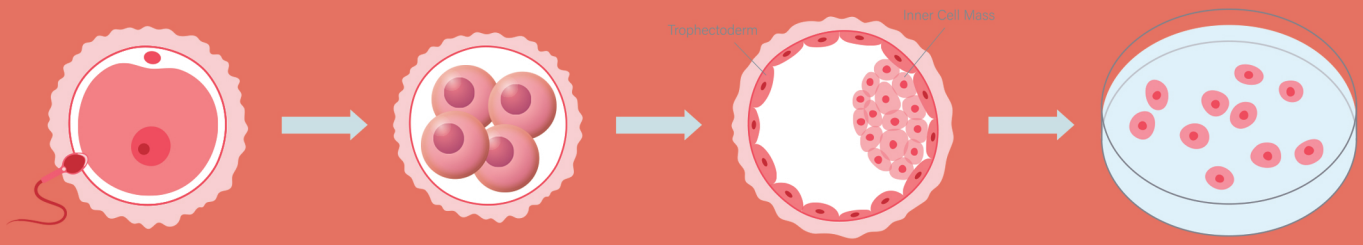
Omics Technologies for Clinical Diagnosis and Gene Therapy

Medical Applications in Human Genetics



Editors:
Syeda Marriam Bakhtiar
Erum Dilshad

Bentham Books

Omics Technologies for Clinical Diagnosis and Gene Therapy

Medical Applications in Human Genetics

Edited by

Syeda Marriam Bakhtiar
Department of Bioinformatics and Biosciences
Faculty of Health and Life Sciences
Capital University of Science and Technology
Islamabad
Pakistan

&

Erum Dilshad
Department of Bioinformatics and Biosciences
Faculty of Health and Life Sciences
Capital University of Science and Technology
Islamabad
Pakistan

Omics Technologies for Clinical Diagnosis and Gene Therapy

Editors: Syeda Marriam Bakhtiar and Erum Dilshad

ISBN (Online): 978-981-5079-51-7

ISBN (Print): 978-981-5079-52-4

ISBN (Paperback): 978-981-5079-53-1

First published in 2022.

need for a court order if at any point you breach any terms of this License Agreement. In no event will any delay or failure by Bentham Science Publishers in enforcing your compliance with this License Agreement constitute a waiver of any of its rights.

3. You acknowledge that you have read this License Agreement, and agree to be bound by its terms and conditions. To the extent that any other terms and conditions presented on any website of Bentham Science Publishers conflict with, or are inconsistent with, the terms and conditions set out in this License Agreement, you acknowledge that the terms and conditions set out in this License Agreement shall prevail.

Bentham Science Publishers Pte. Ltd.
80 Robinson Road #02-00
Singapore 068898
Singapore
Email: subscriptions@benthamscience.net

CONTENTS

FOREWORD

Human diseases especially genetic disabilities have always been a focus of research. Research on genetics and molecular genetics has contributed extensively to improving the overall quality of life and management of patients suffering from genetic diseases. With the onset of Omics and Next generation technologies (NGS), the dream of personalized medicine has almost come to reality. Technological advances in the domains of genomics, transcriptomics, proteomics and metabolomics have enabled scientists to explore the genetic and molecular causes in extraordinary detail. These technologies have contributed immensely to advancements related to early and efficient diagnosis, which have revolutionized clinical practices. Despite the contribution of these technologies, it is always felt that none of these technologies alone have the potential to cope with the biological complexity of human diseases. The integration of multiple technologies and a combination of diverse data types is the new approach that has the potential to provide a more comprehensive understanding of biological systems controlling the onset, progression and impact of diseases.

Initially, the focus of research has been on the early diagnosis and methods by which the symptoms of the disease could be eased off. Gene Therapy, personalized medicine, and precision medicine are very promising concepts but there have always been concerns about the access of the general public to these approaches as well as the application for diverse genetic diseases including rare and common diseases, multifactorial diseases including cancers. Omics Technologies have not only provided the scientists with better opportunities for correct diagnosis but also expanded the options for treatment including gene therapy, pharmacogenomics, single-cell omics, regenerative medicine, stem cell technologies and many more. Integrative approaches utilizing engineering and informatics have also widened the knowledge base required for appropriate treatment and management approaches.

This book compiled and edited by scientists working in various domains of genomics and human genetics will not only provide the researchers with new approaches in conventional methods of genetics-based diagnosis and counselling but will also open new avenues for further exploration of genetic causes and treatment options. This book is a great effort to document state-of-the-art techniques and technologies for disease prediction and early diagnosis to disease treatment and prognosis using integrative Omics.

Shahid Mahmood Baig
Department of Biological and Biomedical Sciences,
Agha Khan University, Karachi
Pakistan
Chairman, Pakistan Science Foundation

PREFACE

Human inherited disorders have been the focus of attention for a long time. Various books have been written with a focus on classical genetic approaches, clinical diagnostic strategies, and counselling, management strategies. With the emergence of Next-generation techniques, a new era was started and lots of developments have occurred in Human Genetics. The perspective has also been widened with emerging OMICS technologies. An integrated approach is being used not only for diagnosis but also for management and therapeutic purposes. This book is an effort to highlight and compile various emerging areas of OMICS technology and its application in the diagnosis and management of human genetic disorders.

The book is planned with three areas of research and implementation i.e., Diagnosis covering conventional strategies to next-generation platforms. This section focuses on the role of Insilco analysis, databases and multi-omics of single-cell which will help in designing better management strategies. Section II covers management and therapeutic interventions starting with genetic counselling and then including more specific techniques such as pharmacogenomics and personalized medicine, gene editing techniques and their applications in gene therapies and regenerative medicine. Section III focuses on case studies and discusses the applications and success of all the above-mentioned strategies on selected human disorders.

Syeda Marriam Bakhtiar
Department of Bioinformatics and Biosciences
Faculty of Health and Life Sciences
Capital University of Science and Technology
Islamabad
Pakistan

&

Erum Dilshad
Department of Bioinformatics and Biosciences
Faculty of Health and Life Sciences
Capital University of Science and Technology
Islamabad
Pakistan

List of Contributors

Adnan Haider	Department of Biological Sciences, National University of Medical Sciences, Rawalpindi, Pakistan
Alvina Gul	Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), Islamabad, Pakistan
Amina Basheer	Department of Industrial Biotechnology,Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), Islamabad, Pakistan
Amjad Ali	Department of Industrial Biotechnology,Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), Islamabad, Pakistan
Ammara Siddique	Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), Islamabad, Pakistan
Amna Naheed Khan	Department of Bioinformatics and Biosciences, Faculty of Health and Life Sciences, Capital University of Science and Technology (CUST), Islamabad, Pakistan
Anam Naz	Institute of Molecular Biology and Biotechnology (IMBB), The University of Lahore (UOL), Lahore, Pakistan
Aqsa Ikram	Institute of Molecular Biology and Biotechnology (IMBB), The University of Lahore (UOL), Lahore, Pakistan
Ambrin Fatima	Department of Biological and Biomedical Sciences, The Aga Khan University, Karachi, Pakistan
Areena Suhail Khan	Department of Biosciences, COMSATS University Islamabad, Park Road, Islamabad, Pakistan
Atif Ali Khan Khalil	Department of Biological Sciences, National University of Medical Sciences, Rawalpindi, Pakistan
Attiya Kanwal	International Islamic University Islamabad (IIUI), Islamabad, Pakistan
Ayaz Khan	National Institute for Biotechnology and Genetic Engineering College, Pakistan Institute of Engineering and Applied Sciences (NIBGE-C, PIEAS), Faisalabad, Pakistan
Aysha Saeed	Department of Biotechnology, Kinnaird College for Women, Lahore, Pakistan
Bibi Nazia Murtaza	Department of Zoology, Abbottabad University of Science and Technology, Abbottabad, Pakistan
Bisma Rauff	Department of Biomedical Engineering, University of Engineering and Technology (UET), Narowal Campus, Lahore, Pakistan
Bushra Bano	Institute of Basic Medical Sciences, Khyber Medical University, Peshawar, Pakistan
Erum Dilshad	Department of Bioinformatics and Biosciences, Faculty of Health and Life Sciences, Capital University of Science and Technology (CUST), Islamabad, Pakistan

Faiza Naseer	Shifa College of Pharmaceutical Sciences, Shifa Tameer-e-Millat University, Islamabad, Pakistan
Fakhra Nazir	Department of Bioinformatics and Biosciences, Capital University of Science and Technology, Islamabad (CUST, Islamabad), Pakistan Department of Biosciences, COMSATS University Islamabad, Sahiwal, Pakistan
Fatima Shahid	Department of Industrial Biotechnology, Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), Islamabad, Pakistan
Fazli Subhan	Department of Biological Sciences, National University of Medical Sciences, Rawalpindi, Pakistan
Hafiza Noor Ul Ayan	National Institute for Biotechnology and Genetic Engineering College, Pakistan Institute of Engineering and Applied Sciences (NIBGE-C, PIEAS), Faisalabad, Pakistan Institute for Cardiogenetic, University of Lubeck, Lubeck, Germany
Hajra Qayyum	Department of Biosciences and Bioinformatics, Capital University of Science and Technology, Islamabad, Pakistan
Hayeqa Shahwar Awan	Department of Industrial Biotechnology,Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), Islamabad, Pakistan
Hira Kazmi	Centre for Human Genetics, Hazara University Mansehra, Pakistan
Humna Masood	Institute of Biochemistry and Biotechnology, University of Veterinary and Animal Sciences-UVAS, Lahore, Pakistan
Huma Tariq	Department of Zoology, Hazara University, Manshera, Pakistan
Ilyas Ahmad	Institute for Cardiogenetic, University of Lubeck, Lubeck, Germany
Iram Anjum	Department of Biotechnology, Kinnaird College for Women, Lahore, Pakistan
Iqra Bashir	Department of Bioinformatics and Biosciences, Faculty of Health and Life Sciences, Capital University of Science and Technology (CUST), Islamabad, Pakistan
Komal Aslam	Department of Biotechnology, Kinnaird College for Women, Lahore, Pakistan Department of Biotechnology, Lahore College for Women University, Lahore, Pakistan
Mahnoor Ejaz	Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), Islamabad, Pakistan
Mahnoor Asif	National Institute for Biotechnology and Genetic Engineering College, Pakistan Institute of Engineering and Applied Sciences (NIBGE-C, PIEAS), Faisalabad, Pakistan
Maria Iqbal	National Institute for Biotechnology and Genetic Engineering College, Pakistan Institute of Engineering and Applied Sciences (NIBGE-C, PIEAS), Faisalabad, Pakistan Centogene GmbH, Rostock, Germany
Maria Shabbir	Healthcare Biotechnology, Attaur Rehman School of Applied Biosciences, National University of Science and Technology, Islamabad, Pakistan

Marriam Bakhtiar	Department of Bioinformatics and Biosciences, Faculty of Health and Life Sciences, Capital University of Science and Technology (CUST), Islamabad, Pakistan
Muhammad Faheem	Department of Biological Sciences, National University of Medical Sciences, Rawalpindi, Pakistan
Muhammad Ilyas	Centre for Omic Sciences, Islamia College University Peshawar, Pakistan Department of Bioengineering, University of Engineering and Applied Sciences, Swat, Pakistan
Muhammad Jawad Hassan	Department of Biological Sciences, National University of Medical Sciences, Rawalpindi, Pakistan
Muhammad Tariq	National Institute for Biotechnology and Genetic Engineering College, Pakistan Institute of Engineering and Applied Sciences (NIBGE-C, PIEAS), Faisalabad, Pakistan
Muhammad Maaz	Department of Bioinformatics and Biosciences, Faculty of Health and Life Sciences, Capital University of Science and Technology (CUST), Islamabad, Pakistan
Muhammad Saad Khan	Centre for Bioresource Research, Islamabad, Pakistan
Muneeba Ishtiaq	Department of Bioinformatics and Biosciences, Capital University of Science and Technology, Islamabad (CUST, Islamabad), Pakistan
Muhammad Naeem	Department of Biological Sciences, National University of Medical Sciences, Rawalpindi, Pakistan
Narjis Khatoon	Department of Bioinformatics and Biosciences, Capital University of Science and Technology, Islamabad (CUST, Islamabad), Pakistan
Naveed Altaf Malik	National Institute for Biotechnology and Genetic Engineering College, Pakistan Institute of Engineering and Applied Sciences (NIBGE-C, PIEAS), Faisalabad, Pakistan
Naveed Altaf Malik	National Institute for Biotechnology and Genetic Engineering College, Pakistan Institute of Engineering and Applied Sciences (NIBGE-C, PIEAS), Faisalabad, Pakistan
Rabbiah Manzoor Malik	Department of Bioinformatics and Biosciences, Capital University of Science and Technology, Islamabad (CUST, Islamabad), Pakistan Wah Medical College, WahCantt, Pakistan
Raees Khan	Institute of Basic Medical Sciences, Khyber Medical University, Peshawar, Pakistan
Syeda Seema Waseem	National Institute for Biotechnology and Genetic Engineering College, Pakistan Institute of Engineering and Applied Sciences (NIBGE-C, PIEAS), Faisalabad, Pakistan
Shahid Mahmood Baig	National Institute for Biotechnology and Genetic Engineering College, Pakistan Institute of Engineering and Applied Sciences (NIBGE-C, PIEAS), Faisalabad, Pakistan
Sadia Nawaz	Institute of Biochemistry and Biotechnology, University of Veterinary and Animal Sciences-UVAS, Lahore, Pakistan

Shumaila Zulfiqar	National Institute for Biotechnology and Genetic Engineering College, Pakistan Institute of Engineering and Applied Sciences (NIBGE-C, PIEAS), Faisalabad, Pakistan Department of Biotechnology, Kinnaird College for Women, Lahore, Pakistan
Shafaq Ramzan	National Institute for Biotechnology and Genetic Engineering College, Pakistan Institute of Engineering and Applied Sciences (NIBGE-C, PIEAS), Faisalabad, Pakistan
Shahid Mahmood Baig	National Institute for Biotechnology and Genetic Engineering College, Pakistan Institute of Engineering and Applied Sciences (NIBGE-C, PIEAS), Faisalabad, Pakistan
Syeda Marriam Bakhtiar	Department of Biosciences and Bioinformatics, Capital University of Science and Technology, Islamabad, Pakistan
Sabba Mehmood	Department of Biological Sciences, National University of Medical Sciences, Rawalpindi, Pakistan
Shifa Tariq Ashraf	Department of Industrial Biotechnology, Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), Islamabad, Pakistan
Sajjad Ahmed	Department of Zoology, Hazara University, Manshera, Pakistan
Shumaila Azam	Department of Bioinformatics and Biosciences, Capital University of Science and Technology, Islamabad, Pakistan
Sahar Fazal	Department of Bioinformatics and Biosciences, Capital University of Science and Technology, Islamabad, Pakistan
Sana Elahi	Department of Bioinformatics and Biosciences, Capital University of Science and Technology (CUST), Islamabad, Pakistan
Sarmad Mehmood	Department of Pathology, CMH Institute of Medical Sciences, Bahawalpur, Pakistan
Sajjad Haider	Department of Chemical Engineering, King Saud University, Riyadh, Saudi Arabia
Syed Babar Jamal	Department of Biological Sciences, National University of Medical Sciences, Rawalpindi, Pakistan
Uzma Abdullah	University Institute of Biochemistry and Biotechnology, PMAS-Arid Agriculture University, Rawalpindi, Pakistan
Zafar Ali	Centre for Biotechnology and Microbiology, University of Swat, Swat, Pakistan

CHAPTER 1

Next-Generation Technologies for Rare Inherited Disorders

Hira Kazmi[1] and **Muhammad Ilyas**[2,3,*]

[1] *Centre for Human Genetics, Hazara University Mansehra, Pakistan*

[2] *Centre for Omic Sciences, Islamia College University Peshawar, Pakistan*

[3] *Department of Bioengineering, University of Engineering and Applied Sciences, Swat, Pakistan*

Abstract: Rare inherited disorders have become a major public health concern in recent years. Owing to a lack of resources, poorly planned primary and basic health care, and inadequate political structures, treatment, and management policies are daunting challenges in many countries. As a result, these diseases need particular attention, especially in less developed areas, where these disorders remain unnoticed. Similarly, the effect of such severe disorders on underprivileged populations is expected to be devastating. Identifying certain genetic markers can provide a valuable explanation for disease etiology, molecular characterization, and pathogenesis. In this chapter, we highlight the importance of next-generation sequencing to explore and recognize the role of novel causative genes in developing successful treatments for the most prevalent rare genetic disorders. DNA methylation and transcriptome markers have been shown to aid in the prediction of common diseases; however, this has not been tested on rare genetic disorders. Since the rate of rare inherited disorders is higher in developing countries, we believe that these populations can provide us with much stronger clues for the genetic and environmental association. These markers, along with other parameters, can be used to systematically build machine learning models to improve risk prediction; this approach has the potential to reshape how we predict disease risk and save many lives around the world.

