flow: Connecting Theory and Empirical Research. *Am. Nat.* **2018,** *191* (1), 1–20.
Korfmann, K.; Gaggiotti, O. E.; Fumagalli, M. Deep Learning in Population Genetics. *Genome Biol. Evol.* **2023**, *15* (2).
Lacy, R. C. Impacts of Inbreeding in Natural and Captive Populations of Vertebrates: Implications for Conservation. *Perspect. Biol. Med.* **1993,** *36* (3), 480–496.
Lancaster, M. L.; Cooper, S. J.; Carthew, S. M. Genetic Consequences of Forest Fragmentation by Agricultural Land in an Arboreal Marsupial. *Landscape Ecol.* **2016,** *31*, 655–667.
Levin, D. A.; Kerster, H. W. Local Gene Dispersal in Phlox. *Evolution* **1974,** 130–139.
Levy, S. E.; Myers, R. M. Advancements in Next-Generation Sequencing. *Ann. Rev. Genomics Hum. Genet.* **2016,** *17*, 95–115.
Lewontin, R.; Kirk, D.; Crow, J. Selective Mating, Assortative Mating, and Inbreeding: Definitions and Implications. *Eugen. Q.* **1968,** *15* (2), 141–143.
Lino, A.; Fonseca, C.; Rojas, D.; Fischer, E.; Pereira, M. J. R. A Meta-Analysis of the Effects of Habitat Loss and Fragmentation on Genetic Diversity in Mammals. *Mamm. Biol.* **2019,** *94*, 69–76.
Liu, H.; Zhang, J. Yeast Spontaneous Mutation Rate and Spectrum Vary with Environment. *Curr. Biol.* **2019,** *29* (10), 1584–1591.
Long, H.; Miller, S. F.; Strauss, C.; Zhao, C.; Cheng, L.; Ye, Z.; ... Lynch, M. Antibiotic Treatment Enhances the Genome-Wide Mutation Rate of Target Cells. *Proc. Natl. Acad. Sci.* **2016,** *113* (18), E2498–E2505.
Lopes, J. S.; Beaumont, M. A. ABC: A Useful Bayesian Tool for the Analysis of Population Data. *Infect. Genet. Evol.* **2010,** *10* (6), 825–832.

Maisonneuve, L.; Beneteau, T.; Joron, M.; Smadi, C.; Llaurens, V. When do Opposites Attract? A Model Uncovering the Evolution of Disassortative Mating. *Am. Nat.* **2021,** *198* (5), 625–641.

Manunza, A.; Diaz, J. R.; Sayre, B. L.; Cozzi, P.; Bobbo, T.; Deniskova, T.; ...; Stella, A. Discovering Novel Clues of Natural Selection on Four Worldwide Goat Breeds. *Sci. Rep.* **2023,** *13* (1), 2110.

Merrick, L.; Meade, K.; Campbell, A.; Muenchrath, D.; Beavis, W. Population Genetics. *Crop Genet.* **2023**.

Monroe, J. G.; Srikant, T.; Carbonell-Bejerano, P.; Becker, C.; Lensink, M.; Exposito-Alonso, M.; ... Weigel, D. Mutation Bias Reflects Natural Selection in Arabidopsis Thaliana. *Nature* **2022,** *602* (7895), 101–105.

Muir, C. D. Technical Comment on "Negative-Assortative Mating for Color in Wolves". *Evolution* **2023,** *77* (2), 627–632.

Nevo, E. Genetic Variation in Natural Populations: Patterns and Theory. *Theor. Popul. Biol.* **1978,** *13* (1), 121–177.

Ota, K. G.; Abe, G. Goldfish Morphology as a Model for Evolutionary Developmental Biology. *Dev. Biol.* **2016,** *5* (3), 272–295.

Potts, W. K.; Manning, C. J.; Wakeland, E. K. Mating Patterns in Seminatural Populations of Mice Influenced by MHC Genotype. *Nature* **1991,** *352* (6336), 619–621

Powell, J. R. Protein Variation in Natural Population of Animals. *Evol. Biol.* **1972,** *8*, 79–119.

Prangle, D. Summary Statistics in Approximate Bayesian Computation, 2015.

Pritchard, J. K.; Stephens, M.; Donnelly, P. Inference of Population Structure using Multilocus Genotype Data. *Genetics* **2000,** *155* (2), 945–959.

Provine, W. B. In *The Origins of Theoretical Population Genetics*; University of Chicago Press: Chicago, IL, 1971.

Provine, W. B. In *The Origins of Theoretical Population Genetics: With a New Afterword*; University of Chicago Press, 2001.

Reed, D. H. Extinction Risk in Fragmented Habitats. *Anim. Conserv. Forum* **2004,** (7), 181–191.

Ricardo, M. C.; Jaqueline, L. F.; Dafne, G. G. M.; Hernández, L. Close Bonds and Social Rank Similarities Favor Non-Random Mating in Captive Stump-Tailed Macaques (Macaca arctoides). *Ethology* **2023**.

Rios Moura, R.; Oliveira Gonzaga, M.; Silva Pinto, N.; Vasconcellos-Neto, J.; Requena, G. S. Assortative Mating in Space and Time: Patterns and Biases. *Ecol. Lett.* **2021,** *24* (5), 1089–1102.

Roberts, D. F. The Genetics of Human Populations. *J. Med. Genet.* **1972,** *9* (3), 382.

Sacristán-Bajo, S.; García-Fernández, A.; Lara-Romero, C.; Prieto-Benítez, S.; Tabarés, P.; Morente-López, J.; ... Iriondo, J. M. Population Origin Determines the Adaptive Potential for the Advancement of Flowering Onset in *Lupinus angustifolius L.* (Fabaceae). *Evol. Appl.* **2023,** *16* (1), 62–73.

Saumaa, S.; Tarassova, K.; Tark, M.; Tover, A.; Tegova, R.; Kivisaar, M. Involvement of DNA Mismatch Repair in Stationary-Phase Mutagenesis During Prolonged Starvation of Pseudomonas Putida. *DNA Rep.* **2006,** *5* (4), 505–514.

Schaner, P.; Richards, N.; Wadhwa, A.; Aksentijevich, I.; Kastner, D.; Tucker, P.; Gumucio, D. Episodic Evolution of Pyrin in Primates: Human Mutations Recapitulate Ancestral Amino Acid States. *Nat. Genet.* **2001,** *27* (3), 318–321.

Schluter, D.; Rieseberg, L. H. Three Problems in the Genetics of Speciation by Selection. *Proc. Natl. Acad. Sci.* **2022,** *119* (30).

Schrider, D. R.; Kern, A. D. Supervised Machine Learning for Population Genetics: A New Paradigm. *Trends Genet.* **2018,** *34* (4), 301–312.

Servedio, M. R.; Brandvain, Y.; Dhole, S.; Fitzpatrick, C. L.; Goldberg, E. E.; Stern, C. A.; Yeh, D. J. Not Just a Theory the Utility of Mathematical Models in Evolutionary Biology. *PLoS Biol.* **2014,** *12* (12).

Sørdalen, T. K.; Halvorsen, K. T.; Harrison, H. B.; Ellis, C. D.; Vøllestad, L. A.; Knutsen, H.; Moland, E.; Olsen, E. M. Harvesting Changes Mating Behaviour in European Lobster. *Evol. Appl.* **2018,** *11* (6), 963–977.

Treindl, A. D.; Leuchtmann, A. Assortative Mating in Sympatric Ascomycete Fungi Revealed by Experimental Fertilizations. *Fungal Biol.* **2019,** *123* (9), 676–686.

Valenzuela, C. Y. Block Mating Selection Model in a Random Mating Population. *Revista Medica de Chile* **1973,** *101* (3), 203–206.

Valenzuela, C. Y.; Harb, Z. Socioeconomic Assortative Mating in Santiago, Chile: A Demonstration using Stochastic Matrices of Mother-Child Relationships Applied to ABO Blood Groups. *Soc. Biol.* **1997,** *24* (3), 225–233.

Vega-Trejo, R.; de Boer, R. A.; Fitzpatrick, J. L.; Kotrschal, A. Sex-Specific Inbreeding Depression: A Meta-Analysis. *Ecol. Lett.* **2022,** *25* (4), 1009–1026.

Vigouroux, Y.; Jaqueth, J. S.; Matsuoka, Y.; Smith, O. S.; Beavis, W. D.; Smith, J. S. C.; Doebley, J. Rate and Pattern of Mutation at Microsatellite Loci in Maize. *Mol. Biol. Evol.* **2002,** *19* (8), 1251–1260.

Wang, X.; Yang, X.; Zhou, L.; Wyckhuys, K. A.; Jiang, S.; Van Liem, N.; Wu, K. Population Genetics Unveils Large-Scale Migration Dynamics and Population Turnover of Spodoptera Exigua. *Pest Manag. Sci.* **2022,** *78* (2), 612–625.

Wedekind, C.; Seebeck, T.; Bettens, F.; Paepke, A. J. MHC-Dependent Mate Preferences in Humans. *Proc. R. Soc. Lond. Ser. B Biol. Sci*. **1995,** *260* (1359), 245–249.