Keywords: Clinical genomics, Genetic counseling, Next-generation sequencing, Prenatal diagnosis, Rare inherited diseases.

1. INTRODUCTION

Rare disorders affect more than 300 million people worldwide, with a diagnosis of less than 0.2 million people [1]. Around 80% of these disorders are genetic in

* **Corresponding author Muhammad Ilyas**: Centre for Omic Sciences, Islamia College University Peshawar, Pakistan and Department of Bioengineering, University of Engineering and Applied Sciences, Swat, Pakistan;
E-mail: milyas@icp.edu.pk

Syeda Marriam Bakhtiar and Erum Dilshad (Eds.)

origin and have no treatment at all. Identification of such disease variants in patients can now be done with greater accuracy and lower cost by using whole genome/exome sequencing [2]. Despite this, researchers and policymakers are still grappling with the problems involving the use and interpretation of genotype data. Genetic variations have been used in molecular diagnostic research in the past, but with a few loci [3]. With lower cost, faster, and more precise sequencing technologies, it is easier to perform diagnostic tests at the single nucleotide level. Researchers around the world have created more sophisticated methods to study the role of variants and their associated environment in complex diseases using human genome data generated by the 1000 genome project and other genome research groups [4]. Clinical researchers are now using genome-wide studies to advance diagnostics and provide improved decision-making tools for patients. This is how genomics' impact on health care ushered in a new era of genetic medicine, also known as personalized medicine. It takes a reasonable amount of time and money to get results from a lab to a professional clinic. According to genetic specialists, it takes more than ten years for the pharmaceutical industry to conduct medical research based on FDA policies [5]. The genome-wide study contributes to the possibility of developing diseases that are widespread in the world's population. These common diseases include diabetes, hypertension, cancer, and cardiovascular diseases [6]. Comprehensive knowledge of the genetic structure of such disorders will help detect the vital mechanism of cells and, in the long run, improve our understanding of how various factors affect an individual's health.

Genetic studies have revolutionized many fields of research, with the total economic value of the human genome project estimated to be 796 billion USD [7]. Until recently, it was time-consuming and costly to carry out tests to detect pathogenic mutations. The recent boom in Next-Generation Sequencing (NGS) technologies has been a key to low-cost, fast, and reliable performance for molecular diagnostics. After the publication of the human genome, the main challenge for researchers working in the field of medical genetics has been to translate and use this mass of data in a clinical setting. However, genetic characterizations, which include transcriptomic and epigenomic studies of populations with unusual genetic disorders, are not yet properly investigated in the demographic and epidemiological studies of rare inherited diseases. Genes and variants that may be used as markers for the pre-diagnostic testing of such disorders should be identified. This will potentially benefit patients and families with such neglected and devastating disorders through pre-screening, genetic counseling, and carrier screening, and will provide a step toward fully personalized medicine and therapy [8].

1.1. Whole Genome/Exome Sequencing

Next-generation sequencing (NGS) technology has become the most groundbreaking research achievement in the science world. It refers to a series of modern DNA sequencing procedures which are making significant progress in the sequencing of millions of genomic fragments by employing particularly parallel reactions [9, 10]. The cost of sequencing a genome with NGS technology is cheaper as compared to the Sanger sequencing method [11]. The costs have plunged in the last few years, quickly exceeding Moore's Law, the standard benchmark for the declining cost of technology [12] (Fig. **1.1**). Numerous technologies, including cutting-edge chemistries, amplification methodologies, and efficient and high-resolution microscopy, have been redesigned to make this possible. In recent years, genome-wide studies using microarray technology have made significant progress [13]. Microarray chip approaches were initially used for gene expression analysis, but they later found widespread use in the analysis of copy number alterations, microRNA studies, mapping of binding sites for protein-protein and DNA-protein interactions, and genotyping of single nucleotide variants [14]. However, NGS technology has made significant advances and is expected to replace the majority of microchip platforms in the long term [15].

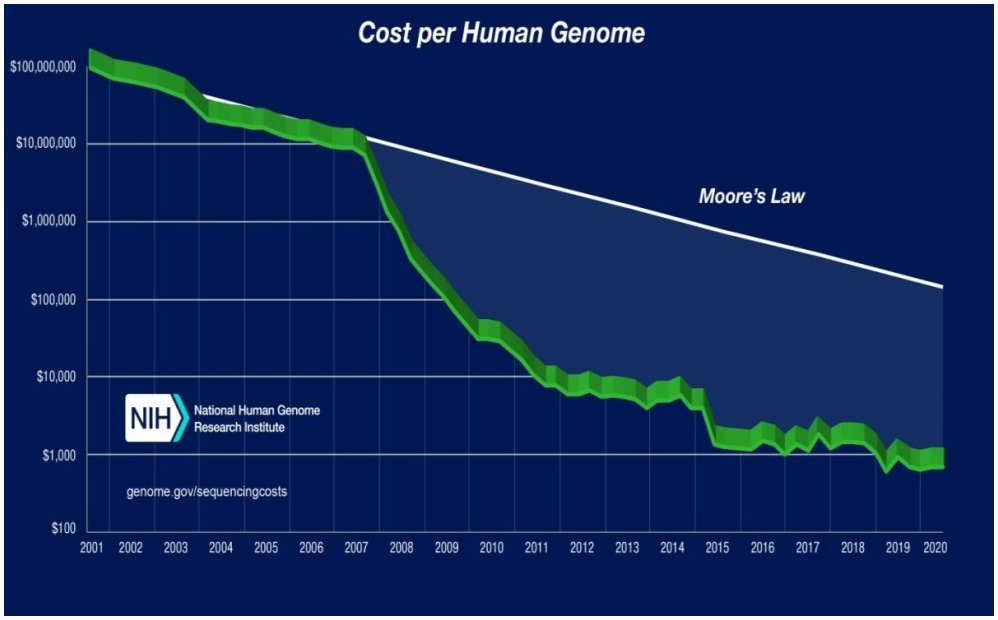


Fig. (1.1). Cost per genome data - August 2020. Data obtained from National Human Genome Research Institute (https://www.genome.gov).

CHAPTER 2

Genetic Testing for Rare Genetic Disorders

Muhammad Tariq[1,*], **Naveed Altaf Malik**[1], **Ilyas Ahmad**[2], **Syeda Seema Waseem** and **Shahid Mahmood Baig**[1]

[1] *National Institute for Biotechnology and Genetic Engineering College, Pakistan Institute of Engineering and Applied Sciences (NIBGE-C, PIEAS), Faisalabad, Pakistan*

[2] *Institute for Cardiogenetic, University of Lubeck, Lubeck, Germany*

[3] *Cologne Center for Genomics (CCG), University of Cologne, Germany*

Abstract: Rare genetic disorders affect a significant proportion of the global population. A large number of these patients are either misdiagnosed or remain undiagnosed which can have potentially adverse effects, including failure to provide anticipatory prognosis and identify potential treatment. With the completion of HGP, genetic testing has fast grown into a diagnostic discipline introducing new and cost-effective diagnostic tests with reasonable accuracy and specificity. NGS technologies, in particular, changed the field of genetic diagnosis by sequencing the entire genome or subset thereof in a single test and accomplishing diagnosis of virtually all diseases, either congenital or late-onset. These technologies have opened up new opportunities and unique challenges. This chapter discusses the importance of genetic testing, its scope, various technologies and approaches and, finally, the opportunities and challenges accompanying the new age genetic tests.

Keywords: aCGH, ARMS-PCR, Genetic disorders, Genetic testing, Massive Parallel Sequencing, NGS, Targeted Gene Panels, WES.

1. INTRODUCTION

Mendelian disorders are more commonly known as rare inherited disorders, especially among researchers working on these disorders. Around 7000 of these disorders are currently known of which a substantial number of disorders are life-threatening or chronically debilitating [1]. Ironically, however, 40 to 82 of every 1000 live births have one or another genetic disorder [2]. If a total load of genetic disorders constituted by all the inherited disorders together is considered, as much as 8% of the world population presents with a genetic disorder before adolescence [3].

[*] **Corresponding author Muhammad Tariq:** National Institute for Biotechnology and Genetic Engineering College, Pakistan Institute of Engineering and Applied Sciences (NIBGE-C, PIEAS), Faisalabad, Pakistan; E-mail: tariqpalai@gmail.com

Syeda Marriam Bakhtiar and Erum Dilshad (Eds.)

Thus, taken together, the so-called rare genetic disorders are not very rare! Moreover, these disorders are amongst the most difficult to diagnose in clinical practice owing to their genetic heterogeneity [4]. For instance, mutations in more than 150 genes have been implicated in inherited hearing loss only [5]. Thus, a large number of patients are either misdiagnosed or remain undiagnosed. This lack of diagnosis can fail to identify any potential therapeutic intervention and assess the risk of recurrence of the disorder in future pregnancies. Therefore, the completion of the Human Genome Project (HGP) in 2003 was heralded as the dawn of an era of genomic medicine [6] wherein individual genetic information will be used for making clinical decisions and delivery of personalized medicine [7]. Promises of rapid detection of mutations and improved diagnosis and prognosis made by proponents of this project fueled patients' desire for a rapid and accurate molecular diagnosis of their disease [8].

But until appropriate genetic tests are available for individual patients, access to the complete human genome alone cannot materialize the dream of personalized genomic medicine.

1.1. Genetic Testing and Its Scope

Ever since the first DNA test in the late 70s [9], genetic testing has fast grown to become an established diagnostic discipline today. Within the last few decades, genetic testing of disease-causing variants has been extensively used in clinical diagnosis and carrier screening of a large number of inherited disorders. In addition, it has also been used for prenatal diagnosis of the fetus in families with a history of a severe disease. Genetic testing is the procedure that detects variations in DNA to determine a patient’s predisposition to develop diseases and disabilities. Although genetic tests were in use before, their applications increased by an order of several magnitudes after the completion of HGP, changing our medical practice for good (reproductive medicine and oncology, for instance) [10]. Today, these applications span a variety of medical disciplines, such as newborn screening for highly penetrant disorders; diagnostic and carrier testing for inherited disorders; screening for adult onset and complex multifactorial disorders; and evaluation of drug dosage, selection and response in pharmacogenetic testing [11].

In this chapter, we discuss the genetic test and its types, and its importance in the diagnosis of rare genetic disorders. The utility of a particular genetic test, in the clinic, depends on the degree of genetic heterogeneity of the disease being investigated and prospects for therapeutic intervention. The selection of genetic tests and platforms is also guided by the nature of the disease, patient’s age, family history and available specimen. Some of the tests can rapidly detect gene

variants previously implicated in similar, or the same, diseases (allele-specific tests) while others are tailor-made for examining the entire coding sequence of one or more genes in search of as yet undiscovered causative variants; each strategy has its strengths and weaknesses [12]. Prenatal whole genome sequencing (WGS), for instance, is useful for detecting carrier status in a large number of heterogeneous rare disorders [13], whereas in single-gene disorders, such as beta-thalassemia, molecular diagnosis can be accomplished in routine by simple and low-cost PCR amplification [14].

1.2. Screening and Diagnostic Testing

A genetic test is different from other clinical tests in that the routine clinical test is purely diagnostic and is meant to select appropriate interventions for a patient. However, a genetic test can serve both as a screening test and as a diagnostic test. As a screening test, it can be used to screen asymptomatic individuals to identify those predisposed to disease or screen for a mutation with potential risk to an unborn child (fetus). As a diagnostic test, it is used to testify the presence of an active disease process. Predictive testing can assess a healthy individual's risk of developing a disease, way before its onset, although it cannot predict its onset and severity [15]. The primary objective of a screening test is to identify individuals who can benefit from further diagnostic testing. Diagnostic tests have higher sensitivity and specificity and specifically, look for a particular clinical condition [16].

1.3. Why Genetic Testing?

For a family with a history of a genetic disorder, genetic testing enables the parents to make decisions during pregnancy based on the information provided by the test, coupled with genetic counselling. An expecting family can, for instance, decide to terminate or continue their pregnancy based on the result of a prenatal genetic test [17]. Postnatal genetic tests are not only useful for inaccurate diagnosis but may also advise on the prognosis of the disease. It has an important influence on a patient's management and therapy decisions. Moreover, a genetic diagnosis can spare further diagnostics and lessen emotional and financial burdens on patients and their families [18].

2. TESTING TECHNOLOGIES

As a natural corollary of the explosion of molecular discoveries during the past few decades, clinicians need to keep up-to-date on research in molecular biology, genetics and genomics so that they understand and interpret the pathophysiology of diseases and how to incorporate genetic testing into their practice in a practical and effective way [12]. When deciding on which test to employ, clinicians, and of

CHAPTER 3

Preimplantation, Prenatal, and Postnatal Diagnosis

Sadia Nawaz[1,*] and **Humna Masood**[1]

[1] *Institute of Biochemistry and Biotechnology,University of Veterinary and Animal Sciences-UVAS, Lahore-54000, Pakistan*

Abstract: Pre-implantation genetic diagnosis (PGD) is a practical alternate evolving approach to prenatal diagnosis and termination of pregnancies in families with a high risk of Mendelian monogenetic and polygenetic disorders. Pre-implantation genetic diagnosis testing is continuing to extend immensely, along with a novel genetic analysis and *in vitro* fertilization approaches are in practice in the medical field throughout the world. However, PGD is regarded as ethically sensitive because repetitive termination of pregnancy causes huge psychological effects on the couples, and also because the low rate of pregnancy and birth makes it unreliable compared to prenatal testing. But it is also helpful in achieving additional goals *e.g.*, improved embryo and gender selection, overcoming the chances of birth of a child with an unknown genetic defect, better understanding of epigenomic regulations and reduction in the monetary burden of society. This chapter focuses on PGD, its procedure, utility and advantages, goals and objectives and the various issues surrounding it. We also discuss the future of this technology at the end of the chapter.

Keywords: Mendelian monogenetic and Polygenetic disorder, Prenatal diagnosis, Preimplantation genetic diagnosis.

1. INTRODUCTION

With the advent of new technologies in genomics, there is an opportunity to resolve or overcome the challenges of genetics and reproductive sciences. It helps to analyze recessive, dominant or X-linked disorders which are the cause of incurable early or late-onset phenotypes or morbidities. For this purpose, different diagnostic methods such as pre-implantation, prenatal and postnatal are offered to prevent the transmission of these genetic anomalies from parents to offspring. These diagnoses are done before or just after birth for the identification of the genetic basis of the disease, which opens up new possibilities to overcome and treat a disease completely or to some extent.

[*] **Corresponding author Sadia Nawaz**: Institute of Biochemistry and Biotechnology, University of Veterinary and Animal Sciences-UVAS, Lahore-54000, Pakistan; E-mails: sadia.nawaz@uvas.edu.pk and shamiryar@hotmail.com

Syeda Marriam Bakhtiar and Erum Dilshad (Eds.)

In this chapter, we briefly overview pre-implantation genetic diagnosis (PGD), prenatal diagnosis (PND), postnatal diagnosis, chromosomal abnormalities, and monogenic and polygenic diseases. We also discuss the diagnostic methods which could be employed before or after birth for the identification of genetic disorders in infants or adults.

1.1. Preimplantation Genetic Diagnosis (PGD)

Medical advances have allowed scientists to achieve preimplantation genetic diagnosis (PGD) clinically, which helps in the selection of genetically unaffected embryos before implantation in the uterus [1]. PGD was first ideated by Robert Edward in the mid-1960s. In 1968, Gardner and Edward conducted experiments for the production of animals of desired sex [2]. By using the sex-specific chromatin patterns, they were able to identify rabbit sex chromatin in the blastocyst stage and then transfer it to the uterus [3]. Later on, it became the first successful PGD test in humans. In the mid-90s, the first preimplantation genetic consultant Alan H. Handyside and his colleagues applied PGD for an X-linked disorder, cystic fibrosis (CF), in Hammersmith Hospital, Landon [4]. Preimplantation genetic diagnosis has advantages over prenatal diagnosis because some families with a high risk of genetically defected progeny in the future choose not to terminate a pregnancy after prenatal genetic testing. Therefore, the basic aim of preimplantation genetic diagnosis is to have healthy offspring and to avoid the invasive termination of pregnancy if the embryo is affected. PGD has applications in three groups of genetic anomalies. The first group comprises patients who are at high risk of transmission of monogenic disorder to their progeny *e.g.* Duchenne muscular dystrophy (DMD) and CF *etc*. The second group consists of families with imbalanced chromosomal abnormalities *e.g.*, translocation, where the number of affected offspring's birth and abortions could be reduced by using PGD. The third group includes patients having aneuploidy in children due to advanced maternal age [5].

Furthermore, there are three professional and highly qualitative laboratories involved in PGD [5] and their details are as follows:

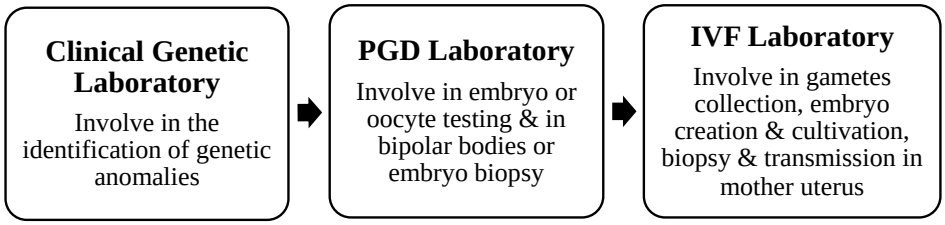


Fig (3.1). Conceptual diagram of professional laboratories involved in pre-implantation diagnosis.

1.1.1. PGD and In vitro Diagnostic Procedures

For the preimplantation genetic diagnostic procedure, family counselling and evaluation are concisely done by IVF members. Afterwards, a variety of procedures can be used for obtaining zygote or oocytes *e.g.*, ovarian hyper-stimulation, *in vitro* fertilization (IVF) techniques *e.g.*, intra-cytoplasmic sperm injection.

1.1.2. Phases of PGD

Embryo biopsy and genetic diagnosis are two phases of PGD; their details are given below.

1.1.3. Embryo Biopsy

PGD biopsy is carried out in two steps: the first step is puncturing the zona-pellucida (a membrane that is present in the surrounding of embryo or oocyte) and the second is the removal of cells from the embryo. The membrane can be breached or punctured by using chemical, mechanical or laser techniques. Different biopsy methods have been reported which are given below [6]:

PB (Polar Body) Biopsy is the first stage of biopsy in which the two polar bodies are removed sequentially or simultaneously. On the 0^{th} day, the first PB is removed from the oocyte in the sequential biopsy method. Then on the 1^{st} day, the second PB is removed after fertilization [7, 8].

Blastomere (Cleavage Stage) Biopsy: The mitotic division of the human zygote occurs after every 24 hours. So, two cells (blastomeres) can be biopsied on the 3^{rd} day without disturbing embryo development and metabolism. Blastomere aspiration and zona-pellucida puncturing are done by using a micro-manipulator [9].

Trophectoderm (Blastocyst) Biopsy: Without harming inner cell mass, 10 to 30 cells in the blastocyst stage can be biopsied on 5^{th} or 6^{th} day. The breaching of zona pellucida is done on the 3^{rd} day and then on the 5^{th} day, cells are removed for preimplantation genetic diagnosis [10].

1.1.4. Genetic Diagnostic Analysis

Different diagnostic techniques such as PCR, FISH, micro-array and next generation sequencing *etc*. are used in PGD. Details are given in the second part of this chapter.

CHAPTER 4

Genetic Counseling in Inherited Disorders

Shumaila Zulfiqar[1,2], Muhammad Tariq[1,*], Naveed Altaf Malik[1], Ayaz Khan[1], Shafaq Ramzan[1], Maria Iqbal[1,3], Iram Anjum[2] and **Shahid Mahmood Baig[1]**

[1] *National Institute for Biotechnology and Genetic Engineering College, Pakistan Institute of Engineering and Applied Sciences (NIBGE-C, PIEAS), Faisalabad, Pakistan*

[2] *Department of Biotechnology, Kinnaird College for Women, Lahore*

[3] *Centogene GmbH, Rostock, Germany*

Abstract: In this chapter, we have focused on the journey of sorting genes and the connotation of genetic counseling started. In a literal sense, we will understand how genetic counseling could contribute to identifying pathogenicity and penetrance of genetic mutation/s in high-risk individual/s or populations. Great strides have been achieved in terms of diagnosis, management, and treatment of various genetic disorders due to rapid advancements in genetic research. The national Thalassemia Prevention Program of Cyprus has been one of the earliest and most celebrated successes in lowering the disease burden and improving life quality and survival rate in patients. The knowledge regarding gene/s and variant/s is quite instrumental for making important reproductive decisions and therapeutic interventions for both rare and common disorders. We also touch upon the associated ethical issues and challenges.

Keywords: Carrier screening, Eugenics, Genetic counseling, Genetic disorders, Genetic testing, Population screening.

1. INTRODUCTION

In 1902, Archibald E Garrod published a study in the journal *The Lancet,* suggesting an autosomal recessive mode of inheritance in a genetic disorder called alkaptonuria. This study enabled researchers to develop an association between genetic disorders and the laws of inheritance published by Gregor Mendel in 1865. From that point forward, our insight into genetic disorders has increased exponentially. Nonetheless, it does not mean that the history of genetic disorders began with the rediscovery of Mendel's laws toward

[*] **Corresponding author Muhammad Tariq**: National Institute for Biotechnology and Genetic Engineering (NIBGE-C), Faisalabad, Pakistan Institute of Engineering and Applied Sciences (PIEAS), Islamabad, Pakistan; E-mail: tariqpalai@gmail.com

Syeda Marriam Bakhtiar and Erum Dilshad (Eds.)

the start of the twentieth century. Earlier paintings, drawings and sculptures show evidence of patients with different conditions like achondroplasia, Robert's condition, hermaphroditism and neurofibromatosis. The famous X-linked blood disorder in Queen Victoria 's family had been known for a long time in the Middle Ages, referred to as “passio transition sanguinis” in the Talmud. Later in the nineteenth century, J L Schónlein named it “haemorrhophilia”; a term later transformed into the current “hemophilia”. In 1866, British doctor, J Langdon Down, characterized mental retardation on the basis of different racial types; Caucasian, Ethiopian, Malaysian and Mongolian. The first three classes were not much persuading, therefore, their use was stopped gradually. However, Down's mongoloid idiot-ism, these days known as Down's syndrome, ended up being a well-characterized disorder. After 93 years of Down’s paper publication, Lejeune found the reason for Down's condition: a little additional chromosome (chromosome 21) [1 - 5]. The scenario before the advent of modern medical science was completely different from today when complex cytogenetic analyses, proper diagnosis and management methods were unavailable. The only available option was to depict and analyze disease occurrence or reoccurrence verbally and communicate them to patients. Later, this method was referred to as genetic counseling.

The practice of advising people regarding genetic conditions started around the twentieth century, soon after William Bateson proposed that the new clinical and medical investigations of heredity be designated as “genetics”. Heredity got interlaced with social changes and structured as a field of ‘eugenics. Although eugenics was conceived with the intention to prevent inherited disorders, later development had deplorable results. Numerous states in the USA had laws ordering sterilization of the affected individuals, others were not permitted to move. By the 1930s, these policies were acknowledged by various different nations including Germany where the killing of the “genetically defective” persons was sanctioned in 1939 [6]. In 1992, a law was approved in Australia suggesting that parents of an affected child cannot have sterilization or hysterectomy by themselves to avoid defective child birth. Permission would be required from the court to do so. However, even after 10 years of this legislation, illegal sterilization of minors was noticed [7]. Edwin Black has termed the eugenic policies as “eugenics wars” and suggested a negative impact on people’s life due to isolation, discrimination, genocide and violation of basic human rights. However, supporters of eugenics like Comfort proposed that it drives towards elimination of disease and a healthier society [8].

The term “genetic counseling” was formally coined in the post eugenic era by Sheldon Reed in his popular book *Counseling in Medical Genetics* published in 1955 [9, 10]. According to the World Health Organization, the first-ever genetic

counseling service began at Heredity Clinic in the University of Michigan, USA in 1940. It is rather impossible to exactly speculate how the participating individuals would have acted regarding outcome of counseling (eugenic or dysgenic). Later, multiple genetic counseling centers were opened across the globe. (https://www.who.int/genomics/professionals/counselling). In 1963, phenylketonuria became the first ever genetic disorder for which a newborn screening was performed [11]. From then onwards, the idea prevailed that genetic counseling was used far earlier as genetic testing. The American Society of Human Genetics (ASHG) defines genetic counseling as a process of facilitating individuals to combat medical, psychological and familial implications of genetic disorders [12]. General objectives of genetic counseling, agreed upon by all professionals, are; 1. prevention of birth defects and genetic disorders, 2. promoting adaptation to a genetic condition without compromising on psychological well-being of counselees 3. counselees should make their own reproductive choices, however, decisions that reduce the impact of certain genetic condition/s should be encouraged [13].

With the collaboration between clinicians and geneticists, there has been an increase observed in genetic counseling . Human Genome Project opened new horizons to identify individuals carrying mutant allele/s responsible for inherited disorders before disease appearance. These advancements have broadened the prospects to delineate inherited disorders by analyzing human chromosomal abnormalities, providing facilities to analyze certain disorders in the first trimester of pregnancy (prenatal) and the initiation of screening programs for certain diseases in high-risk populations [14]. Various diagnostic centers provide services like karyotyping, dermatoglyphic analysis, syndrome recognition and biochemical tests related to the inborn errors of metabolism. Carrier screening and genetic counseling are offered to individuals with a positive family history of a certain genetic condition. Couples at risk of having affected children with genetic disorder(s) can be provided with reproductive options to avoid affected births. In this way, genetic counseling makes significant contributions to disease prevention as well as management.

As the initial step, the regular counseling process comprised a single interview where pedigree is prepared and a prediction/estimation of recurrence risk is provided to counselees. Later, as the psychological difficulties associated with counseling emerged, the research circle was expended to various other aspects including procedures, reasoning, psychodynamics and morals of advising. In some countries, the meeting and counseling are done by the same individual; in others, the counseling approach is a multiple tier process requiring interviews and multidisciplinary process. Mostly, the counselor lacks formal training in the methods/techniques of counseling, even if such trainings would be a benefit, but

CHAPTER 5

Genome-Wide Association Studies (GWAS)

Hafiza Noor Ul Ayan[1,2] and **Muhammad Tariq**[1,*]

[1] *National Institute for Biotechnology and Genetic Engineering College, Pakistan Institute of Engineering and Applied Sciences (NIBGE-C, PIEAS), Faisalabad, Pakistan*

[2] *Institute for Cardiogenetic, University of Lubeck, Lubeck, Germany*

Abstract: Genome-wide association studies (GWAS) are designed to find associations between genomic variants and a phenotype, usually a complex multifactorial disease. The idea for association studies in a large cohort was floated after linkage analysis, which proved extremely successful in the identification of causative genes for rare disorders, but it did not come up to expectations in the case of common complex disorders where causative alleles are less frequently aggregated in families. Ever since their advent in 2005, GWAS have transformed gene identification ventures in complex disease genetics over the past fifteen years, giving rise to several powerful associations for complex traits and disorders. Association studies are based on the "common disease common variant" hypothesis which assumes that genomic variation with low penetrance and high population frequency are involved in the causation of common complex disorders. Although GWAS, complemented with the downstream functional assessment of the variants, have been successful in identifying novel disease-causing genes and biological mechanisms, the field has also received intense criticism over the years, especially its failure in tracing the so-called 'missing heritability'. Therefore, further functional studies are mandatory to precisely establish a link between risk alleles and a phenotype. This chapter broadly covers an introduction of GWAS, their successes and limitations, and various important factors affecting the design and results, followed by challenges in the post-GWAS era.

Keywords: Genome-wide Association Studies, Linkage Disequilibrium, Multifactorial Diseases, Missing Heritability, SNPs, WGS.