Weeks, A. R.; Stoklosa, J.; Hoffmann, A. A. Conservation of Genetic Uniqueness of Populations may Increase Extinction Likelihood of Endangered Species: The Case of Australian Mammals. *Front. Zool.* **2016,** *13*, 1–9.

Weis, A. E.; Winterer, J.; Vacher, C.; Kossler, T. M.; Young, C. A.; LeBuhn, G. L. Phenological Assortative Mating in Flowering Plants: The Nature and Consequences of its Frequency Dependence. *Evol. Ecol. Res.* **2005,** *7* (2), 161–181.

Wielgoss, S.; Barrick, J. E.; Tenaillon, O.; Wiser, M. J.; Dittmar, W. J.; Cruveiller, S.; ... Schneider, D. Mutation Rate Dynamics in a Bacterial Population Reflect Tension between Adaptation and Genetic Load. *Proc. Natl. Acad. Sci.* **2013,** *110* (1), 222–227.

Willi, Y.; Van Buskirk, J.; Schmid, B.; Fischer, M. Genetic Isolation of Fragmented Populations is Exacerbated by Drift and Selection. *J. Evol. Biol.* **2007,** *20* (2), 534–542.

Xu, S. Basic Concept of Population Genetics. In *Quantitative Genetics*; Springer International Publishing: Cham, 2022; pp 25–37.

Yamasaki, M.; Tenaillon, M. I.; Vroh Bi, I.; Schroeder, S. G.; Sanchez-Villeda, H.; Doebley, J. F.; ... McMullen, M. D. A Large-Scale Screen for Artificial Selection in Maize Identifies Candidate Agronomic Loci for Domestication and Crop Improvement. *Plant Cell* **2005,** *17* (11), 2859–2872.

Zei, G.; Guglielmino, C. R.; Siri, E.; Moroni, A.; Cavalli-Sforza, L. L. Surnames as Neutral Alleles: Observations in Sardinia. *Hum. Biol.* **1983,** 357–365.

Zilko, J. P.; Harley, D.; Hansen, B.; Pavlova, A.; Sunnucks, P. Accounting for Cryptic Population Substructure Enhances Detection of Inbreeding Depression with Genomic Inbreeding Coefficients: An Example from a Critically Endangered Marsupial. *Mol. Ecol.* **2020,** *29* (16), 2978–2993.

CHAPTER 12

Sexuality in Bacteria

MAHNOOR IMRAN, MARIA BUTT, TANZEELA REHMAN, BUSHRA TABASSUM, and UZMA QAISAR

School of Biological Sciences, University of the Punjab, Lahore, Pakistan

ABSTRACT

Sexual reproduction is common among living creatures to bring variability, diversity, and evolution. Sexual reproduction involves mixing of DNA from two parent organisms through the process of mating followed by recombination of genetic material to produce new combinations in the off-springs. The mating process varies widely among various forms of life. At one time it was thought that bacteria had no sex, no DNA or nucleus. Though some observers claimed that they had beheld the tiny creatures in the act of conjugation, the prevailing view denied them. With the development of new techniques, equipment, and research, it is now a profound fact that bacteria undergo sexual process. Biochemists had found that bacteria contain deoxyribonucleic acid (DNA), the helical molecule that is found in chromosomes. Cytologists, in turn, had shown that the DNA in bacteria is concentrated in bodies as it is in the nuclei of other cells. This DNA is indeed the bearer of hereditary characteristics in bacteria. Scientists were able to transfer a purified extract of DNA from one strain of *Pneumococcus* sp. to another. Bacteria exchange their genetic material with each other by different methods. Three methods of sexual genetic transfer in bacteria, that is, conjugation, transduction, and transformation are discussed in detail in this chapter.

12.1 INTRODUCTION

For many years, it was thought that bacteria were unique organisms with no genetic makeup or nuclei, although their biochemical resemblance with other

Molecular Genetics, Structures, Mechanisms, and Functions, Volume 1: Principles of Gene Manipulation and Genomics. Tariq Ahmad Bhat, Abdul Rauf Shakoori, and Jameel M. Al-Khayri (Eds.)
© 2026 Apple Academic Press, Inc. Co-published with CRC Press (Taylor & Francis)

life forms established the basis of comparative biochemistry. Evidences have shown that bacteria possess characteristics more similar to other life forms, such as genes, nucleus, and in some cases mechanism of sexual reproduction for recombination of characters. Most of the microbes, including prokaryotes, without any nucleus reproduce asexually. Similarly, bacteria are also grouped under prokaryotes because they are unicellular and do not have a well-defined cell system. Bacterial cells grow in an exponential manner, in each division (generation) which takes around twenty minutes, it doubles its number. Bacteria are prokaryotes that show both sexual and asexual modes of reproduction. During asexual reproduction, bacterial cells divide through mitotic cell division giving rise to daughter cells which are exactly similar to their parents. While in sexual reproduction, the genetic material of two similar organisms is combined and mixed to give rise to new genetic combinations. This exchange of genetic material allows the bacteria to develop new characteristics and genetic makeup. In this way, they can survive and adapt to the environmental changes. In this chapter, we are going to discuss the sexual reproduction in bacteria.

During the course of evolution, the bacterial ability of surviving in the new environment and to cope up with the environmental challenges depends on the acquisition of new genes from other bacteria and environment via horizontal gene transfer (HGT) (Lawrence, 2002). The presence of xenobiotic degradation and antibiotic resistance genes in bacteria is generally because of the horizontal gene transfer and environmental pressure triggered by spread of increased amount of these substances in the surrounding environment (Aminov, 2011; Mazel, 2001). In the same way, pathogenicity islands are DNA segments with virulence genes acquired by a variety of pathogens including *Helicobacter pylori*, *Yersinia pestis, Salmonella typhimurium, Dichelobacter nodosus,* and variants of *Escherichia coli* at some time in the past.

After the *E. coli* genome sequence success, a wide range of analysis techniques for frequency of HGT (horizontal gene transfer) of genetic material was made (Lawrence and Ochman, 1998). It was predicted that the separation of phylogenetic families of *Salmonella enterica* and *E. coli* occurred around hundred million years ago. Later, the HGT has resulted in the acquisition of new characters and finally separation of phylogenetic lineages. For example, the phenotypic characters that differentiate *S. enterica* from *E. coli* (lactose utilization, citrate utilization, pathogenicity, and indole production) are all because of the acquisition of chromosomal genes as a result of the divergence from common ancestor (Lawrence and Ochman, 1998). Study on

the difference in the codon utilization and base composition patterns allowed an assessment of the frequency of HGT. For example, *E. coli* has acquired 17.6% of genes through horizontal gene transfer at a rate of 16 kilo bases/ million years. These are just estimations as the codon utilization and base composition similarity of transferred DNA to that of *E. coli* could not be measured (Ochman and Groisman, 1994).

The horizontal transfer of genes in a variety of bacterial species has been detected in labs. The most important transfer is by conjugation, in which a portion of Ti plasmid from *Agrobacterium tumefaciens* is transferred between different kingdoms, that is, to plants (Zupan and Zambryski, 1995) and yeast (Bundock and Hooykaas, 1996). DNA transfer to *S. cerevisiae* through conjugation can be affected by the *E. coli* F-plasmid (Heinemann and Sprague, 1989). It was observed that the kanamycin resistance gene in transgenic plants can transform *Acinetobacter* species (kanamycin resistance) (Gebhard and Smalla, 1998). The bacteria can transfer the DNA to other species by using three methods: conjugation; contact-dependent DNA transfer that occurs in most bacteria, transformation; uptake and incorporation of naked DNA and transduction; host DNA encapsulated in bacteriophage (vector) that transfers it to recipient DNA.

These methods for genetic transfer have enhanced our knowledge about molecular genetics of bacteria and served as a tool for genetic engineering technology. In the mid-1980s, biotechnology for recombinant DNA technology was completely developed. But the public concern about the accidental release of transgenic DNA in the environment and its possible consequences was still an issue which shows that we still have very minimal knowledge on the transfer of genetic material in the surrounding natural environment.

This knowledge is essential for the study of likelihood of the release of non-recombinant microbes deliberately into the environment for such agricultural purposes as plant growth stimulation (*Agrobacterium, Pseudomonas, Rhizobium, Azospirillium*), bioremediation of xenobiotic-polluted sites (*Alcaligene*), insect control (*Bacillus thuringiensis*), nitrogen fixation control of phyto-pathogenic fungi and bacteria, and weed control by phyto-pathogenic fungi (Davison, 1999).

It was recommended that the characteristics of the microorganisms released in the environment might be enhanced by genetic modification (Wilson and Lindow, 1993). The situation is unlike from unintentional release of microbes in the environment because in this case the bacteria are normally disabled or tamed laboratory bacteria which are unable to survive

in the natural environments. On the other hand, the deliberate release of (recombinant or natural) microorganisms would have an aim of stable conservation and ability to function in the particular environment. Thus, a balanced evaluation of the level of HGT in the environment is required. Data is needed regarding the survival, persistence, stress, nutrition of the bacteria with particular niche (van Veen et al., 1997).