1. INTRODUCTION

During the last century, conventional approaches, such as linkage analysis using PCR and conventional sequencing, were extensively used to map genomic regions co-segregating with disease phenotypes. These techniques were quite successful

* **Corresponding author Muhammad Tariq**: National Institute for Biotechnology and Genetic Engineering College, Pakistan Institute of Engineering and Applied Sciences (NIBGE-C, PIEAS), Faisalabad, Pakistan;
E-mail: tariqpalai@gmail.com

in identifying genes underlying Mendelian disorders (*e.g.*, cystic fibrosis). However, linkage analyses were not found very fruitful in case of complex traits and diseases (*e.g.*, cardiovascular diseases, cancer and Parkinson's disease), wherein multiple genetic variants add to the phenotype, as well as the environmental risk factors (*e.g.*, diet, smoking, alcoholism and sedentary lifestyle) which also contribute to disease picture. These disorders and individuals with these disorders are less frequently aggregated in families, thus, the corresponding causal alleles are less likely to be shared among related patients [1]. This limitation of linkage analysis necessitated novel approaches that could be applied to dissect the genetic architecture of complex multifactorial diseases in unrelated patients. This inspired Genome-wide Association Studies, popularly known as 'GWAS' in 2005 [2].

1.1. Rationale

GWAS are designed to identify genomic locations harboring variants (a marker allele; typically, a SNP) associated with disease risk [3]. It is basically a case-control study in which genotypes associated with the certain genomic variants have different frequencies among affected and phenotypically healthy individuals. GWAS are usually designed for analyzing large cohorts of unrelated individuals or nuclear families since a large number of patients with similar disease phenotype is relatively easy to collect within a population, compared to within families (a requirement for linkage analysis). Assume that causal variants are not rare, theoretical arguments support the power of GWAS in elucidating the genetic basis of multifactorial disorders [4]. Therefore, GWAS are primarily designed around the hypothesis "common disease common variant (CDCV)", which argues that complex traits underlie genetic variants with low penetrance but high population frequency [5].

Identification of the CFH gene variant, causative for age-related macular degeneration (AMD) in European population, was among the earliest success stories of GWAS, published back in 2005 [2]. This study analyzed >100,000 SNPs in 96 affected and 50 healthy individuals to discover a strong association between a common intronic variant and the disease. However, subsequent sequencing of the identified gene revealed an exonic variant that changes tyrosine at amino acid 402 by histidine (p.Y402H). The coding variant has a larger impact with a relative risk of 7.4 in homozygous patients in comparison to individuals with wild type genotype. This study, undoubtedly, emphasized the potential of GWAS to explore underlying genetics of complex disorders. Later in 2007, Welcome Trust Case Control Consortium (WTCCC) embarked on seven GWAS initiatives, simultaneously, for seven complex diseases in the UK population (type

1 and type 2 diabetes, coronary heart disease, hypertension, bipolar disorder, rheumatoid arthritis and inflammatory bowel disease) [6]. In their landmark paper, the consortium demonstrated that for reproducible discoveries using GWAS, large sample sizes, a thorough study design and rigorous criteria are a must. By 2010, GWAS had identified more than 3000 loci for over 250 diseases and phenotypes [7]. In the following 10 years, GWAS accelerated gene discovery exponentially with more than 4300 research papers from 4500 association studies implicating around 55,000 genomic loci for over 5000 diseases [7]. Most of the GWAS data are publicly available and several user-friendly data portals are accessible that help scientists to analyze GWAS data easily.

Association depends on linkage disequilibrium (LD), which is the co-segregation of a marker allele with a causal variant, by virtue of their genomic proximity, across a population [8]. LD, a correlation structure, in genomic variants is a reflection of evolution, limited population size, mutation, natural selection and recombination rate. The statistical power of associations studies depends on several factors *e.g.*, sample size, distribution of effect sizes of causal variants segregating in the population, minor allele frequency (MAF) of the variant and LD between SNP variant and as yet unknown causative variant [9].

The core objective of GWAS is to improve the understanding of disease biology to pave way for prevention or better treatment. However, in most cases, GWAS findings do not necessarily have a utility in the prevention or treatment of disease. GWAS detect variants that are associated with phenotypic differences, therefore, the association between a genomic variant and a disease phenotype does not necessarily provide direct information about the causal gene or disease mechanism. Identification of causal variant frequently requires follow up studies to narrow down the associated region. However, novel analytical methods and molecular technologies provide us with the opportunities to fill the gap between sequence and consequence [9].

GWAS, over the past decade, have successfully detected associations of thousands of SNPs with human diseases [10]. However, the associated alleles typically have high frequency in a population with a MAF >5% [11]. Moreover, an overwhelming majority of SNPs (~90%), associated with different traits, are found in the non-coding parts of the genome [12] which further complicates identification of relevant genes, causal variants and mechanisms [13]. Nonetheless, the limitations of GWAS can be partially overcome by using cohorts of small families or sporadic patients from an isolated geographical region [14] and by using whole genome sequencing (WGS). Data generated by WGS can better capture low frequency rare variants. Moreover, it facilitates capturing variants that are in strong LD with SNPs on a genotyping array. Therefore, WGS

CHAPTER 6

Regenerative Medicine

Hajra Qayyum[1,*] and **Syeda Marriam Bakhtiar**[1]

[1] *Department of Biosciences and Bioinformatics, Capital University of Science and Technology, Islamabad*

Abstract: Regenerative medicine (RM) is defined as a replacement and revival of human cells, tissues, or organs to reinstate or reconstruct their normal physiology. RM is regarded as a solution to provide healthy substitutes for a malfunctioning/failed organ or a tissue. It is emerging as the suitable substitute for organ transplantation. Transplantation seems impractical due to the limited availability of donors as significant disparities lie between the number of patients that require transplantation and the availability of organs from the donor, so there was a gap created. Therefore, to comply with these needs, RM has emerged as a new science to create biological replacements and exploit the body's ability of regeneration to recover and sustain normal function in diseased and damaged tissues. This chapter overviews RM in terms of adopted strategies, its clinical applications in organ engineering along with inherited challenges and their plausible solutions.

Keywords: Cell-therapy, Grafting, Organ engineering, Pluripotent stem cells, Regenerative medicine, Stem cell therapy, Tissue culturing.

1. INTRODUCTION

Regenerative medicine (RM) is described as a replacement and revival of human cells, tissues, or organs to reinstate or reconstruct their normal physiology [1]. The term “regenerative medicine” was first formulated by William Haseltine in 1999 to represent an interdisciplinary field that is an amalgam of apprehensions derived from various fields. These fields included tissue engineering, cell transplantation, stem cell biology, biomechanics prosthetics, nanotechnology, and biochemistry [1, 2].

Besides normalizing congenital anomalies, RM could be employed to repair or replace different body tissues and organs damaged by aging, ailment, or a shock. To date, promising preclinical and clinical results for the treatment of chronic

* **Corresponding author Hajra Qayyum:** Faculty of Health and Life Sciences, Capital University of Science and Technology, Islamabad, Pakistan; E-mail: hajraqayyum92@gmail.com

Syeda Marriam Bakhtiar and Erum Dilshad (Eds.)

disorders and acute injuries encourage the use of RM as an effective treatment option. It can also ameliorate ailments that can affect a multitude of organ systems and also provide support in case of dermal wounds, cardiovascular diseases and traumas, cancer, and many more.

Another treatment option that can provide the solution is the grafting of healthy organs and tissue to compensate for a malfunctioning/failed organ or a tissue. However, this option seems impractical due to the limited availability of donors as there remain significant disparities between the number of patients with diseased or damaged organs that require replacement and the availability of organs to be transplanted. As the lifespan of the population progresses and the number of new cases of organ failure rises, the supply/demand ratio for the organs intensifies. Sometimes the situation gets worsens because of serious immune complications, however, these challenges can theoretically be overcome by the application of RM strategies [3]. To comply with these needs, RM emerged as a new science to create biological replacements and exploit the body's ability of regeneration to recover and sustain normal function in diseased and damaged tissues [2].

2. APPROACHES TO RM

The field of RM involves a variety of approaches which comprises the use of materials and de novo produced cells and various other combinations to replace malfunctioning tissue, effectively replacing it structurally and functionally, or to promote the process of tissue healing. While adult humans have a limited regenerative ability in comparison to lower vertebrates, the body's innate healing response may also be used to encourage regeneration [2, 3]. Three strategies are widely adopted to pursue the aim of RM [1, 2]. These strategies are:

1. Cell-based therapy

2. Use of biomaterials

3. Scaffold implantation seeded with cells

2.1. Cell-based Therapy

All human cells are derived from the same origin *i.e.*, zygote. During development, these cells progressively differentiate into more specialized cells with specific cellular functions. This ability of differentiation possessed by cells is termed “cell potency”. Cell-based therapies exploit this cellular property of potency. It involves injecting novel and healthy cells into pathological tissues. The injected cells are either the pre-differentiated or the undifferentiated stem

cells which can later get differentiated depending upon the underlying circumstances [1].

The pre-differentiated cells are extracted by the patient's specific tissues as they are ready to implant just by expansion. However, it is difficult to get a substantial number of these cells in vitro or *in vivo* as they lose the usual microenvironment required for proliferation. Hence, these cells are less likely to be utilized in the future even when they do not associate with rejection and noticeable inflammatory responses [1]. On the other hand, Stem Cells (SC) can proliferate extensively, with the ability of self-renewal while they keep their undifferentiated state until they are induced to differentiate into a specific cell type. SCs are divided into various types such as autologous, allogenic, and xenogeneic depending upon the source of their origin [1].

2.1.1. Adult Stem Cells

In case of histological injuries, adult stem cells (ASCs) are derived from the tissues of an adult human body, which then perform corrective functions, restoring normal tissue functioning. Among these cells, mesenchymal stem cells (MSCs) are of special importance as they have the potential to get differentiated into other cell types that are particularly required for the amelioration of ailments related to bone, cartilage, nerves, muscles, cardiovascular, blood, and gastrointestinal [1].

2.1.2. Pluripotent Stem Cell-Based Cell Therapies

Pluripotent Stem Cells (PSCs) have the potential of infinite proliferation and getting differentiated into the cells of the three germ layers. With these two attributes, PSCs become the most suitable source for cell therapies in case of different diseases and injuries. Two types of human PSCs are in clinical use: embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs) [4].

2.1.2.1. Embryonic Stem Cells

These cells are derived from the inner most cell of an embryo during the blastocyst stage and can proliferate extensively. ESCs keep their pluripotent state intact until they are induced to get differentiate into any of the three embryogenic germ layers. Human ESCs are usually extracted from a surplus of embryos during in-vitro fertilization. Another approach to extract ESCs could be therapeutic cloning or somatic cell nuclear transfer (SCNT) which includes the transferal of the somatic cell nucleus into an oocyte [1]. Lastly, SCs can also be obtained from amniotic fluid or placenta via the process of amniocentesis. Such cells are known as amniotic fluid stem cells (AFSCs). Human ESCs have two concerns from the

CHAPTER 7

Emerging OMICS and Genetic Disease

Muhammad Jawad Hassan[1,*], **Muhammad Faheem**[1] and **Sabba Mehmood**[1]

[1] *Department of Biological Sciences, National University of Medical Sciences, Rawalpindi, Pakistan*

Abstract: Multiomics also described as integrative omics is an analytical approach that combines data from multiple 'omics' approaches including genomics, transcriptomics, proteomics, metabolomics, epigenomics, metagenomics and Meta transcriptomics to answer the complex biological processes involved in rare genetic disorders. This omics approach is particularly helpful since it identifies biomarkers of disease progression and treatment progress by collective characterization and quantification of pools of biological molecules within and among the various types of cells to better understand and categorize the Mendelian and non- Mendelian forms of rare diseases. As compared to studies of a single omics type, multi-omics offers the opportunity to understand the flow of information that underlies the disease. A range of omics software and databases, for example WikiPathways, MixOmics, MONGKIE, GalaxyP, GalaxyM, CrossPlatform Commander, and iCluster are used for multi-omics data exploration and integration in rare disease analysis. Recent advances in the field of genetics and translational research have opened new treatment avenues for patients. The innovation in the next generation sequencing and RNA sequencing has improved the ability from diagnostics to detection of molecular alterations like gene mutations in specific disease types. In this chapter, we provide an overview of such omics technologies and focus on methods for their integration across multiple omics layers. The scrupulous understanding of rare genetic disorders and their treatment at the molecular level led to the concept of a personalized approach, which is one of the most significant advancements in modern research which enable researchers to better comprehend the flow of knowledge which underpins genetic diseases.

Keywords: Genomics, Metabolomics, Multiomics, Proteomics.

1. INTRODUCTION TO OMICS AND GENETIC DISEASE

Rare diseases do not mean very "rare" as 8000 types of rare diseases have been described. An estimated 262.9–446.2 million people are living with a rare disease globally (Global Genes. RARE Diseases: Facts and Statistics [Available: https://globalgenes.org/rare-diseases-facts-statistics/). Rare disease classification

[*] **Corresponding author Muhammad Jawad Hassan:** Department of Biological Sciences, National University of Medical Sciences, Rawalpindi, Pakistan; E-mail: jawad.hassan@numspak.edu.pk

Syeda Marriam Bakhtiar and Erum Dilshad (Eds.)

and definition vary globally; in Europe, it is any disease with an incidence of less than one in 2000 (RARE DISEASES - a major unmet medical need. Luxembourg: European Commission, 2017), in the United States (US), they are conditions affecting fewer than 200,000 people (Global Genes. RARE Diseases: Facts and Statistics- Available at: https://globalgenes.org/rare-diseases-facts-statistics/], and in China, they define it as a disorder prevalent in less than one in 500,000 within the population [1]. Numerous common issues faced by patients ranked under the 'rare' umbrella. However, the problems faced by patients and the disease burden imposed on the country's population remain the same, as obtaining a diagnosis, suitable medical care and access to support services is always a great challenge for patients with a rare disease. Next-generation sequencing techniques have increased the chances of identification of causative mutation in these patients as compared to single gene testing alone, with a significant increase in diagnostic rates, and it is believed that "integrated multi-omic analysis" may further increase this diagnostic yield. On the other hand, "multi-omic" approach aids the explanation of genotypic and phenotypic heterogeneity, which is not possible with single omic analyses.

2. ADVANCED TECHNIQUES IN "OMICS"

The recent advancement in technology has revolutionized many fields of research. "OMICS" study that includes genomics, transcriptomics, metabolomics and proteomics, has also gained the attention with potential to analyse the molecular data at a higher pace. The adopted advanced high-throughput techniques for "OMICS" data are now often incorporated in the routine biological research. The newly advanced techniques help us to analyse and generate tera- to penta-bytes data files on a routine basis. These big-datasets lead to the understanding of biological systems and solving biological question. In this chapter, we are discussing the applications of these techniques in different biological research.

2.1. Emerging Omics Techniques: Genomics and Transcriptomics

2.1.1. Genomics

The first discipline of OMICS that appears at first is genomics, which deals with the study of entire genomes unlike genetics which focuses on single genes or individual variants. Genomics has been utilized in different facets of research and clinical applications which range from diagnostics, pharmaceuticals, pharmacogenomics, disease prevention, gene therapy, and developmental biology to comparative and evolutionary genomics. The Whole Genome Sequencing (WGS) data generation initiated with the Human Genome project (HGP), which was initiated in 1990 and completed in 2003. With time, the genome sequencing

technology has evolved which transformed the DNA sequencing industry. During the start of the 21st century, completion of WGS of a single haploid human genome took many years with spending of more than 3 billion dollars [2]. The advancement in the sequencing technology over the period of time has improved. These days, WGS is as cheap as a few thousand dollars with some days of work. Advance NEXT Generation Sequencing (NGS) techniques like nanopore (*e.g.* MinION), SMRT (single molecule real-time; *e.g.* sequel system) and Semiconductor (*e.g.* Ion S5 sequencer) are the keys for reducing the cost and time [3 - 5].

Unlike the traditional Sanger sequencing technology, the NGS methods are more popular and in use these days. The most frequently used NGS to date is sequencing by synthesis (SBS) method. These methods utilize a solid support which contains micro wells where the sequencing reactions occur. The individual DNA molecules which have to be sequenced are distributed in these micro wells or linked to a solid support. The DNA fragment sequence is identified while its complementary strand is being synthesized, thus, each of the sequenced DNA fragment serves as a template. The nucleotides in the reaction are labelled or upon incorporation identified *via* chemical reaction, which can be detected or imaged.

Roche 454 pyro sequencing was the first SBS based NGS technology developed in 1996 that was introduced in the market (1999) for genomics purposes. This technology was based on the detection of pyrophosphate (PPi) as a by-product upon insertion of a nucleotide. The PPi released is converted into ATP by the enzyme present in the reaction. This ATP is utilized as a substrate by luciferase-catalysed reaction to emit light [6]. The 454-sequencing technology was generally utilized for genomics especially metagenomics samples, since it has long read lengths of up to 600-800 nt which typically can be achieved with high throughput of 25 million bases with 99% or more accuracy in a 4 hrs reaction run, thus facilitating genome assembly.

Since, there are multiple complicated steps involved, so there are error chances while reading homopolymer sequences longer than 6bp, thus, pyrosequencing was replaced by Illumina sequencing which has become industrial standard. Illumina sequencing also supports a range of procedures which include exome and targeted sequencing, genomic sequencing, metagenomics, Chromatin immunoprecipitation sequencing and RNA sequencing. Illumina technology was first established by Solexa and Lynx Therapeutics. In this technique, genome is chopped into DNA fragments, and combined with designed adapters (forked adapters) at the both ends of DNA fragments [7]. This technique sequencing is known as “bridge amplification” wherein the DNA fragments have a length of about 500bp ligated with adapters. The adapters are linked on a solid support (Flow cell) where the

CHAPTER 8

Integrated Bioinformatics and Computational Biology Approaches: Applications in Diagnosis and Therapeutics

Fatima Shahid[1], Shifa Tariq Ashraf[1], Hayeqa Shahwar Awan[1], Amina Basheer[1] and **Amjad Ali[1,*]**

[1] *Department of Industrial Biotechnology, Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), Islamabad, Pakistan*

Abstract: The advent of bioinformatics and integrated biology approaches has given rise to new avenues of diagnostic and therapeutic regimes. Living systems have been explored to identify disease-associated biomarkers that facilitate the early diagnosis of perilous medical conditions. Likewise, gene networks are pondered upon to obtain better insights into biochemical systems that can assist in the prediction and testing of the effects of various interactions within the systems. Genomics and proteomics-based approaches are being explored to facilitate the early diagnosis of cancers, shifting the paradigm towards noninvasive diagnostic alternatives. Bioinformatics has also fueled pharmacogenomics and pharmacogenetics-based strategies that have in turn contributed to the development of personalized medications. Similarly, the reverse vaccinology approach has emerged as a prominent option to combat deadly pathogens that were otherwise unrestrainable. This chapter highlights the fruits of integrated bioinformatics in diagnosing and treating detrimental conditions.

Keywords: Integrated bioinformatics, Precision medicine, Computational diagnostic tools, *In silico* Therapeutics, Reverse vaccinology.

1. INTRODUCTION

Bioinformatics and computational biology are integrative fields that apply *in silico* methods to explore the large assembly of biological data, for instance, genetic sequences, protein samples, and cell populations to devise advanced predictions. These computational methods include mathematical modelling, simulation, and analytical methods. Computational models can be constructed for

* **Corresponding author Amjad Ali**: Department of Industrial Biotechnology, Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), Islamabad, Pakistan; E-mail: Amjad_uni@gmail.com

Syeda Marriam Bakhtiar and Erum Dilshad (Eds.)

the potent diagnosis of disease progression and in response to therapeutics at a personalized level.

1.1. Exploration of Disease-Associated Biomarkers

Biomarkers are the distinct characteristics that are evaluated and serve as indicators of biological, pharmacological, as well as pathogenic processes in response to therapeutic interventions or exposure. Biomarkers are used as a tool for the prediction of disease severity, and drug response, as well as in the drug development process [1]. For instance, the detection of specific autoantibodies in the patient's blood is a potential biomarker used to detect autoimmune diseases. For rheumatoid arthritis (RA), rheumatoid factors are considered predominant diagnostic biomarkers. Hence, for detection purposes, anti-citrullinated protein/peptide antibodies (ACPAs) formed against the patient's own body citrullinated proteins are considered. These ACPAs can be detected at early stages as potential biomarkers in the blood, before the onset of symptoms associated with RA. Therefore, these biomarkers can be helpful in the prognosis of RA [2].

Biomarkers have been discovered through different high-throughput techniques to comprehend transcriptomics such as microarray, gene expression data, and genomics strategies that include genome sequencing, genome annotation, along with proteomics techniques and metabolomics frameworks which comprise mass spectrometry and nuclear magnetic resonance imaging, respectively. Proteins, peptides, genetic markers, and/or histological data have the potential to act as biomarkers. Generally, a genetic biomarker is a single gene or a collection of genes [3, 4]. Some of these have been described as follows:

Human Epithelial-growth Factor (HER2) is a protein receptor present on the cells of the human body. In colorectal cancer tumors, HER2 gene and its receptors are overexpressed by 3-4% while in the case of breast tumors, it is overexpressed up to 20-30%. Therefore, this gene is used as a predictive and prognostic marker in case of both cancers [5]. Likewise, carbohydrate antigen125 (CA125) is a high molecular weight glycoprotein which serves as a serum biomarker. In patients with epithelial ovarian cancer, the expression of CA125 increases as compared to a normal state. Moreover, by overexpression of this CA125 different signaling pathways including P13K/AKT and ERK pathways are stimulated. Similarly, Human Epididymis protein 4 (HE4) is another protein coded by the WFD2 gene that acts as a protease inhibitor, and overexpression of this protein in ovarian cancer has been observed that can, in turn, irritate EFGR and MAPK signaling pathways. Due to these significant roles, HE4 was approved by the Food and Drug Administration (FDA) as a monitoring biomarker in women with epithelial ovarian cancers. In contrast to HE4, the concentration of ApoA1 is decreased in

the serum of patients with ovarian cancer and can serve as a potential biomarker. Likewise, BRCA1, BRCA2, and p53 are genetic biomarkers that are attributed to breast cancer. However, the BRCA1 gene, located on chromosome 17q12-21 is pondered upon and hypermethylation of BRCA1 was found to be involved in both ovarian as well as breast tumors. Moreover, p53 is another potential biomarker that is a tumor suppressor gene and reportedly 50% cases of ovarian cancer have a mutation in this gene [6].

Similarly, glycated Haemoglobin (HbA1c) is a traditional biomarker for the diagnosis of diabetes and pre-diabetes. Fructosamine (FA) is a keto-amine that is formed by the glycosylation of total serum proteins majorly albumin. When the concentration of glucose is high, the concentration of FA increases. Hence, FA can serve as another biomarker for the screening of diabetes. Similarly, HDL-C, another biomarker of diabetes is a major lipoprotein that is associated with insulin secretion. If the concentration of HDL-C is lower, it will eventually lead to the development of diabetes from the pre-diabetic state [7]. Moreover, Alzheimer's disease is the most common neurodegenerative disorder and in the United States, it is the 6^{th} leading cause of death [8]. Some biomarkers associated with Alzheimer's are cerebrospinal fluid (CSF), amyloid β, phosphorylated tau, and tau proteins [9].

1.2. Computational Models as Tools to Identify Key Biomarkers

Biomarkers are the driving force for the prognosis, diagnosis, and treatment of a disease and are identified by different computational and bioinformatics tools that are curated. Resultantly, these tools can help scientists in the early diagnosis and treatment of a disease [3].

The advent of DNA microarray technology has aided in the transition from systematic methods to biological discovery; these transitions have already started to have a deep impact on biological research, medicine, and pharmacology. The quantitative data about the complete cell transcription profile pledges to become an effective way of studying the basic biology, disease diagnosis, drug development facilitation, and modify therapeutics to specific pathologies. These data can help generate the database containing living processes information [10].

Researchers have worked on survival related gene network modules [11]. In a study, a deep learning-based risk stratification model was developed for lung cancer by choosing an illustrative gene from survival-related network modules. Similarly, a new computational model was presented, based on SimRank and density-based clustering recommender model for prediction of miRNA-associated diseases (SRM-DAP) [12]. Another sample-specific method was

CHAPTER 9

Multi-omics Data Integration: Applications in Systems Genomics

Anam Naz[1,*], Ammara Siddique[2], Aqsa Ikram[1], Bisma Rauff[5], Huma Tariq[3] and **Sajjad Ahmed[4]**

[1] *Institute of Molecular Biology and Biotechnology (IMBB), The University of Lahore (UOL), Lahore, Pakistan*

[2] *Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), Islamabad, Pakistan*

[3] *Department of Zoology, Hazara University, Manshera, Pakistan*

[4] *Department of Health and Biological Sciences, Abasyn University, Peshawar, Pakistan*

[5] *Department of Biomedical Engineering, University of Engineering and Technology (UET), Lahore, Narowal Campus, Pakistan*

Abstract: Interpretation of molecular differences and intricacy at multiple stages, for instance proteome, genome, epigenome, metabolome, and transcriptome is needed for a thorough understanding of disease and human health. Biology has been reliant on data produced at these stages, which is collectively referred to as multi-omics data, after the emergence of sequencing techniques. Among all the aspects of biology, rapid development in high-throughput data initiation has enabled to carry out research on multi-omics systems biology. Metabolomics, proteomics, and transcriptomics data can provide answers to the targeted biological queries about the expression of metabolites, proteins, and transcripts, independently. A concise summary of multi-omics data sources, challenges in datasets integration, and visualization portals is also discussed.

Keywords: Genomics, Metabolomics, Proteomics, Transcriptomics, Omics.

1. INTRODUCTION

The approach of systematic multi-omics integration can methodically annotate, model, and assimilate such considerable sets of data. Multi-omics data availability has modernized the domain of biology and medicine by generating opportunities for different methods of an integrated system. Analysis of clinical and multi-omics data has acquired the lead in deriving productive understandings of cellular functions. Multi-omics data integration provides necessary information regarding

* **Corresponding author Anum Naz**: Institute of Molecular Biology and Biotechnology, University of Lahore, Lahore, Pakistan; E-mail: anam.naz@imbb.uol.edu.pk

Syeda Marriam Bakhtiar and Erum Dilshad (Eds.)

biomolecules from respective layers and appears to be capable of understanding complex biology holistically as well as systematically [1]. Integrated approaches link the exclusive omics data, either in a simultaneous or sequential way, in order to comprehend the interaction of molecules [2]. They are helpful in evaluating data flowing towards different levels and therefore assist in linking genotype to phenotype variance. Integrative approaches are capable of studying the biological processes holistically, hence improving predictive accuracy and disease phenotype prognostics and therefore can ultimately help in improved therapy and inhibition [3].

In order to analyze the multifaceted biological methods, integrative methodology should be adopted which links multiomics data for highlighting the functions and interconnections of engaged biomolecules. With the advancement in highthroughput practices and accessibility of multiomics data, numerous promising methods and tools have been established for data analysis, interpretation, and integration. Multiomics integration approach has been reviewed and comprehensively reviewed in studies related to animals [4], humans [5] microbes [6], and their combinations [7]. In contrast, multiomics integration in plants is difficult owing to metabolic assortment, poor annotation of large genomes, and occurrence of various symbionts having complex networks for interaction. Several inclusive analyses are present particularly on plant multiomics integration and their practical use in precision plant breeding, green systems biology, and other biological and biotechnological applications [8]. Moreover, the development of highthroughput techniques and considerable multiomics data ahead of broad data biology might remain vast, and possibly troublesome for the inexperienced investigators. Omics data derived from inadequately categorized species are frequently feed to the software exclusive of any appropriate curation and unaware of technology's constraint, that might result in inaccurate interpretations.

The integration of data related to multi-omics draw holistic knowledge of biological methods and disorders comes with several tests. Heterogeneity in distinct omics data, considerable sets of data leading towards computation of thorough study, and absence of correct information which assists in highlighting different tool sets and software make analysis and integration of multiomics data a challenging job. Multiomics data are generated using variety of programs, and therefore formats and storage of data show considerable differences. Tools for analysis of multiomics integration require particular data, and hence individual omics data needs proper preprocessing.

1.1. Advanced Techniques in "Omics"

Advancements in omics technologies have led to tremendous success in molecular characterization of wide range of complicated human disorders for instance, cancer. Analysis of multiomics integration employing genomics, proteomics, epigenomics, and metabolomics, is attributing advancement of precision medicine in clinical setups. Several patho-mechanisms of cancer development, treatment resistance and risk factors have been revealed and the information has been used in taking right treatment decisions. The application of integrated omics analyses has been limited which has impeded the possible miraculous developments in the diagnostic and research avenues. A lot of efforts are still needed to improve the methodical infrastructure for commendable generation, evaluation, and annotation of multiomics integration data. The proper utilization of this modern approach will lessen the burden of wrong treatments against any particular molecular disease [9].

1.2. Omics-Driven Targeted Therapy

Some latest findings have potentiated the significance of omics-driven targeted treatment in patients suffering from non-small cell lung cancer. The comparison of results between patients suffering from lung adenocarcinoma who were treated by using omics driven directed therapy and individuals who received typical therapies supports this modern approach of decision making. FFT model (fast and frugal decision tree model) was established to estimate the effects of omics-based approach on the treatment of lung cancer patients. Omics-driven therapy decisions were positively related with better overall survival rate and progression free survival of the patients [10].