12.2 CONJUGATION

12.2.1 HISTORICAL ASPECTS

Bacterial conjugation is one of the modes of transfer of genetic material between bacteria and the only mode involving cell-to-cell connection. This transfer of genetic material was reported in *Escherichia coli* by Lederberg and Tatum in 1946. The discovery was made during the search of sexual recombination in bacteria, in which the genetic markers from both the partners were transferred to the progeny. Conjugation is a sexual mode of DNA transfer in which the chromosomal content of two distinct bacterial cells is brought close to each other in a programmed manner (Llosa et al., 2002). Unlike higher organisms in which whole sets of chromosomes of parent cells are involved during sexual reproduction, the bacterial conjugation involves only some part of the genome from one cell (donor) and complete genome of the partner cell (recipient). Thus, this transfer is partial as only small portion is transferred and it is polar, from one cell (donor) to the other cell (recipient) (Fig. 12.1). After this the two cells are separated and additional changes in genetic material (by recombination and genetic organization) of recipient cell take place. The genetic material transfer takes few minutes to hours. In 1958, Tatum, Beadle, and Lederberg received Noble Prize for their work on conjugation.

They observed, when two different autotrophs (nutritional mutants) were grown on the same minimal media together, they produced a phototroph (wild type). This process required cell-to-cell contact. In 1957, Anderson observed this process in an electron microscope. Later on, conjugation was reported in many other bacteria. Bacterial cell showing conjugation is dimorphic; they have two cell types the male donor (F^-) and the female recipient (F^+). The male donor cell type contains the sex pili (1–4 in number) on their surface with fertility factor also known as sex factor. The fertility factor (F-factor) contains the genes for sex pili formation and gene transfer. The F factor and sex pilus is absent in female recipient cells (Curtiss III, 1969).

When the two cell types come closer to each other, the F factor allows the male donor to produce thin tube-like projections called pilus/conjugation tube (Tatum and Lederberg, 1947). The pilus is used to draw a connection between the two cell types which allow the transfer of genetic material (a small piece of DNA) from male donor to female recipient. During this exchange of genetic material, the female recipient often gets the genetic advantage, that is, in the form of antibiotic resistance genes. It normally takes 6–8 min. Gene exchange can occur by two methods.

12.2.1.1 STERILE MALE METHOD

The plasmid with the fertility factor replicates. One copy of the plasmid transfers into the recipient cell. The recipient cell also becomes donor and this procedure of reversible sex is known as seduction.

12.2.1.2 FERTILE MALE METHOD

The plasmid containing fertility factor gets integrated into the chromosomal DNA and this attachable plasmid is known as episome. This process of integration happens because some nucleotides of the chromosomal DNA and the F factor are compatible with each other. The donor cell with fertility factor integrated into its chromosome is known as high frequency of recombination (Hfr), or super male. It has thousand times more recombination frequency in comparison to normal F^+.

The nonintegrated F^+ plasmids are disintegrated. The breakage of integrated F^+ factor at one end of bacterial chromosome allows replication. After replication, a copy of the free end of chromosome transfers into the recipient cell through sex pili or conjugation tube. The fertility factor is the last to transfer and usually it does not pass into the recipient cells because the whole chromosome is not transferred into the recipient cells. Few genes transfer in limited time, that is, 1 in 7 min, 2 in 9 min, 3 in 10, and so on. Conjugation produces incomplete zygotes known as partial zygotes (Skaar et al., 1957).

12.2.2 STAGES OF BACTERIAL CONJUGATION

Some phases of conjugation have been investigated. A brief review of the stages is given below:

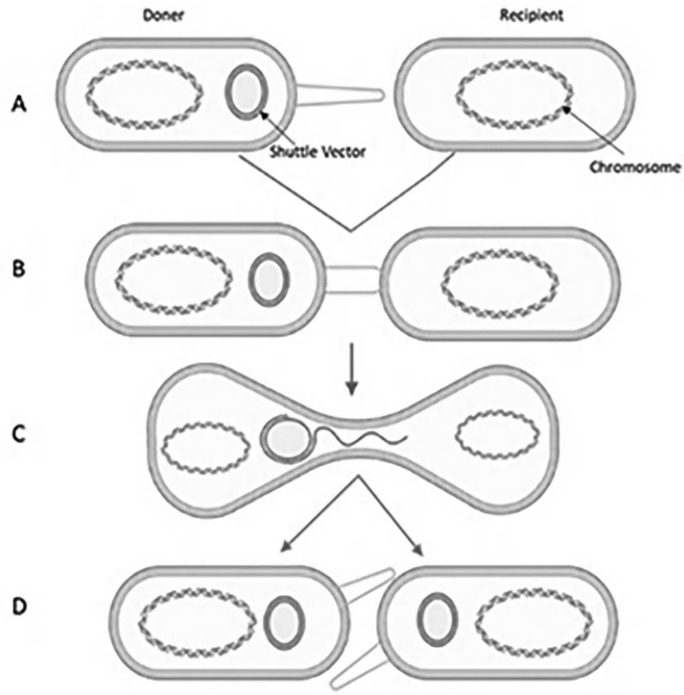


FIGURE 12.1 Schematic diagram of bacterial conjugation. (A) Donor cell produces pilus; (B) pilus attaches to recipient cell, brings the two cells together; (C) the mobile plasmid is nicked and a single strand of DNA is then transferred to the recipient cell; (D) both cells circularize their plasmids, synthesize second strands and reproduce pili. ⏎

Source: Adapted from Deng et al., 2017.

12.2.2.1 CELL-TO-CELL ATTACHMENT

Several genes are involved for cell-to-cell attachment of two bacteria. In case of gram-negative bacteria, genes for sex pili are responsible for the contact between two cells involved in genetic transfer. For example, in *E. coli* 37 genes from tra/trb operon are involved in the production of 15 gene products that are responsible for production and assembling of the pili. Some conjugative plasmids produce pili similar to the F pili (F-like). These pili range from 5 to 12 nm in diameter and have central hollow hole with a diameter of 2–3 nm.

In gram-positive bacteria like *Enterococcus faecalis*, the cell attachment does not involve the pili. The recipient cell produces a diffusible pheromone that causes aggregation and clumping of the donor cells, followed by gene transfer events. In *Streptomyces* overlapping growth with each other results in conjugation (Trevors, 1997).

12.2.2.2 RELAXOSOME FORMATION

For transfer of DNA also called Dtr, a DNA–protein complex is formed (relaxosome) at part of the DNA known as oriT, which defines the site wherever DNA is going to be nicked (Lawley et al., 2004). The nicking enzymes have two domains: one for the recognition of the oriT region and the other domain produces nick nearby and becomes attached to the nick site by covalent bond. Some other proteins along with the enzyme form a relaxosome and help in promoting the nicking activity by destabilizing the duplex. The protein complex (relaxase) may have a regulated activity. The initiation of DNA transfer activates the helicase activity.

12.2.2.3 DNA TRANSFER

The gene transfer from donor to recipient takes place in the form of single-stranded DNA. This phenomenon is regulated by mobilization genes or Mob genes and type IV secretion system (T4SS) (Alvarez-Martinez and Christie, 2009). The complex contains the 5′ end from the transferred strand. When oriT transfers into the recipient, another nick is formed and the original 5′ of the recipient attaches to the 3′ end (newly formed). This forms a circle with single strand, which is further replicated into a double strand with the help of primase activity of MobA (from plasmid R1162) (Perwez, 1998).

12.2.2.4 DONOR CONTROL OF CONJUGATION

Most conjugation system genes are regulated negatively (most cells do not express) in a particular population. For example, if one cell in a given population is able to conjugate and transfer its DNA to new host, it will stimulate the transfer system in the new host allowing speedy conjugation and rapid production of plasmid in the new population. Afterward, the process of conjugation is repressed and gene expression slows down.

12.2.3 REQUIREMENTS FOR CONJUGATION

There is an extensive range of variations among the details of conjugation mechanism within subgroups of bacteria, but there is a common theme that applies to every group. Below are mentioned some common functions

required for conjugation process, whether an element is conjugative (can transfer itself on its own) or needs help in transfer (mobilize able). A necessary site of DNA is required called oriT also known as basis of mobility. If a cell contains all the suitable conjugation functions, any DNA or gene in it attached to the oriT can be transferred (even the element with all other functions but without oriT is not transferable). This transfer is known as chromosome/plasmid mobilization. The position of oriT where cleavage occurs is known as nick. Additional proteins are also required for conjugational donorsA nicking enzyme also known as trans esterase recognizes the oriT region. This enzyme works in multi-protein complex with other required activities. This enzyme is known as nickase/relaxase and the enzyme with the oriT region is known as relaxosome. Few examples of nicking enzymes are TrwC enzyme from R388 and MobA from R1162.

Protein that couples the transport apparatus with the complex of DNA and relaxes is required. T4SS coupling protein (T4CP) acts as a mediator/coupling protein. Some examples are TraG from RP4**,** TraD from F factor, and VirD4 from the Ti plasmid. Mating pair formation (MPF) is subgroup of the T4SS. They are involved in forming the intracellular transport apparatus. The T4SS system is a complicated association of several gene products, i.e., pilin subunits. As a part of symbiotic or pathogenic lifestyle, other families of T4SS transport proteins. Integrases are also significant for conjugational elements. They use these integrases to remove the DNA from the donor and incorporate in the recipient. These recombinases are site specific. They identify the bases specifically on the end of the elements for deletion and recircularization, and determine the particular sequence to reintegrate in the genetic material of recipient cells.