1.3. Meta-omics

The integrated OMIC discipline is crucial for deep understanding of a disease and thus is helpful in combating many diseases. Gut microbiome shows a significant influence on human well-being and illness and modern expansion of omics methodologies, involving shotgun metagenomics, phylogenetic microbiome profiling based on markers, metatranscriptomics, metagenomics, and metaproteomics, has led to the proficient representation of microbial communities. Modern omic methodologies provide deep taxonomic resolution of the microbial taxa, reveal the taxa specific functions and metabolic actions within an intricate microbiome. Applications of metaomics techniques to clinical samples of microbial diseases have led to the identification of disease-causing microbial species and culprit metabolites. This identification helps the

CHAPTER 10

Single Cell Omics

Erum Dilshad[1,*], Amna Naheed Khan[1], Iqra Bashir[1], Muhammad Maaz[1], Maria Shabbir[2] and **Marriam Bakhtiar[1]**

[1] *Department of Bioinformatics and Biosciences, Faculty of Health and Life Sciences, Capital University of Science and Technology (CUST), Islamabad, Pakistan*

[2] *Healthcare Biotechnology, Attaur Rehman School of Applied Biosciences, National University of Science and Technology, Islamabad, Pakistan*

Abstract: Recent advances are nowadays providing opportunities to examine the complexities of organs and organisms at the single-cell level. The conventional cell-based analysis mainly examines the cellular processes from the bulk of cells but single-cell omics provides a more detailed insight into individual cell phenotypes, thus giving a link between the phenotype and genotype of cells. Single-cell analysis can be performed at genome, epigenome, transcriptome, proteome and metabolome levels and thus makes it possible to come across mechanisms not seen during the sequencing of bulk tissues. Researchers need to isolate single cells before the initiation of single-cell analysis. For this, various strategies like FACS, MACS, LCM, micro-manipulation and micro-fluids are used for cell isolation depending upon their physical properties and cellular biological characteristics. The analysis of single-cell data at multiple levels gives us an unusual view of multilevel transformation at the single-cell level and thus providing a better chance to discover novel biological processes. High throughput analysis of single cells at genome, transcriptome and proteome levels provides unique and important insights into cell variability and diverse processes like development, genetic expressions and severity of different symptoms in disease pathogenesis.

Keywords: Metabolomics, Omics, Proteome, Single-cell, Transcriptome.

1. INTRODUCTION

All living beings are composed of several groups of individual cells; thus, a cell can be termed a basic entity of life. The study to explore the phenotypes of different cells and the capability to observe the comportment of organs and organisms at the single-cell stage is imperative to develop and comprehend the evolving practices of these cell communities and is essential to understand the functions of diverse biological systems [1].

[*] **Corresponding author Erum Dilshad**: Department of Bioinformatics and Biosciences, Faculty of Health and Life Sciences, Capital University of Science and Technology (CUST), Islamabad, Pakistan; Tel: 9251111555666-324; E-mail: dr.erum@cust.edu.pk

Acquiring single cells population from a heterogeneous population of tissues or cells is called single-cell omics. These populations of cells can be of various cellular states, to study either their normal development or disease mechanisms. Single-cellomics can be analyzed at genome, epigenome, transcriptome, proteome and metabolome levels with specific approaches; applications for each of the technology are mentioned in Table **10.1** [2].

Table 10.1. Single Cell-Omics Techniques with approaches and applications [2].

Life Level	Technique Level	Approaches	Applications
Single Cell	Genome	PCR, MALBAC, MDA	CNV, SNV, indel
	Epigenome	ATAC, Chlp, RRBS	Methylome, histone modifications
	Transcriptome	FISH, RT-PCR, sequencing	Cell states and types
	Proteome	Barcoding, Western blotting, staining	Cellular functions
	Metabolome	Competition assays	Genotype-Phenotype correlations

1.1. Single-cell genomics

The branch of science that deals with the all-inclusive study of genome and genome culture of a targeted sample organism is said to be genomics. The analysis of any specific cell within a tissue sample by using omics procedures is known as single-cell genomics. The single-cell genome sequencing technique is used to observe the physical features including the structural or morphological dissimilarities, aneuploidies, mutations, and genome recombination [3, 4]. The single-cell genomics technique is important for the revelation of various interactions between genetic mechanisms and cell lineage in standard and diseased tissues [5, 6]. For creating an effective treatment strategy for cancers, a precise estimation of prognosis is necessary, thus single-cell genomics has helped in allowing many new prognostic elements to be identified and confirmed. Breast cancer is the best example whereby the basis of propagated cells of a tumor in breast cancer was traced [7] and thus, now, single-cell technology is the best tool to provide a prognosis, unlike in the past [8]. The significance of SCG can be illustrated in Fig. (**10.1**).

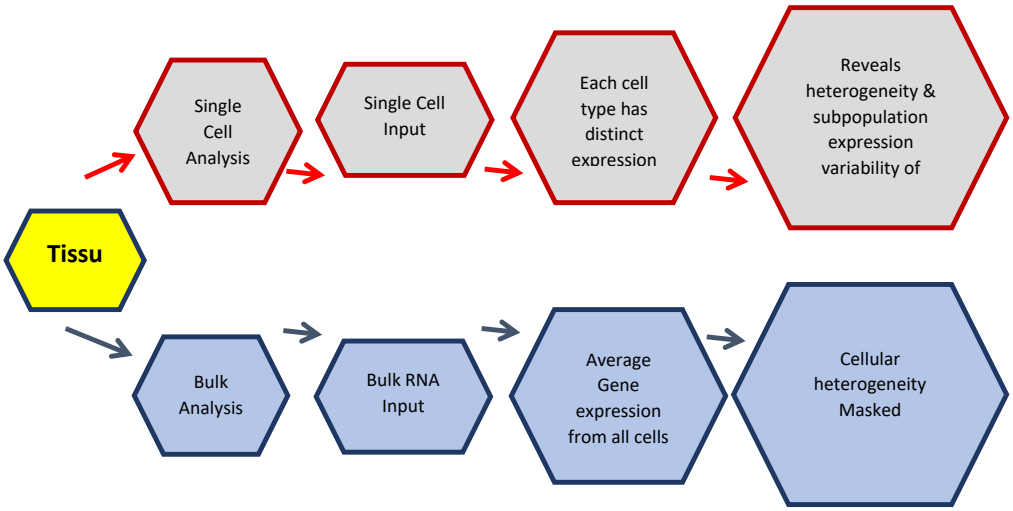


Fig. (10.1). Significance of Single Cell Genomics.

Single-cell transcriptomics

Transcriptomics is an extensive study of entire genome transcriptions for a specified sample living cell, tissue, body part, or an organism as a whole at a certain time of a developmental phase or a specified physiological situation. It is a persuasive technology used to unveil the distinct gene expressions and various RNA-associated configurations through diverse, initial embryonic development and then reprogramming. The genomics technology is also able to link a cell's genotype to its phenotype, and that is the reason for the detection of hundreds of transcripts in an array of cells and tissues [9, 10]. Single-cell transcriptomics has been functional in a range of infections including tumors, cancers, and inflammatory diseases.

The three most widely used methods for computing single-cell gene expression are single-cell quantitative PCR (qPCR), mRNA in situ hybridization, and single-cell RNA sequencing. Single-cell RNA-seq or common scRNA-seq can define the gene expression network in marked cells by combining it with over knockdown, expression, or knockout of a gene of interest [11 - 13]. Furthermore, the tools can also be used to acquire transcriptomic evidence from intra-tumoral cells and to find out the sub-populaces within a body lump by identifying alleged cancer stem cells. Besides, single-cell RNA-seq is considered a favorable technique to boost clinical diagnosis and prediction, and thus is feasible to provide a realistic plus perfect target therapy [14, 15].

CHAPTER 11

Pharmacogenomics

Shumaila Azam[1], Sahar Fazal[1,*], Attiya Kanwal[2], Muhammad Saad Khan[5], Narjis Khatoon[1], Muneeba Ishtiaq[1], RabbiahManzoor Malik[1, 3], Sana Elahi[1] and **Fakhra Nazir[1, 4]**

[1] *Department of Bioinformatics and Biosciences, Capital University of Science and Technology, (CUST), Islamabad, Pakistan*

[2] *International Islamic University Islamabad (IIUI), Islamabad, Pakistan*

[3] *Wah Medical College, WahCantt, Pakistan*

[4] *Centre for Bioresource Research, Islamabad, Pakistan*

[5] *Department of Biosciences, COMSATS University Islamabad, Sahiwal, Pakistan*

Abstract: The ongoing development in new genotyping methods necessitates an understanding of their potential benefits and limits in terms of pharmacogenomics utility. We give an overview of technologies that can be used in pharmacogenomics research and clinical practice in this chapter. The Human Genome Project's completion has paved the way for the development of clinical instruments for patient evaluation. Pharmacogenomics may enable the identification of patients who are most likely to benefit from a specific drug, as well as those for whom the expense and risk are greater than the advantages. Both drug therapy's safety and efficacy may improve. In the future, genotyping may be used to tailor drug treatment for large groups of individuals, lowering drug treatment costs and improving therapeutic efficacy and overall health.

Keywords: Clinical Pharmacogenetics Implementation Consortium (CPIC), Dutch Pharmacogenetics Working Group (DPWG, Pharmacogenomics (PGx), PharmGKB, US Food and Drug Administration (FDA).

1. INTRODUCTION

Scientists have been attempting to recognize the causes of ailments at the genetic level since Mendel's discovery of genes in 1865 and the Human Genome Project (HGP) in 2003 [1]. Following the completion of the HGP, new projects have continued to use the genetic sequencing data obtained from that project [1]. We now have more knowledge than ever before about how our genes influence development, growth, health, and even drug metabolism [2]. In 2003, The

* **Corresponding author Sahar Fazal**: Department of Bioinformatics and Biosciences, Capital University of Science and Technology, Islamabad, Pakistan; E-mail: sahar@cust.edu.pk

Syeda Marriam Bakhtiar and Erum Dilshad (Eds.)

National Human Genome Research Institute instigated the project “Encyclopedia of DNA Elements” [3] the aim of which is to locate all coding elements/sequences in the human genome [3]. In many inherited diseases, next-generation sequencing has established a potent method in disease-related variants identification, whereas to detect variants in previously identified genes associated with disease, selective sequencing is found to be useful [4, 5].

The Cancer Genome Atlas [6] has helped us better understand the molecular markers associated with many types of cancer. The TCGA has also proven valuable in detecting molecular similarities between different cancer patients and types of cancer, as well as in documenting the heterogeneity of all cancers [6]. In addition, a rising number of organizations and universities are creating pharmacogenetics testing clinical facilities and support infrastructures [7, 8]. The Pharmacogenomics Research Network (PGRN) is made up of three major center-grant programs with the goal of improving precision medicine by identifying and understanding genetic variants that influence therapeutic effects and cause adverse medication effects in patients [9]. Two of the three PGRN projects aim to make pharmacogenomics and precision medicine services more accessible [9]. To obtain information, The PharmGKB base (https://www.pharmgkb.org/) and a PGRN hub (http://www.pgrn.org/) have been developed to coordinate the activities of the new PGRN [9]. The scientific community's access to genomic data is the major hurdle to full adoption of pharmacogenomics [10]. ClinVar [11] is a free resource of papers and supporting information on the connections between human genetic variations and phenotypes. This database offers information on how human genetic diversity affects patients' health [11]. ClinVar works with the Clinical Genome Resource (ClinGen; https://www.clinical genome.org/) to determine the clinical importance of genes and their variations in precision medicine and research [12]. The NOMAD registry is a database of healthy people's genetic information that allows researchers to discuss and exchange their discoveries. Thanks to breakthroughs in genomics, scientists can now generate therapeutic therapies and diagnostic tests at a faster rate than ever before [13, 14].

Due to the ability to sequence, DNA and RNA at a faster rate and at a low cost, clinicians can efficiently detect and identify rare disorders [4, 5, 14]. Rare disease patients may have to wait a long period for an accurate diagnosis [15, 16]. Rather than putting patients in the dark with delayed therapy for their sickness, genetic testing gives them peace of mind and hope [17, 18]. The National Institutes of Health will launch “All of Us,” a new programme in the spring of 2018 that will collect data from at least one million people in the United States in order to speed up research and improve health. This programme will have a substantial impact on the future delivery of precision medicine [19]. Data provided through genetic

testing is exceedingly difficult to understand [21]. The biggest challenge, according to the scientific community [21], is to understand the effects after acquiring the DNA sequence and finding a genetic mutation. Genetics' ultimate goal is to discover a link between genetic variation and a specific illness or metabolic pathway [22]. On the other hand, the pool of new variants is enormous, with the possibility of having no established link to a specific condition [22]. As a result, genetics must transform genomic data into information that doctors can utilize to make decisions [22]. Physicians indicated they did not know how to apply pharmacogenetic testing in clinical practice in a recent study [23, 24]. Researchers and scientists in this sector must provide information, guidance, and education to physicians who lack confidence when ordering pharmacogenetic testing in order to bridge the gap between science and healthcare [20].

Before prescribing a treatment based on a single biomarker, a physician must have a thorough understanding of the patient (integrated medicine) [20]. In response to the growing need for pharmacogenomics information and guidance, the FDA has released a list of pharmacogenomic biomarkers in drug labelling on its website [25]. Biomarkers identified on the FDA website include differences in the germline or somatic genes (polymorphisms, mutations), genetic etiological functional impairments, variations in gene expression, and chromosomal abnormalities [26]. Some drug labels state what a clinician can do based on biomarkers, which may or may not require pharmacogenetic testing [26]. Specific variants in the genetic code are linked to variances in medication reactivity, the chance of an individual having certain drug side effects, and variations in the rate and degree of drug metabolism. We are at a vital juncture in medical science and pharmacy practice. Pharmacogenomics (PGx) is the study of the connection between the human genome and the science of pharmacology [27]. Most pharmacy schools' curricula have just lately introduced genetics and PGx courses. Previous generations of pharmacists may not have been exposed to the genetic concepts that drive PGx.

The goal of this chapter is to go through fundamental principles in genetics and genetic variation in order to lay the groundwork for comprehending crucial and well-illustrated gene-drug interactions. Wherever possible, we employ drugs with genetic testing guidelines on the labeling or professional clinical practice recommendations from the US Food and Drug Administration (FDA). There will also be several genuine, online PGx references offered. The moment has come for pharmacists and scientists to embrace these developments in drug therapy and prepare to translate gene-drug therapy. The Human Genome Project's findings spawned pharmacogenomics. The Human Genome Project (HGP) began in 1990 as a global effort to identify and comprehend the structure of every human gene. The HGP is without a doubt our lifetime's greatest scientific achievement, having

CHAPTER 12

Biomaterials in Gene Therapy for Soft and Hard Tissues

Sarmad Mehmood[1], Sajjad Haider[2], Muhammad Naeem[3], Raees Khan[4], Muhammad Faheem[3], Bushra Bano[4], Syeda Marriam Bakhtiar[5], Atif Ali Khan Khalil[3], Fazli Subhan[3], Syed Babar Jamal[3,*] and **Adnan Haider[3,*]**

[1] *Department of Pathology, CMH Institute of Medical Sciences, Bahawalpur, Pakistan*

[2] *Department of Chemical Engineering, King Saud University, Riyadh, Saudi Arabia*

[3] *Department of Biological Sciences, National University of Medical Sciences, Rawalpindi, Pakistan*

[4] *Institute of Basic Medical Sciences, Khyber Medical University, Peshawar, Pakistan*

[5] *Department of Bioinformatics and Bioscience, Capital University of Science and Technology, Islamabad, Pakistan*

Abstract: Bone healing and formation are under the control of growth factors. Among these factors, bone morphogenetic proteins (BMPs) have a vital role in bone and cartilage maintenance and formation. BMP itself belongs to the superfamily of transforming growth factor β (TGFβ). Although, the use of recombinant BMPs has no significant association with the treatment of bone fractures, arthroplasty, pseudoarthrosis or other bone-related diseases. Recent advancements in genetic engineering have led to the foundation of gene therapy. Gene therapy uses genes to be incorporated in the living cells to replace defective genes or manipulate gene expression for therapeutic purposes. Gene therapy is emerging for the treatment of diseases with approval in Europe where it is in the marketing surveillance phase (Phase IV Clinical trial). This technique has also been tested for the incorporation of osteogenic genes in stem cells for repairing spinal fusion and recovering defects in bones in preclinical models. Therefore, gene therapy has the potential for the treatment of different diseases and has the advantage over the use of recombinant proteins. In this chapter, we have discussed gene therapy, its mechanism, delivery system and its use in tissue engineering (soft and bone tissue) for clinical application.

Keywords: Gene Delivery, Gene Therapy, Scaffolds, Polymeric Composites.

[*] **Corresponding authors Adnan Haider and Syed Babar Jamal:** Department of Biological Science, National University of Medical Sciences, Rawalpindi, Pakistan; E-mail: adnan.haider@numspak.edu.pk and Department of Biological Science, National University of Medical Sciences, Rawalpindi, Pakistan;
E-mail: babar.jamal@numspak.edu.pk

Syeda Marriam Bakhtiar and Erum Dilshad (Eds.)

1. INTRODUCTION

Many clinical conditions have characteristic features of bone growth under the stimulus of osteogenic cytokines and proteins [1]. In the last decade, *in vivo* investigations have shown that growth factors can stimulate bone healing and bone formation [2]. Osteogenic cytokines are studied recently for understanding the osteoinductive ability of demineralized bone matrix [3], along with bone morphogenetic proteins (BMP) interactions with muscle bone [4]. BMPs are members of the transforming growth factor (TGF) superfamily and play an important role in bone and cartilage maintenance and development. Recent studies have revealed that the BMP pathway has also been associated to adult skeletal muscle mass control. Therefore, BMPs come forward as an essential player involved in the regulation of both homeostasis and muscle formation. BMP-2 and BMP-7 have already been approved for the treatment of spinal fusion and non-union fractures [5]. Recombinant BMPs presented higher healing rates and lesser infections with reduction in failure risk [6], also exhibiting a higher fusion rate contrary to autograft [7]. On the other hand, protein delivery has shown potential in the field of bone tissue engineering. Application of the recombinant proteins are often challenged regarding delivery as this protein based therapy is not feasible in terms of shorter half-life of protein along with poor retention in the defected site [8]. Moreover, expensive and high doses are required for protein delivery contrary to conventional bone repairing methods [9]. Apart from that, the use of recombinant BMPs, bone loss, and delayed bone healing are all linked to revision joint arthroplasty, tumor excision, spine and trauma. Keeping the aforementioned hurdles, it is pivotal for the researchers working in the field to look for the alternate method through which bone repair and formation can be stimulated.

Recent advancement in gene therapy has gained much popularity and attraction as the first gene therapy based product (Glybera®, uniQure, Amsterdam, The Netherlands) was approved in Europe and proved successful in post marketing surveillance phase (Phase IV Clinical trial) [10]. Glybera®, a gene delivery system created by Amsterdam-based uniQure, has been approved for patients with a rare lipid processing impairment known as lipoprotein lipase deficiency (LPLD), which affects only 1 to 2 people in a million [11]. Gene therapy has shown to have a promising potential in overcoming hurdles which occur in conventional protein-based delivery system. Different animal studies have highlighted the potential of gene therapy to be capable of delivering osteogenic drug molecules to precise anatomical levels for a longer period of time [12]. In preclinical models, genetic engineering of stem cells with the addition of osteogenic genes showed promising outcomes in repairing spinal fusion, advanced fractures, and the recovery of defective bones [13].

Gene delivery exploits protein synthesis by using natural cellular machinery which is a central dogma of life. Furthermore, endogenously synthesized proteins may have greater efficacy as compared to recombinant and exogenous counterparts [14]. Nucleic acid delivery to the site provides osteogenic growth factors with an increase in retaining time at the site [15]. In this book chapter, we have discussed in detail about the mechanism and current strategies opted for bone tissue engineering including both viral and non-viral delivery techniques providing future prospective on the potential use of tissue engineering *via* gene therapy in clinical use.

2. MECHANISM OF GENE THERAPY

In gene therapy, gene is transferred to suitable cells which ultimately penetrates the nucleus. Vectors are required for carrying genes and protecting them from negative charges of the genes and nucleases. Gene delivery is broadly classified into viral and non-viral categories depending upon the type of vector used, with every type having its own benefits and limitation, for example, viral vectors exhibited higher transfection efficiency along with toxicity and immunogenicity making patients vulnerable to the side effects. However, non-viral vectors have shown lesser transfection efficiency with non-immunogenic and safe results.

2.1. Delivery of Gene for Gene Therapy

2.1.1. Gene Release

In the gene therapy process, an abnormal copy of gene, which is responsible for a particular disease, is replaced by the normal copy of the gene. Releasing of genes into stem cells is the major hurdle in gene therapy. Vectors, which are efficient, specific, able to be purified in higher concentrations and non-immunogenic are used for releasing genes or a number of genes depending upon the requirement into stem cells. Vectors must not trigger inflammatory responses and should rectify deficiencies, inhibit harmful activities and increase normal processing of cells. Furthermore, it should be safe for both persons being administered with the gene and the person administering the gene of interest harboring vector. Conclusively, the vector should be able to express transgene for patient's entire life in general [16].

2.1.2. Hematopoietic Stem Cells and Gene Therapy

Because of their ability of self-renovation and longevity, hematopoietic stem cells have proved ideal for gene therapy. Using gene therapy technology, induced

CHAPTER 13

Induced Pluripotent Stem Cells

Ambrin Fatima[1,*], **Uzma Abdullah**[2] and **Zafar Ali**[3]

[1] *Department of Biological and Biomedical Sciences, The Aga Khan University, Karachi, Pakistan*

[2] *University Institute of Biochemistry and Biotechnology, PMAS-Arid Agriculture University, Rawalpindi, Pakistan*

[3] *Centre for Biotechnology and Microbiology, University of Swat, Swat, Pakistan*

Abstract: A limiting factor for the identification of disease mechanisms and development of new therapies has been the access to a model system/s that can faithfully recapitulate key features of the disease and more precise clinical translations of new treatments. Stem cells in this regard are very promising, but the ethical issues related to totipotent embryonic stem cells and functional constraints to unipotent somatic stem cells have led to focus on induced pluripotent stem cells to avoid both functional and ethical constraints. The introduction of human Induced Pluripotent Stem Cell (iPSC) technology provides a model system to replicate diseases in humans. In this technology, human somatic cells can be "reprogrammed" by the transgene expression of four transcription factors into stem cells called iPSC. In this chapter, it will be discussed how iPSCs can be used for disease modelling, drug discovery and regenerative medicine.

Keywords: Disease modelling, Embryonic Stem Cells, Induced Pluripotent Stem Cells.

1. HUMAN EMBRYONIC STEM CELLS

A stem cell is defined by two key characteristics namely; self-renewability (unlimited potential to make copies of itself) and differentiability (give rise to almost any human cell type). Human embryonic stem cells (hESCs) are pluripotent stem cells derived from the inner cell mass of embryos (blastocyst stage 4–5 days post fertilization, at the stage of 50-150 cells) [1]. hESCs are mostly obtained from donors after they are deemed unsuitable for implantation and appropriate consent has been obtained. For the isolation of hESCs, the inner cell mass is separated from the trophectoderm and initially plated onto a layer of feeder cells (non-dividing cells that provide growth factors, adhesion molecules, and extracellular matrix) to form an hESC cell line [1, 2].

[*] **Corresponding author Ambrin Fatima**: Department of Biological and Biomedical Sciences, The Aga Khan University, Karachi, Pakistan; E-mail: ambrin.fatima@aku.edu

Syeda Marriam Bakhtiar and Erum Dilshad (Eds.)

The first hESC line was developed by James Thomson in 1998, from the cells derived from surplus embryos after *in vitro* fertilization [3]. The hESCs are pluripotent, and under optimal laboratory conditions can be differentiated *in-vitro* into any cell type of the body. The pluripotency and unlimited capacity for self-renewal of hESCs have created a huge interest both in disease modelling and as possible tools for regenerative medicine (Fig. **13.1**). However, this excitement has been attenuated due to ethical and political considerations as well as few scientific limitations that have reduced the scientific progress in the field of hESCs.

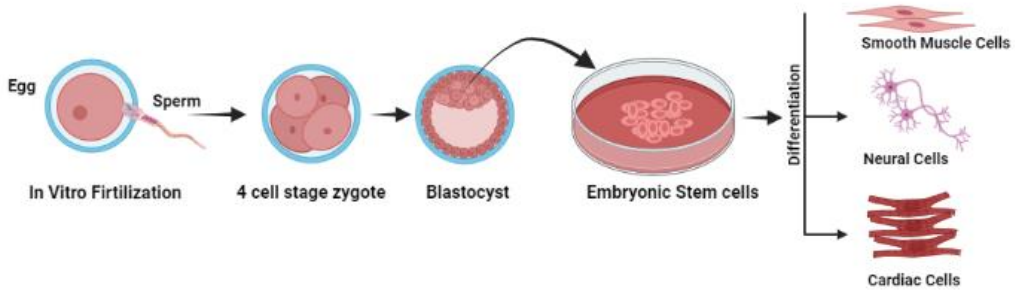


Fig. (13.1). General scheme of development of embryonic stem cell lines and their pluripotency.

New doors were opened with the development of methods for reprogramming somatic cells with properties akin to ESCs. These are called induced pluripotent stem cells (iPSCs). This technology has changed the prospects of the field of stem cell and regenerative medicine.

1.1. Human Induced Pluripotent Stem Cells: Origins and Properties

The modern work on iPSCs is built on the pioneering work of Sir John Gurdon [4] [5], a British developmental biologist, who showed programing of somatic nucleus to a pluripotent state through the process of somatic cell nuclear transfer. In 2006, two Japanese researchers Shinya Yamanaka and Kazutoshi Takahashi extended Goudon's work and attempted to generate mouse stem cells in the laboratory conditions. They achieved this by employing forced and transgenic expression of transcription factors (*i.e. Oct 3/4*, *Sox2*, *c-Myc* and *Klf4*) in murine dermal fibroblasts [6]. It was the first proof of concept and received with huge enthusiasm. The same process of generating human iPSCs (hiPSCs) was repeated in 2007, this time human fibroblasts were used for generating hiPSCs [7].

This time, through virus mediated over-expression of a human ortholog by using Yamanaka factors (OCT4, SOX2, KLF4 and c-MYC) (Fig. **13.2**), they were able to induce pluripotency in cultured dermal fibroblasts. The generated pluripotent

stem cells which were highly analogous to human embryonic stem cells in terms of transcriptional activity, epigenetic patterns and morphology [7 - 9]. Therefore, these cells were termed as induced pluripotent stem cells (iPSC). In 2012, Shinya Yamanaka and John Gurdon received the Nobel Prize in physiology or medicine for their work of reprograming somatic cells into stems cells that are capable of differentiating into any cell of the body.

The iPSCs technology offers a substitute of accessible tissue types and thus enormous possibilities in research and therapeutics. However, it also has certain inherent challenges, such as all cells do not complete this forced reprogramming journey, some remain in a partially reprogrammed state. Moreover, the stressful reprogramming process, involving the increased expression of oncogenes like KLF-4 and c-MYC, can introduce gene chromosomal anomalies. Therefore, reprogrammed iPSCs are subsequently screened for morphology, pluripotency, chromosomal stability, genetic aberrations and differentiation potential [9].

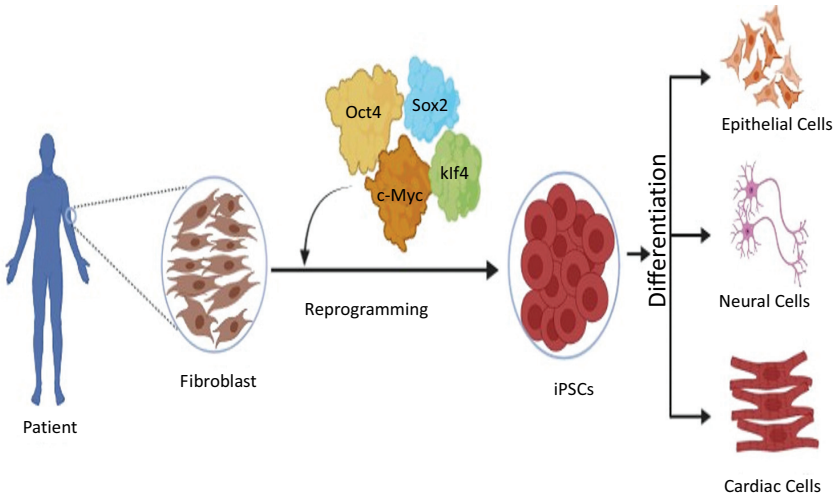


Fig. (13.2). A schematic representation of the human iPSC generation from patient fibroblast by introduction of cocktail of reprogramming factors, such as (OCT4, SOX2, KLF4, and C-MYC) and pluripotency of generated iPSCs.

1.2. Applications of Induced Pluripotent Stem Cells

Surpassing initial expectations, 15 years since discovery, iPSCs have revolutionized biology and proved an invaluable tool for disease modelling, developmental biology, regenerative medicine and drug discovery and has proven its universality for cellular reprogramming in different somatic cells.

CHAPTER 14

Hemoglobinopathies

Mahnoor Asif[1], Sadia Nawaz[2,*] and **Muhammad Tariq[1]**

[1] *National Institute for Biotechnology and Genetic Engineering College, Pakistan Institute of Engineering and Applied Sciences (NIBGE-C, PIEAS), Faisalabad, Pakistan*

[2] *Institute of Biochemistry and Biotechnology, University of Veterinary and Animal Sciences-UVAS, Lahore-54000, Pakistan*

Abstract: Hemoglobinopathies are a group of inherited blood disorders characterized by compromised hemoglobin function. Hemoglobin is a 64kDa protein, consisting of four globin polypeptides each containing one heme molecule; blood acquires its red color from this heme molecule. Two of the four polypeptide chains are α-globin chains, whereas the other two are β and γ chains during adult and fetal life, respectively. Hemoglobin carries oxygen to respirating cells and tissues in vertebrates and defects in genes encoding this protein result in a variety of disorders, ranging from mild asymptomatic to severe fatal phenotypes. This chapter reviews various hemoglobinopathies underlying mutations in globin genes. We also provide a brief note of the traditional and contemporary diagnostic approaches and screening, both prenatal and postnatal, with a specific focus on recent advances in this regard. We have summarized various therapeutic strategies, from transfusion and iron chelation to CRISPR-driven genome editing aimed at reactivating fetal hemoglobin in adults. The chapter concludes with a brief account of the future challenges and prospects for developing a therapy for hemoglobinopathies a clinical reality.