12.2.4 CONJUGATIVE PLASMIDS

Conjugative plasmids also known as self-transmissible plasmids contain all the information and machinery required for their independent transfer in the recipient cells. RP4, a conjugative plasmid that is able to transfer DNA in many host species, can transfer between different bacterial species. Its example is gram-positive bacteria and *Enterobacterium,* since most members in this group are pathogenic or else produce useful things like cheese and antibiotics. All the members from the *Enterobacteriaceae*, including *E. coli*, *Salmonella, Shigella*, and *Klebsiella*, are able for conjugational transfer. Conjugational transfer between archaea has also been observed (Clewell, 2013).

Inter-kingdom conjugational transfer occurs in nature with the help of Ti plasmid that induces tumor and crown gall disease in plants. The tumor-producing DNA gets incorporated in the plant genome. Transfer among *E. coli* and yeast has also been reported. Some gram-positive bacteria can exchange a double-stranded DNA in the conjugation process in contrast to the transmission of single strand through T4SS. This process is much simpler because of the noninvolvement of the nicking enzymes and involving fewer steps, but the mechanism is not yet clear (Wilkins et al., 1993). Conjugative plasmids require ancillary gene providing support to the conjugational machinery. In some circumstances the conjugative plasmids contain several accessory genes other than those involved in conjugation. Like antibiotic resistance genes present on plasmids in humans, such factors that are able to transfer are known as R factors. Incompatibility group (Inc) includes the plasmids that are incapable to replicate in the same cell. *Enterobacteriaceae* contains 30 different Inc groups. Replication controller features are responsible for this incompatibility.

12.3 TRANSDUCTION

Transduction is the phenomenon in which foreign DNA is transferred into a cell by means of a virus or by viral vector. This transfer of genetic material through virus is an example of horizontal gene transfer (HGT). This process does not involve any physical interaction between two cells (i.e., donor and acceptor cells). Molecular biologists use this technique to transfer foreign gene in host cells for biotechnological benefits. The virus infecting the bacterial cell is termed as phage or bacteriophage.

The phages that transfer DNA are known as transfusing phages, but all are not transfusing. The phenomenon of transduction is capable of transferring DNA fragments of 10–1000s of kilo bases. Transduction has a lower host range for DNA transfer as compared to other methods because of high specificity of phages for cell-surface receptors. The transduction process is resistant to DNases while transformation is not and transduction does not require any physical contact between two bacteria (Lederberg, 1956).

12.3.1 HISTORY

When a provirus turns into virulent virus causing the lysis of the host cell, they carry some of the genetic material from host cell. When such virus

infects a new cell/host, it is able to transfer the genetic material from the previous host to the new one (Fig. 12.2). This process of gene transfer through viruses was first revealed by Lederberg and Zinder in 1952.

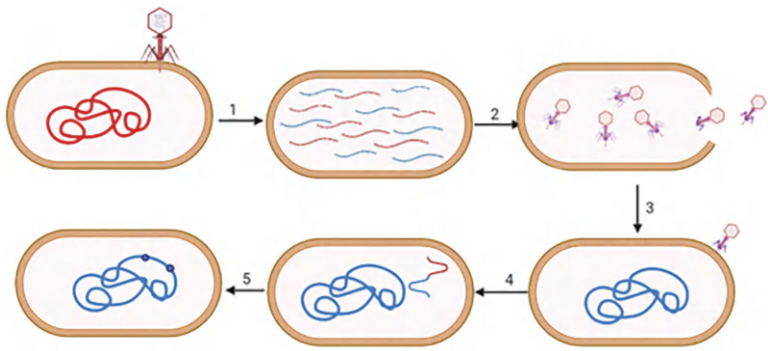


FIGURE 12.2 Bacterial Transduction. (1) Attachment of phage to the host cell (2) Replication of viral particles in the host cell (3) Packing and assembly of recombinant phage genome by the host (4) Infection of recipient cell with the phage (5) Release of recombinant phage genome in recipient cells (6) Integration of genome in recipient cells. ⏎

Source: Adapted from Schneider, 2021.

They discovered the process of transduction through an experiment, commonly known as U-tube experiment. In this experiment, a "u"-shaped tube was taken and separated into two compartments with a filter, allowing only virus particles to pass through the filter. One side of the tube contained the lysogenic bacteria with the latent virus and the other side had the susceptible bacteria. The lysogenic bacteria are lysed and the virus particles are released, which pass through the filter and infect the bacteria. This caused destruction of the susceptible bacteria and release of the viral particles. These particles carried the DNA from the susceptible bacteria and traveled back to the lysogenic bacteria and transferred the genetic material. The two bacterial cells were separated by filter so no conjugation occurred. DNase was also added to the tube to destroy the free DNA, so there was no possibility of transformation (Lederberg, 1955).

Transduction takes place through lytic or lysogenic cycle in bacteria. In the lysogenic cycle, the viral/phage genetic material gets incorporated into the bacterial chromosome by covalent bond formation. It remains dormant for many generations. If it is induced by any means, that is, UV light exposure, it can activate itself and cleave off from the bacterial genome to initiate

the lytic cycle. It ends by lysing the bacterial cells and releases the phage elements. Lytic cycle results in the assembly of new phage particles released by host cell lysis (Lederberg and Lederberg, 1953).

12.3.2 STEPS OF BACTERIAL TRANSDUCTION

The DNA transfer depends on the viral infection by a stepwise manner (Morse et al., 1956). Bacteriophage infects the bacterial cells. The virus overtakes host machinery to replicate either directly as in case of lytic cycle or indirectly by lysogenic cycle. During bacteriophage assembly in the host cells some part of the bacterial DNA get incorporated into the viral genome mistakenly. In lysogenic cycle, when the incorporated viral DNA detaches from the bacterial chromosome it takes some part of the bacterial DNA flanking the viral genome and gets assembled in the bacteriophage. When these bacteriophages infect other host cells, they incorporate their DNA along with the donor DNA into the new cells. The bacterial DNA can be incorporated into the recipient chromosomal DNA or remains as a plasmid in the form of extrachromosomal DNA. It can also get incorporated into the viral DNA. The fate of transfer DNA depends on the nature of bacteriophage and the recipient cell DNA.

12.3.3 FATE OF DNA INSERTED INTO THE RECIPIENT CELL

When the foreign DNA gets incorporated into the recipient cell, it has three fates; DNA will be absorbed by the cells and recovered for spare parts; if the DNA was a plasmid before, it can circularize and in the new cell form extra chromosomal DNA; if the foreign DNA sequence is homologous to the recipient's DNA it can exchange genetic material from the chromosomal DNA of recipient cells. This recombination is random and depends on viral genome size (Lacks et al., 1967).

12.3.4 TYPES OF TRANSDUCTION

Transduction is common in both lytic and lysogenic cycles. It has two types, that is, generalized transduction in which any part of DNA can be taken by the phage and specialized transduction, in which phage carries only the specific part of DNA.

12.3.4.1 GENERALIZED TRANSDUCTION

In generalized transduction any random part of the bacterial DNA gets incorporated into the viral genome by mistake. It can occur by both lytic and lysogenic cycles. The generalized transduction occurs at the lytic stage of phage. It is the packing of bacterial DNA in viral envelope. It takes place in two ways, either headful packing or recombination. If the phage undergoes lytic cycle, the virus infects the cell and starts replicating its DNA but if by mistake, the bacterial DNA gets incorporated into the viral capsid (for encapsulation of viral DNA) it leads to generalized transduction (Thierauf and Perez, 2009).

If virus uses headful packaging, it fills the nucleo-capsid with the genetic material. The viral packaging may insert the bacterial DNA into the new virion. When this new infects other bacteria, the genetic material of host bacteria gets recombined with the recipient DNA. Transduction of *E. coli* by P1 phage is an example of generalized transduction. Generalized transduction is used to study gene mapping, mutagenesis, and linkage study.

12.3.4.2 SPECIALIZED TRANSDUCTION

In specialized transduction, only specific part of the bacterial DNA gets incorporated into the virus. Specialized transduction occurs through lysogenic cycle. When the prophage (viral DNA incorporated into the bacterial genome by lysogenic cycle) gets excised from the bacterial DNA, some parts of the bacterial DNA flanking the viral genome are also cleaved off. The newly packaged virus particle contains recombinant DNA from virus as well as bacteria. When this virus infects another recipient cell, it transfers the recombinant genetic material into it. The recipient cell contains DNA from both viral and host bacterial genome. Transduction of *E. coli* by $\boldsymbol{\lambda}$ phage is an example of specialized transduction which was discovered by Esther Lederberg. The genes that get transferred (donor genes) depend on where the phage genome is located on the chromosome (Zahler, 1982).

12.3.5 LATERAL TRANSDUCTION

Another type of transduction known as lateral transduction involves very long fragments of DNA to be transferred. It has the capacity to transfer more genes at greater frequency as compared to specialized and generalized transduction.

The lateral transduction has only been reported in *Staphylococcus aureus*. In this process, the prophage replicates itself in situ, in turn replicating the nearby bacterial DNA before excision. After replication, the packaging of the phage DNA and adjacent bacterial DNA occurs in situ. Successive packing events take place leading to packaging of many kilo-base of bacterial genome packaged in the viral particles which are then transferred to new bacterial strains (Chen et al., 2018).

If the bacterial DNA transferred in the phage particles is enough for homologous recombination, it will get inserted into the recipient genome. Multiple copies of phage DNA are produced during in situ replication, so some of the phage DNA excise normally (instead of in situ packaging) producing normal infectious viral particles.

12.3.6 APPLICATIONS OF TRANSDUCTION

Transduction is important for genetic engineering. It is an important tool for inserting or modifying genes in animal or plant cells for production of desired characteristics. Transduction can be used for gene therapy. It is important for the molecular biology research. It can be used to treat genetic diseases by modification of genetic error.

12.4 TRANSFORMATION

Transformation is the one of the three methods of HGT, in which external genetic material enters from one bacterium to another. Other two are conjugation (genetic transfer through direct contact between cells) and transduction (transfer through a viral vector). Transformation is the process in which naked DNA present in the environment is taken up by the recipient cells, which usually results in the expression of the newly acquired gene in the recipient cell. The cells that uptake the free DNA are known as transformed cells. The transformation process does not require living host cells but only free DNA in the surrounding environment.

During extreme conditions, some bacteria normally release their DNA in the environment to be absorbed by other cells. The recipient cells' or competent cells' capacity of acquiring the foreign DNA also depends on the environment. Transformation is needed to transfer gene of interest in the cells and this process can transfer 10–1000 of kilo bases of DNA.

12.4.1 TRANSFORMATION PRINCIPLE

Transformation depends on the natural ability of bacteria to release the DNA which is taken up by other competent cells. The transformation success depends on competency of the cells (ability to incorporate DNA). Several bacteria, like *E. coli*, can be made competent artificially (by calcium or electroporation) to increase their transforming capacity. The heat shock or electric shock increases the cell membrane permeability which allows DNA entry. The free DNA that enters the recipient cell can be incorporated into the bacterial chromosome and become a part of the genome or it can be inserted into the plasmid and replicate autonomously. Usually, a selectable marker is present on the plasmid used to select the transformed cells from non-transformed cells.

In artificial transformation, the gene of interest is first inserted into the plasmid. This plasmid is used to transform the competent host cells. After transformation, the bacterial cells containing the foreign plasmid can be detected by growth media supplemented with specific antibiotic.

12.4.2 HISTORY

Bacterial transformation was discovered by Frederick Griffith in 1928. He was interested in knowing whether heat-killed strains of bacteria were able to act as vaccine for mice against pneumonia. He discovered that the nonvirulent *Streptococcus pneumoniae* strains could be made virulent by exposing them to the heat-killed virulent strains. He discovered that some "transforming principle" was involved in making the harmless strains virulent. In 1944, Colin MacLeod, Oswald Avery, and Maclyn McCarty identified this transforming principle as "genetic." They isolated the DNA from virulent *S. pneumoniae* strains and injected it in the harmless strains. They were able to make the harmless strains virulent and named this phenomenon of DNA uptake as transformation (Fig. 12.3). The results of this experiment were not accepted by the scientific society until the discovery of other genetic transfer methods, transduction (1952) and conjugation (1946).

It was originally thought that *E. coli* was not able to transform. But in 1970, Akiko Higa and Morton Mandel showed that *E. coli* can uptake the DNA without the use of helper phage after treating it with calcium chloride. Mandel and Higa's method of transformation was later improved by Douglas Hanahan. Introduction of artificial competency in *E. coli* opened wide area

of research for transforming bacteria for molecular cloning research and this procedure is routinely used in laboratories.



FIGURE 12.3 Bacterial transformation. Uptake of naked DNA by the recipient cell. ⏎

Source: Adapted from Dubnau & Blokesch, 2019.

In 1980, the electroporation method was developed for transformation protocols. Later on, transformation of animal was developed when growth hormone was injected in the mouse embryo in 1982. In 1970, tumor-inducing capacity of *Agrobacterium tumefaciens* was discovered. It was discovered that the Ti plasmid was involved in inducing tumor in plants. These genes were removed from the plasmid to introduce novel genes that can be transferred in the plants by infecting them with the *Agrobacterium tumefaciens*. Electroporation, microinjection, and particle bombardment methods were developed for transformation in plant cells.

12.4.3 WHY DO BACTERIA TAKE UP DNA AND WHERE DOES IT COME FROM?

Three models can explain for the evolution of DNA uptake systems. Bacteria acquire novel genetic information, gene functions, virulent strains, or antibiotic resistance for their survival. Bacterial DNA under damaged conditions can uptake the surrounding foreign DNA for their repair. Carbon, nitrogen, and phosphorous from DNA can be taken up by bacteria to fulfill the food requirements. Free DNA from dead organism is abundant in the environment. Viable organism can also secrete DNA; some strains of *Neisseria gonorrhoeae* use a type IV secretion system to release DNA into the medium.

12.4.4 NATURAL COMPETENCE AND TRANSFORMATION

In nature, the bacteria carry sets of gene responsible to bring the DNA in close vicinity of the cell membrane. The incorporation of the foreign DNA into the cell requires DNA translocase complex, type IV pili and type II secretion

system assembly at the cytoplasmic membrane. There is some difference in the mechanism of DNA uptake in Gram-positive and Gram-negative bacteria due to the difference in the cell wall structure of both the groups.

The DNA binds to the receptors on the surface of the competent cells and crosses the membrane through DNA translocase. Single-stranded DNA can pass the membrane and the other strand is degraded by the nucleases in this process. This single-stranded DNA can be integrated into the bacterial genome by RecA. Due to the presence of an extra membrane, DNA transfer in Gram-negative bacteria requires the presence of channels formed by secretin on the outer membrane.

In general, natural transformation is an energy requiring complex process. The process of competence development in *Bacillus subtilis* requires 40 different genes and the transferred DNA length is greater than 1271 kb. The ability for natural transformation seems to occur in a number of prokaryotes, and so far 67 prokaryotic species (in seven different phyla) are known to undergo this process. Transformation in *Haemophilus influenzae* occurs most efficiently at the end of exponential growth as bacterial growth approaches stationary phase. Transformation in *Streptococcus mutants*, as well as in many other streptococci, occurs at high cell density and is associated with biofilm formation. Competence in *Bacillus subtilis* is induced toward the end of logarithmic growth, especially under conditions of amino acid limitation.

12.4.5 TRANSFORMATION OF GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIA

The best studied species among the transformable gram-positive bacteria are *Streptococcus pneumoniae*, *Streptococcus sanguinis*, and *B. subtilis*. Genetic transformation has been reported in a number of genera of gram-negative bacteria, including *Haemophilus*, *Neisseria*, *Moraxella*, and *Acinetobacter*. Of these, transformation in the genus *Haemophilus* (and in particular, *H. influenzae* and *H. parainfluenzae*) is by far the best studied (Smith et al., 1981).

12.4.6 STEPS OF TRANSFORMATION

12.4.6.1 COMPETENCE

Competence is the ability of a cell to uptake the DNA from the environment in a DNase-free state. Most bacteria become transformable under some specific growth conditions; so competence is an induced state. In

Gram-positive bacteria, for example if *Streptococcus* strains are grown in the suitable growth medium, the culture becomes competent at a given time and concentration (10^7–10^8 cells/mL) (Hotchkiss, 1954). This is due to the secretion of a factor called competence factor (CF) or activator. It releases continuously in the cell culture and increases when the cell density increases. So it is autocatalytic. CF binds to the surface of cell receptors and induces competence. In Gram-negative bacteria, competence development does not depend on external factors or proteases and it is regulated internally. The conditions for competence development vary between species. For example, in *H. influenzae*, conditions for competence development require end of cell division and continued protein synthesis. When exponentially growing cells were transferred to nongrowth media, the rate of competence development was 100% (Smith et al., 1981).

12.4.6.2 BINDING OF DNA

Binding of DNA to the cell can be reversible or irreversible. In reversible binding, the bound DNA can be easily removed undamaged by elution. It can occur in two ways: the DNA can interact with the non-transformation-related elements, that is, cell envelops or it can bind to the transformation-related elements (this binding can become irreversible, however the DNA remains exposed and sensitive to DNases) (Seto and Tomasz, 1974). In Gram-positive bacteria, that is, competent pneumococcal cells bind the DNA but do not convert in a DNase-resistant form in the presence of EDTA. On the other hand, the mutant pneumococcus strain lacking the membrane bound endonucleases binds DNA efficiently by are blocked for DNA uptake.

The DNA uptake by the competent cells depends on various factors, including pH, temperature, and optimal growth conditions. Transformation normally requires 100 mM Na^+ and K^+ and 1mM Ca^{2+} for DNA uptake. For example, competent *H. influenzae* cells bind single-stranded DNA under low pH (4.5–4.8) and 1M EDTA (Scocca et al., 1974).

12.4.6.3 UPTAKE OF DNA

DNA uptake requires conversion of DNA into a form that cannot be eluted in a DNase-resistant manner. This absorption of DNA requires tightly bound DNA to the cell surface and removal of DNA from external environment. This condition could be met if the DNA becomes enclosed in the

cell compartment, that is, vesicles, cytoplasmic, or periplasmic space. In Gram-positive bacteria, that is, the pneumococcal species, DNA uptake is completed when double-stranded DNA is nicked from both the strands. The first nick is produced during the binding of DNA and the second nick occurs by the help of membrane-bound endonucleases present at the cell surface (Lacks et al., 1975). In Gram-negative bacteria, there is a specificity of DNA uptake. Only homologous DNA is taken up by the cell. In *Haemophilus* and *Neisseria* species, a large amount of DNA is bound to the competent cells, but only DNA from closely related species is absorbed by the cell (Sparling, 1966).

12.4.6.4 ECLIPSE FORMATION

Eclipse formation is seen only in Gram-positive bacteria. When the DNA is taken up by the cells in single-stranded form and used to transform other cells, its transformation capacity is lost for some time. And the DNA is said to be in an eclipse form, this activity is regained when the DNA is reverted back to duplex state during integration (Morrison et al., 1983). In pneumococcal cells, the eclipse complex is formed between the single-stranded DNA and a 19.5 kDa protein that is formed in large quantity during competence development (Morrison and Baker, 1979). This complex is quite stable and insensitive to the ribonucleases. Susceptibility of this DNA to digestion by nucleases is 50–1000 fold less than the naked DNA. In Gram-negative bacteria (*Haemophilus* and *Neisseria* species), the DNA is taken up by the cells in a duplex form during transformation, so there is no eclipse formation.

12.4.6.5 INTEGRATION

The single-stranded foreign donor DNA is integrated into the recipient chromosome by recombination events. The integration mechanism works efficiently and nonselectively on any homologous DNA taken up by the cells. Studies have shown that the single-stranded donor DNA displaces a homologous strand of the recipient DNA to form a hetero-duplex (Smith et al., 1981). In *B. subtilis,* the hetero-duplex was visualized under an electron microscope. The process of integration can be separated into 4 stages on the basis of mutational analysis, first is eclipse complex translocation into the close proximity of recipient chromosome, then formation of unstable donor

recipient complex, formation of stable donor recipient complex, and finally the covalently closed complex formation.

During the integration process, one strand of donor DNA is integrated into the host genome while the other strand is degraded and released in the environment. The displaced strand of host is also degraded, but the exact mechanism of displacement is not known. However, the unintegrated strand in Gram-negative bacteria does not lose the transforming activity (Darlington and Bodmer, 1968).

12.5 CONCLUSION

The contents of bacterial genomes are affected by numerous processes, including "sex," which, as previously mentioned, can be defined in bacteria as the inheritance of DNA from any source aside from their one parent cell. Taken together, the features and consequences of sex in bacteria need not overlap with any of those that we normally associate with sex. Unlike the sexual process of higher eukaryotes, bacterial sex does not require the formation or fusion of gametes, cell contact, recombination, or even the creation of genetic variation or new combinations of chromosomal genes, and it is not associated with reproduction. Sex in bacteria is simply the uptake of any genetic material that might eventually be vertically or horizontally transmitted, and now, with the recent availability of complete genome sequences, we have begun to understand its pervasive role in the adaptation and diversification of bacterial lineages.

KEYWORDS

- **sexual reproduction**
- **recombination**
- **conjugation**
- **hereditary**
- **transformation**
- **transduction**
- **genetic material**

REFERENCES

Alvarez-Martinez, C. E.; Christie, P. J. Biological Diversity of Prokaryotic Type IV Secretion Systems. *Microbiol. Mol. Biol. Rev.* **2009,** *73*, 775–808.

Aminov, R. I. Horizontal Gene Exchange in Environmental Microbiota. *Front. Microbiol.* **2011,** *2*, 158.

Bundock, P.; Hooykaas, P. J. Integration of Agrobacterium Tumefaciens T-DNA in the Saccharomyces Cerevisiae Genome by Illegitimate Recombination. *Proc. Natl. Acad. Sci.* **1996**, *93*, 15272–15275.

Chen, J.; Quiles-Puchalt, N.; Chiang, Y. N.; Bacigalupe, R.; Fillol-Salom, A.; Chee, M. S. J.; Fitzgerald, J. R.; Penadés, J. R. Genome Hypermobility by Lateral Transduction. *Science* **2018**, *362*, 207–212.

Clewell, D. B. In *Bacterial Conjugation*; Springer Science & Business Media, 2013.

Curtiss III, R. Bacterial Conjugation. *Ann. Rev. Microbiol.* **1969,** *23*, 69–136.

Darlington, A. J.; Bodmer, W. F. Events Occurring at the Site of Integration of a DNA Molecule in Bacillus Subtilis Transformation. *Genetics* **1968,** *60*, 681.

Davison, J. Genetic Exchange between Bacteria in the Environment. *Plasmid* **1999,** *42*, 73–91.

Deng, Y.; Zhang, X.; Zhang, X. Recent Advances in Genetic Modification Systems for Actinobacteria. *Appl. Microbiol. Biotechnol.* **2017,** *101*, 2217–2226. DOI: 10.1007/s00253-017-8156-1

Dubnau, D.; Blokesch, M. Mechanisms of DNA Uptake by Naturally Competent Bacteria. *Ann. Rev. Genet.* **2019,** *53*, 217–237.

Gebhard, F.; Smalla, K. Transformation of Acinetobacter sp. Strain BD413 by Transgenic Sugar Beet DNA. *Appl. Environ. Microbiol.* **1998,** *64*, 1550–1554.

Harper, S. T. Abedon, B. H. Burrowes, & M. L. McConville (Eds.), *Bacteriophages: Biology, Technology, Therapy* (pp. 151–192). Cham: Springer International Publishing.

Heinemann, J. A.; Sprague, G. F. Bacterial Conjugative Plasmids Mobilize DNA Transfer Between Bacteria and Yeast. *Nature* **1989,** *340*, 205–209.

Hotchkiss, R. D. Cyclical Behavior in Pneumococcal Growth and Transformability Occasioned by Environmental Changes. *Proc. Natl. Acad. Sci. U. S. A.* **1954,** *40*, 49.

Lacks, S.; Greenberg, B.; Carlson, K. Fate of Donor DNA in Pneumococcal Transformation. *J. Mol. Biol.* **1967,** *29*, 327–347.

Lacks, S.; Greenberg, B.; Neuberger, M. Identification of a Deoxyribonuclease Implicated in Genetic Transformation of Diplococcus Pneumoniae. *J. Bacteriol.* **1975,** *123*, 222–232.

Lawley, T.; Wilkins, B. M.; Frost, L. S. Bacterial Conjugation in Gram-Negative Bacteria. *Plasmid Biol.* **2004,** 203–226.

Lawrence, J. G. Gene Transfer in Bacteria: Speciation without Species? *Theor. Popul. Biol.* **2002,** *61*, 449–460.

Lawrence, J. G.; Ochman, H. Molecular Archaeology of the Escherichia coli Genome. *Proc. Natl. Acad. Sci.* **1998,** *95*, 9413–9417.

Lederberg, J. Recombination Mechanisms in Bacteria. *J. Cell. Comp. Physiol.* **1955,** *45*, 75–107.

Lederberg, J. Genetic Transduction. *Am. Sci.* **1956,** *44*, 264–280.

Lederberg, E. M.; Lederberg, J. Genetic Studies of Lysogenicity in Escherichia coli. *Genetics* **1953,** *38*, 51.

Llosa, M.; Gomis-Rüth, F. X.; Coll, M.; Cruz, Fdl. Bacterial Conjugation: A Two-Step Mechanism for DNA Transport. *Mol. Microbiol.* **2002,** *45*, 1–8.

Mazel, D. Bacterial Genetics and Antibiotic Resistance Dissemination. *Antibiotic Development and Resistance*, 2001; pp 91–105.

Morrison, D.; Baker, M. Competence for Genetic Transformation in Pneumococcus Depends on Synthesis of a Small Set of Proteins. *Nature* **1979,** *282*, 215–217.

Morrison, D. A.; Lacks, S.; Guild, W.; Hageman, J. Isolation and Characterization of Three New Classes of Transformation-Deficient Mutants of Streptococcus pneumoniae that are Defective in DNA Transport and Genetic Recombination. *J. Bacteriol.* **1983,** *156*, 281–290.

Morse, M.; Lederberg, E. M.; Lederberg, J. Transduction in Escherichia coli K-12. *Genetics* **1956,** *41*, 142.

Ochman, H.; Groisman, E. The Origin and Evolution of Species Differences in Escherichia coli and Salmonella Typhimurium. In *Molecular Ecology and Evolution: Approaches and Applications*; Springer, 1994; pp 479–493.

Perwez, T. The MobB Protein of Broad-Host-Range Plasmid R1162 has Two Distinct Functions: Stabilization of Relaxosome and Enhancement of Transfer, The University of Texas at Austin, 1998.

Schneider, C. L. Bacteriophage-Mediated Horizontal Gene Transfer: Transduction. In *Bacteriophages: Biology, Technology, Therapy*, 2021; pp 151–192.

Scocca, J.; Poland, R.; Zoon, K. C. Specificity in Deoxyribonucleic Acid Uptake by Transformable Haemophilus Influenzae. *J. Bacteriol.* **1974,** *118*, 369–373.

Seto, H.; Tomasz, A. Early Stages in DNA Binding and Uptake During Genetic Transformation of Pneumococci. *Proc. Natl. Acad. Sci.* **1974,** *71*, 1493–1498.

Skaar, P. D.; Richter, A.; Lederberg, J. Correlated Selection for Motility and Sex-Incompatibility in Escherichia coli K12. *Proc. Natl. Acad. Sci. U. S. A.* **1957,** *43*, 329.

Smith, H. O.; Danner, D. B.; Deich, R. A. Genetic Transformation. *Ann. Rev. Biochem.* **1981,** *50*, 41–68.

Sparling, P. F. Genetic Transformation of Neisseria Gonorrhoeae to Streptomycin Resistance. *J. Bacteriol.* **1966,** *92*, 1364–1371.

Tatum, E.; Lederberg, J. Gene Recombination in the Bacterium Escherichia coli. *J. Bacteriol.* **1947,** *53*, 673–684.

Thierauf, A.; Perez, G. Generalized Transduction. In *Bacteriophages*; Springer, 2009; pp 267–286.

Trevors, J. Molecular Evolution in Bacteria: Surfaces, Cathodes and Anodes. *Antonie van Leeuwenhoek* **1997,** *71*, 363–368.

van Veen, J. A.; van Overbeek, L. S.; van Elsas, J. D. Fate and Activity of Microorganisms Introduced into Soil. *Microbiol. Mol. Biol. Rev.* **1997,** *61*, 121–135.

Wilkins, B.; Lanka, E.; Clewell, D. In *Bacterial Conjugation*, 1993.

Wilson, M.; Lindow, S. Release of Recombinant Microorganisms. *Ann. Rev. Microbiol.* **1993,** *47*, 913–944.

Zahler, S. A. *Specialized Transduction in Bacillus subtilis*; Academic Press, New York, 1982.

Zupan, J. R.; Zambryski, P. Transfer of T-DNA from Agrobacterium to the Plant Cell. *Plant Physiol.* **1995,** *107*, 1041.



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Index

A

Agriculture
  biotechnology benefits, 31–32
    benefits, 32
    limitations, 33
Alleles
  frequency, 203–204
Amplification, 46
Analysis techniques
  DNA libraries, construction and screening
    gel quantification method, 64–65
    PCR, 65–66
    southern blotting, 66
Apoptosis, 137–138
  importance, 138
  senescence, 138–139
Arboreal habitat, 208
  Hardy–Weinberg theorem (HWT), 209
  mutation, 209–210
  natural population growth, 210
  selection, 210
Artificial chromosomes
  bacterial artificial chromosomes (BACs), 115–116
  P1-derived artificial chromosome (PAC), 115
  yeast artificial chromosomes (YACs), 114–115
Artificial restriction enzymes, 80
Artificial selection, 224–225
Assortative mating, 217–218
Autosomal loci, 216

B

Bacterial artificial chromosomes (BACs), 95–96, 115–116
Bacterial conjugation
  fertile male method, 237
  historical aspects, 236–237
  stages of, 237–239
    cell-to-cell attachment, 238
    DNA transfer, 239
    donor control, 239
    relaxosome formation, 239
  sterile male method, 237
Biotechnology, 29
  advances, 31
  agriculture, 31–32
    benefits, 32
    limitations, 33
  benefits
    efficiency, 32–33
    increased crop yields, 32
    reduced losses, 32
  definition, 30
  diagnosis, 35
  history, 31
  limitations
    ethical, 33
    high cost, 33
    lack of regulation, 33–34
  medicine
    antibody engineering, 34
    gene therapy, 34
    personalized, 35
  transgenic animals, 36–37

C

Cell exclusion, 47
Cell-to-cell attachment, 238
Centimorgan (cM), 144, 145
Cloning vectors, 53–59
  properties, 110–111
  types, 111
Conformation-sensitive capillary electrophoresis (CSCE), 169–170
Conjugation
  plasmids, 240–241
  requirements for, 239–240
Crossing over, 147
  class I and II, 155–156
  molecular mechanism, 156–157
  sex chromosomes, 157
Crossover frequency, 135
CSGE (conformation-sensitive gel electrophoresis), 169

D

dCMP hydroxymethylase and DNA, 150
DNA libraries, construction and screening, 51
  analysis techniques
    gel quantification method, 64–65
    PCR, 65–66
    southern blotting, 66
  cloning vectors, 53–59
  DNA, 63–64
  extraction, 64
  molecular analysis, 63
  protein, 68–69
    analysis, 69–70
    gel electrophoresis, 70–71
    western blotting, 71
  RNA, 67
    northern blotting, 67–68
    SEQ, 67
  screening methods, 59–63
DNA mutations, 163
DNA technology, 41
  gene therapy, 43
  methods, 44–45
  recombinant antibodies, 43–44
  role of vectors, 43
  steps in formation
    amplification, 46
    cell exclusion, 47
    cutting, 45–46
    inclusion, 47
    isolation, 45
    joining, 46–47
Dolly breakthrough, 5
  abnormalities, 9–10
  animal model, 7
  benefit, 6
  improved, 8–9
  xenotransplantation, 7–8
Donor control, 239
Double-stranded breaks (DSBs), 152
Drosophila melanogaster, 162, 163

F

Fertile male method, 237
Forces changing gene
  dispersive process, 220
  migration, 220–221
  mutation, 221–223
  systematic process, 219
Frameshift mutations, 193

G

Gel electrophoresis, 70–71
Gel quantification method, 64–65
Gene and cell division, 119
  apoptosis
    importance, 138
    senescence, 138–139
  division, 129
    apoptosis, 137–138
    crossover frequency, 135
    gene mapping, 136
    maximum mutation probability, 139
    Morgan's unit, 136
    phases, 129–130
    recombination frequency, 135
    techniques, 136–137
  human genome arrangement, 121
    distribution, 122
    exon characteristic, 123
    number, 121–122
  meiosis
    process, 133–134
  methods, 124–126
  mutation, 126–127
    epigenetic changes, 127–129
  phases
    interphase, 130
    meiosis, 132–133
    mitosis, 130–132
  project, 123, 124
    impact, 124
    methodology, 123
    results, 124
  understanding, 120
Gene mapping, 136, 145
  chromosomes assort, genes, 146
  crossing over, 147
    class I and II, 155–156
    molecular mechanism, 156–157
    sex chromosomes, 157
  homologous chromosomes, 147
  linkage analysis
    meiosis indicators, 151

parametric linkage analysis, 150–151
linkage maps, 147, 148
dCMP hydroxymethylase and DNA, 150
restrict fragment period polymorphisms (RFLPs), 149
Thomas Morgan, 149
recombination
double-stranded breaks (DSBs), 152
gene conversion, phenomenon, 152
interchromosomal recombination, 151, 152
recombinational repair, 153–154
RNA *virus* recombination, 153
synapsis, 152
recombination frequency, 147, 148
dCMP hydroxymethylase and DNA, 150
restrict fragment period polymorphisms (RFLPs), 149
Thomas Morgan, 149
Gene mutations
effects of, 194
mutation frequency, 195
single-nucleotide polymorphism, 195
sublethal, 195
methods, 186
base substitution, 187
deletion, 187
insertions, 187–188
sources of
analogues, 190
chemical mutagens, 189–190
intercalating agents, 192
modifying agents, 191–192
physical mutagens, 188–189
spontaneous and induced, 187
types
frameshift mutations, 193
missense mutations, 193–194
nonsense mutations, 194
silent or synonymous mutations, 192–193
Genetic engineering, 1
applications, 13–14
cancer treatment, 15–16
crop production, 17–18
gene therapy, 16–17
mediated iPSC models, 16
produce, 17
biotechnological, 2–3
cloning, 4
CAS system, 12–13
cloning, 10–11
genome editing, 11–12
dolly breakthrough, 5
abnormalities, 9–10
animal model, 7
benefit, 6
improved, 8–9
xenotransplantation, 7–8
ethical issues, 20–21
human health, 10
humanitarian, 19–20
mechanism, 4–5
perilous, 18–19
somatic cell nuclear transfer (SCNT), 4
Genetic engineering tools, 75
classification
type I, 79
type II, 79
type III, 79–80
cleavage
applications, 82
components, 77–78
restriction enzymes, 78
vectors used, 82
nomenclature
pattern, 78
plasmids
types, 84–89
principles, 77
restriction enzymes
artificial restriction enzymes, 80
classification, 79
cleavage, 80–82
nomenclature, 78
type II
palindromic sequences, 79
types
Ti plasmid, 89–90
vectors, 82–83
bacterial artificial chromosomes (BACS), 95–96
bacteriophages, 90–93
cloning vector, 83
cosmids, 94–95
phagemids, 94

plasmids, 84
yeast artificial chromosomes (YACS), 97–98
Genetic linkage, 144
Mendel's second law, 145
Genotype frequency, 204–205
dispersive process, 220
migration, 220–221
mutation, 221–223
systematic process, 219
Gram-negative bacteria, 248
Gram-positive bacteria, 248

H

Hardy-Weinberg theorem (HWT), 209, 210–212
applications, 213
assortative mating, 217–218
assumptions, 212
autosomal loci, 216
connected loci, 216
disassortative mating, 218
equation, 199
genotype and gene frequencies, 213
heterozygotes, 214
inbreeding coefficient, 219
multiple allelic systems, 215–216
nonrandom mating, 217
recessive and dominant genes, 214
test for agreement, 215
Homologous chromosomes, 147
Horizontal gene transfer (HGT), 234, 235
HRM (high-resolution melting analysis), 170–171
Human genome arrangement, 121
distribution, 122
exon characteristic, 123
number, 121–122

I

Illumina genome analyzer, 174
Increased crop yields, 32

L

Linear energy transmission (LET), 165
Linkage analysis
meiosis indicators, 151
parametric linkage analysis, 150–151

M

Maximum mutation probability, 139
Mediated iPSC models, 16
Medicine
antibody engineering, 34
gene therapy, 34
personalized, 35
Meiosis
process, 133–134
Missense mutations, 193–194
Morgan's unit, 136
Mutant chromosome sequence (MutChromSeq), 175–176
Mutation detection, 166
cleavage analysis, 171
conformation-sensitive capillary electrophoresis (CSCE), 169–170
CSGE (conformation-sensitive gel electrophoresis), 169
dot blot SNP analysis, 171
exome capture, 176–177
HRM (high-resolution melting analysis), 170–171
Illumina Genome Analyzer, 174
Mutant Chromosome Sequence (MutChromSeq), 175–176
MutMap technology, 174–175
next-generation sequencing, 173–174
quantitative trait loci (QTLs), 176
single-strand conformation polymorphism (SSCP), 167
targeting induced local lesions in genomes (TILLING), 171–172
temperature gradient capillary electrophoresis (TGCE), 169
temperature gradient gel electrophoresis (TGGE), 168–169
whole genome sequencing (WGS), 172–173
Mutation frequencies, 164, 166
linear energy transmission (LET), 165
plant tissue, 166
radiation frequency, 166
Mutations
defined, 185

N

Natural selection, 223
Nonsense mutations, 194

P

P1-derived artificial chromosome (PAC), 115
Plants population
  genetics and evolution of, 206–207
  mobility of genes, 207–208
  organism, mating process, 208
Population genetics
  alleles
    frequency, 203–204
  arboreal habitat, 208
    Hardy–Weinberg theorem (HWT), 209
    mutation, 209–210
    natural population growth, 210
    selection, 210
  forces changing gene
    dispersive process, 220
    migration, 220–221
    mutation, 221–223
    systematic process, 219
  genetic drift, 225–226
  genotype frequency, 204–205
    dispersive process, 220
    migration, 220–221
    mutation, 221–223
    systematic process, 219
  Hardy–Weinberg theorem (HWT), 210–212
    applications, 213
    assortative mating, 217–218
    assumptions, 212
    autosomal loci, 216
    connected loci, 216
    disassortative mating, 218
    genotype and gene frequencies, 213
    heterozygotes, 214
    inbreeding coefficient, 219
    multiple allelic systems, 215–216
    nonrandom mating, 217
    recessive and dominant genes, 214
    test for agreement, 215
  history of, 205–206
  origin of species, 200
    Darwin's theory, 201–202
    Mendel's theory, 202
  plants population
    genetics and evolution of, 206–207
    mobility of genes, 207–208
    organism, mating process, 208
  selection
    artificial selection, 224–225
    natural selection, 223
  speciation, 226
Protein, 68–69
  analysis, 69–70
  gel electrophoresis, 70–71
  western blotting, 71

Q

Quantitative trait loci (QTLs), 176

R

Reduced losses, 32
Restrict fragment period polymorphisms (RFLPs), 149
Restriction enzymes and vectors, 101
  artificial chromosomes
    bacterial artificial chromosomes (BACs), 115–116
    P1-derived artificial chromosome (PAC), 115
    yeast artificial chromosomes (YACs), 114–115
  cloning vectors
    properties, 110–111
    types, 111
  definition
    applications, 108
    nomenclature, 107
    recognition sequences, 105–106
    restriction modification system, 104
    types, 106
  historical background, 102
  ligase
    uses, 110
  tools, 102–103
    cloning vectors, 110
    definition, 103–104
    ligase, 109
  types
    artificial chromosomes, 114
    cosmids, 113
    phage vectors, 113
    phagemids, 114
    plasmids, 112
    type I, 106

type II, 106–107
type III, 107
type IV, 107
RNA *virus* recombination, 153

S

Screening technique, 166
cleavage analysis, 171
conformation-sensitive capillary electrophoresis (CSCE), 169–170
CSGE (conformation-sensitive gel electrophoresis), 169
dot blot SNP analysis, 171
exome capture, 176–177
HRM (high-resolution melting analysis), 170–171
Illumina Genome Analyzer, 174
Mutant chromosome sequence (MutChromSeq), 175–176
MutMap technology, 174–175
next-generation sequencing, 173–174
quantitative trait loci (QTLs), 176
single-strand conformation polymorphism (SSCP), 167
targeting induced local lesions in genomes (TILLING), 171–172
temperature gradient capillary electrophoresis (TGCE), 169
whole genome sequencing (WGS), 172–173
Sex chromosomes, 157
Sexuality in bacteria
bacterial conjugation
fertile male method, 237
historical aspects, 236–237
stages of, 237–239
sterile male method, 237
conjugation
plasmids, 240–241
requirements for, 239–240
transduction
applications of, 245
bacterial transduction, 243
DNA, recipient cell, 243
generalized transduction, 244
history, 241–243
lateral transduction, 244–245
specialized transduction, 244
transformation
binding of DNA, 249
competence, 248–249
DNA uptake, 249–250
eclipse formation, 250
gram-negative bacteria, 248
gram-positive bacteria, 248
history, 246–247
integration, 250–251
natural competence and, 247–248
principle, 246
Silent or synonymous mutations, 192–193
Single-strand conformation polymorphism (SSCP), 167
Southern blotting, 66
Speciation, 199
Spontaneous mutations
defined, 162
forward and reverse genetic approach, 163–164
mutation detection, 166
cleavage analysis, 171
conformation-sensitive capillary electrophoresis (CSCE), 169–170
CSGE (conformation-sensitive gel electrophoresis), 169
dot blot SNP analysis, 171
exome capture, 176–177
HRM (high-resolution melting analysis), 170–171
Illumina genome analyzer, 174
Mutant Chromosome Sequence (MutChromSeq), 175–176
MutMap technology, 174–175
next-generation sequencing, 173–174
quantitative trait loci (QTLs), 176
single-strand conformation polymorphism (SSCP), 167
targeting induced local lesions in genomes (TILLING), 171–172
temperature gradient capillary electrophoresis (TGCE), 169
temperature gradient gel electrophoresis (TGGE), 168–169
whole genome sequencing (WGS), 172–173
mutation frequencies, 164
linear energy transmission (LET), 165

plant tissue, 166
radiation frequency, 166
screening technique, 166
cleavage analysis, 171
conformation-sensitive capillary electrophoresis (CSCE), 169–170
CSGE (conformation-sensitive gel electrophoresis), 169
dot blot SNP analysis, 171
exome capture, 176–177
HRM (high-resolution melting analysis), 170–171
Illumina genome analyzer, 174
Mutant chromosome sequence (MutChromSeq), 175–176
MutMap technology, 174–175
next-generation sequencing, 173–174
quantitative trait loci (QTLs), 176
single-strand conformation polymorphism (SSCP), 167
targeting induced local lesions in genomes (TILLING), 171–172
temperature gradient capillary electrophoresis (TGCE), 169
whole genome sequencing (WGS), 172–173
Sterile male method, 237

T

Targeting induced local lesions in genomes (TILLING), 171–172
Temperature gradient capillary electrophoresis (TGCE), 169
Temperature gradient gel electrophoresis (TGGE), 168–169
Transduction
applications of, 245
bacterial transduction, 243
DNA, recipient cell, 243
generalized transduction, 244
history, 241–243
lateral transduction, 244–245
specialized transduction, 244
Transformation
binding of DNA, 249
competence, 248–249
DNA uptake, 249–250
eclipse formation, 250
gram-negative bacteria, 248
gram-positive bacteria, 248
history, 246–247
integration, 250–251
natural competence and, 247–248
principle, 246
Transgenic animals, 36–37

V

Vectors, 82–83
bacterial artificial chromosomes (BACS), 95–96
bacteriophages, 90–93
cloning vector, 83
cosmids, 94–95
phagemids, 94
plasmids, 84
yeast artificial chromosomes (YACS), 97–98

W

Western blotting, 71
Whole genome sequencing (WGS), 172–173

X

Xenotransplantation, 7–8

Y

Yeast artificial chromosomes (YACs), 97–98, 114–115