Keywords: CVS, Hemoglobin, Hemoglobinopathies, Sickle cell anemia, Thalassemia.

1. INTRODUCTION

In 1840, Friedrich Ludwig Hünefeld accidentally discovered what he referred to as plate-like crystals in the dried blood of humans or pigs [1]. Later, Hoppe-Seyler named these crystals Hemoglobin (Hb) [2]; soon afterwards, Claude Bernard recognized them as oxygen-carrying molecules [3]. The discovery of the detailed 3-D structure of this molecule by X-ray crystallography took more than 100 years [4]. For this long-awaited discovery, Max Perutz, along with Sir John Kendrew, won the 1962 Nobel Prize in chemistry. A hemoglobin molecule consists of four

[*] **Corresponding author Sadia Nawaz**: Institute of Biochemistry and Biotechnology, University of Veterinary and Animal Sciences-UVAS, Lahore, Pakistan; E-mails: sadia.nawaz@uvas.edu.pk and shamiryar@hotmail.com

Syeda Marriam Bakhtiar and Erum Dilshad (Eds.)

polypeptide chains, two alpha (α) chains, each comprising 141 residues and two beta (β) chains, each containing 146 amino acids. A polypeptide chain combined with a heme constitutes a monomer or a subunit of hemoglobin. The complete molecules consists of four subunits, closely joined as in a 3D jigsaw puzzle, forming a tetramer [5]. The structure of hemoglobin is strictly conserved to maintain the affinity of Hb-O_2 [6]; aberrations during the synthesis of α or β chains result in hemoglobinopathies, a family of autosomal blood disorders, inherited in recessive fashion. Thalassemia is highly prevalent in Far East, India and in the Mediterranean areas, whereas Sickle Cell Disease (SCD) is prevalent in sub-Saharan Africa [7].

Their spread across different parts of the world, especially in the US and Europe is largely the result of migration. Globally, hemoglobinopathies have one of the highest incidences among monogenic disorders, with ~5% people carrying a disease allele. Some of the hemoglobin variants cause disease only in certain circumstances or under stress, while other variants are highly penetrant [8, 9].

1.1. Structure and Genetics of Hemoglobin Synthesis

Hemoglobin is a 64-64.5 kDa protein, carries oxygen to respirating cells and tissues, in all vertebrates except Channichthyidae [10]. Human hemoglobin consists of four globin (peptide) chains, each containing one heme molecule; blood acquires its red color from this heme molecule (Fig. **14.1**). Heme molecule is a porphyrin ring of carbon, hydrogen and nitrogen, containing an iron atom in the center like a jewel. The globin chains are produced differentially during ontogenesis; hence, these are different at various stages of life *i.e.*, embryonic, fetal and adult [11]. The adult hemoglobin molecule (HbA) has two β- and two α-globin chains whereas the predominant hemoglobin during fetal life is HbF containing two gamma (γ) globulin chains instead of β chains.

The switching of Hb F to Hb A takes place within 30 to 40 weeks of conception [12] and the transition is completed by six months postnatal life, with around 1% Hb F, ≤ 3% Hb A2 (α2δ2) and 97% Hb A (α2β2) in a normal adult [13]. This switching between different hemoglobin reflects the physiological adaptation according to differing oxygen requirements during ontogenesis [12].

Globin genes are located on chromosome 11 and 16 in two remarkably conserved clusters (Fig. **14.2**). Alpha gene cluster is located on 16p13.3 whereas beta gene cluster is present on 11p15.4 [14, 15]. The mechanism underlying the expression regulation of these globin genes is not completely understood, but it is seen that at the 8^{th} week of embryonic development, the two chains epsilon (ε) and zeta (ζ) are switched off in definitive cells and cannot be reactivated. Gamma (γ) genes

undergo independent silencing, analogous to ε gene with some degree of competition, by modification in relative transcription, between β- and γ-globin genes. This relative transcription changes during development, with balance shifting from the expression of γ-globin in the fetus to β-globin gene in adults [12]. More than 1000 hemoglobinopathies have already been reported, most of which are asymptomatic. The most clinically relevant of these are SCD, Thalassemias, Erythrocytosis, Cyanosis, and Hemolytic anemias [16]. Thalassemia (208 million patients including ~4.7 million severely affected [17] and SCD (with around 3.2 million homozygous and 46 million heterozygous patients) [16] are the most common hemoglobinopathies.

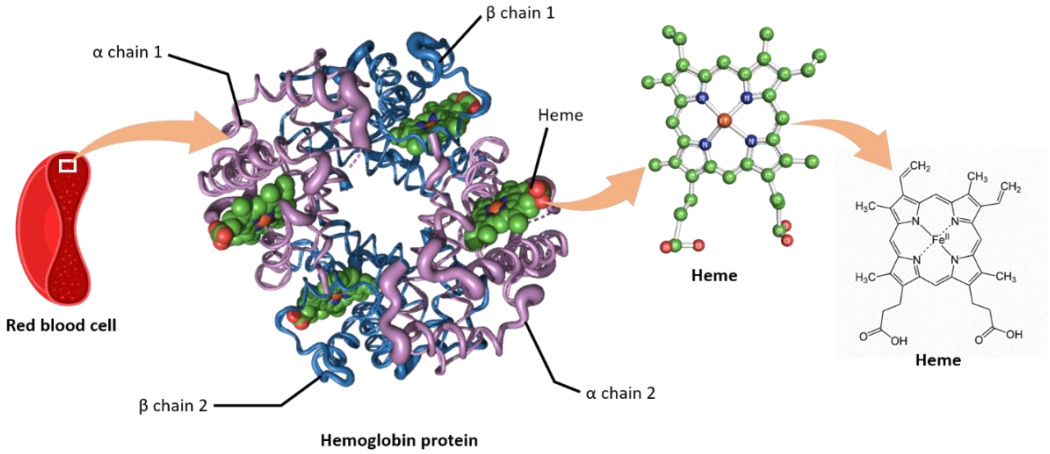


Fig. (14.1). Structure of Hemoglobin Molecule.

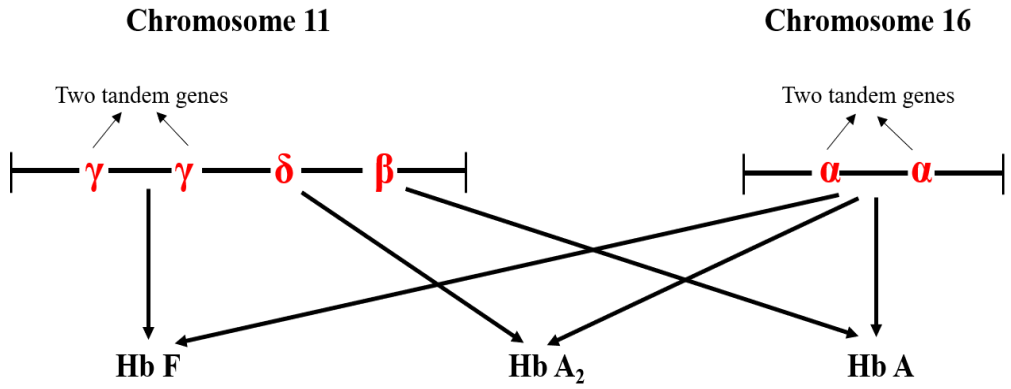


Fig. (14.2). Normal adult hemoglobin.

CHAPTER 15

Metabolic Syndromes

Mahnoor Ejaz[1,*], Areena Suhail Khan[2], Faiza Naseer[3] and **Alvina Gul[1]**

[1] *Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), Islamabad, Pakistan*

[2] *Department of Biosciences, COMSATS University Islamabad, Park Road, Islamabad, Pakistan*

[3] *Shifa College of Pharmaceutical Sciences, Shifa Tameer-e-Millat University, Islamabad, Pakistan*

Abstract: Metabolic Syndromes (MetS) are recognized as a cluster of risk factors which are known to increase the likelihood of obesity, type 2 diabetes (T2D) and cardiovascular disorders (CVDs). It is significant to understand disease pathology in order to discover a pathological mechanism leading to the development of MetS. Elevated triglycerides, increased blood pressure, hyperglycemia (increased blood glucose levels), low levels of High-density lipoprotein (HDL) cholesterol and elevated waist circumference are key parameters in diagnosing MetS. Various therapeutic interventions have been developed for treating metabolic diseases like polypills which are commonly known as combination pills, along with the fixed dose combinations. In addition to pharmacological handling, surgical treatment is also showing success in treating MetS such as Bariatric treatment. With the emerging experimental techniques, gene therapy allows the replacement of a defective gene with a healthy one, which may eventually reverse the disease. Leptin Gene Therapy, ZFN Gene Editing, CRISPR/ Cas9 genome editing are different platforms of gene therapy which are showing promising results in treating the metabolic disease. Novel experimental approaches and pharmacological treatments can provide a better insight into metabolic syndrome and its related complications, thereby reducing its global burden.

Keywords: CRISPR/Cas9 genome editing, CVDs, Metabolic Syndrome, Obesity, Polypills, Type 2 Diabetes.

1. INTRODUCTION

Metabolic syndrome (MetS) emerges as an epidemic and a vital public well-being concern. It is characterized not as a disorder, but as a common entity which includes a group of metabolic risk factors which are likely to increase the chances of type 2 diabetes (T2D) and cardiovascular disorders (CVD) [1]. Various other

[*] **Corresponding author Mahnoor Ejaz**: Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), Islamabad, Pakistan; E-mail: mahnoorejaz@hotmail.com

comorbidities (like proinflammatory/prothrombic state, NAFLD, cholesterol gallstone disease, reproductive disorders) are known to be associated with MetS [2]. Several lifestyle factors including overconsumption of food, inactive lifestyle and abdominal obesity may cause MetS [1].

Although the presence of a unifying pathogenic mechanism that can help understand the disease pathophysiology is still unclear, it is extremely likely that insulin resistance and abdominal obesity play a vital role in promoting the development of MetS since studies have suggested a two-fold increased risk of CVD and a five-times increased T2D risk [3].

A number of terms are incorporated to define metabolic syndrome, and almost all of them include hypercholesterolemia, central obesity, hypertension, decreased HDL-cholesterol levels, resistance to insulin and increased plasma triglycerides [1, 4].

1.1. Background

MetS begin as a notion rather than a diagnosis. In 1920 a Swedish physician, Kylin observed a relationship among hypertension, hyperglycemia and gout, and later found visceral obesity linked to CVD and T2D [5]. In 1988, Reaven described Syndrome X as the group of risk factors for T2D and CVD- an addition of the theory of insulin resistance [6]. In 1989, Kaplan retitled the syndrome "The Deadly Quartet" adding obesity or visceral obesity as a major abnormality [7]. The syndrome was renamed once again as "The Insulin Resistance Syndrome" in 1992 [8].

1.2. Components of MetS

The idea of metabolic syndrome has numerous empirical applications, one such is the clinical assessment to categorize individuals at a greater risk of obesity, T2D or CVD as shown in Fig. (**15.1**). Therefore, it helps identify a particular subgroup of patients with a common pathophysiology. Thus, the term assists as a shorthand for clinicians for the pooled underlying biological processes [9].

1.2.1. Obesity

Various health experts assert that if obesity had not been a major public health issue that it is today, MetS would never have been put forth. By 2030, about 20% of world's population is estimated to be obese [10]. The elevated prevalence rate of obesity is linked to increase in insulin resistance and MetS [11]. The loss of insulin secretory capacity (β-cell function loss or reduced β-cell mass) along with

genetic, social and environmental factors is identified to be involved in the development of the disease [12 - 15]. This universal burden insists fundamental nutrition and lifestyle modifications that are based on a sound comprehension of MetS pathology [16]. To represent abdominal obesity, measurement of body mass index (BMI) and waist circumference (WC), though widely utilized, are not the specific parameters. However, a more accurate predictor is visceral fat area (VFA) which has shown to be more strongly correlated with MetS [17].

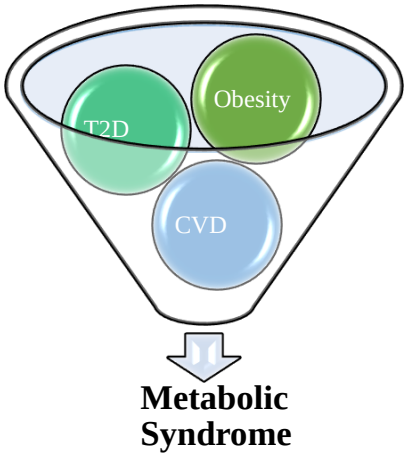


Fig. (15.1). Components of MetS. Individuals at a risk of obesity, T2D and CVD are more likely to develop MetS.

1.2.2. Type2 Diabetes

In the last 30 years, T2D prevalence has tripled, affecting millions worldwide and is predicted to rise by 54% in the next ten years [10]. T2D and MetS are known to increase the likelihood of CVD and various carcinomas [17]. Elevated blood glucose concentration marks this disorder, where defected insulin secretion is regarded as the key pathophysiological factor [11]. The degree of prevalence of T2D is one of the criteria, which is preset in all the three defined versions of MetS and hence most closely related [18]. MetS individuals have a five times greater incidence to T2D and the presence of insulin resistance has an additive effect increasing the risk 6-7 times [19]. Human body's inability to respond to elevated glucose concentration has a major role in T2D pathogenesis [20]. Loss of skeletal muscles- major insulin utilization and uptake sites increase resistance to insulin thereby elevating risk [17].

1.2.3. Atherosclerotic Cardiovascular Disease (ASCVD)

Various studies have revealed that patients with MetS are more likely to develop ASCVD [21]. It accounts for about 30% of worldwide mortality rate hence a vital

CHAPTER 16

Intellectual Disabilities

Zafar Ali[1,*], **Uzma Abdullah**[2] and **Ambrin Fatima**[3]

[1] *Centre for Biotechnology and Microbiology, University of Swat, Swat19130, Pakistan*

[2] *University Institute of Biochemistry and Biotechnology (UIBB), PMAS-Arid Agriculture University, Rawalpindi. Pakistan*

[3] *Department of Biological and Biomedical Sciences. The Aga Khan University, Karachi, Pakistan*

Abstract: Intellectual disability (ID) is caused by the disruption of neurodevelopmental processes. Its diagnosis and severity are defined in terms of an Intelligence Quotient score of ≤70. ID has diverse presentations and clinical overlaps with other cognitive disorders such as autism spectrum disorder and microcephaly. ID has a diverse etiology encompassing both environmental and genetic insults to the developing brain. The precise diagnosis is challenging but crucial for prognosis and risk assessment for future pregnancies. The suspected cases of genetic ID often follow a strategic series of tests for diagnosis. There is no effective cure for this disorder except in the cases of early diagnosed metabolic disorders. The available therapies are mostly aimed at easing the symptoms and improving the quality of life.

Keywords: Autism spectrum disorder, Genetic diagnosis, Global developmental delay, Intelligence Quotient, Microcephaly, Neurodevelopment, Supportive therapy.

1. INTRODUCTION

Human brain development is a complex process involving cellular proliferation, differentiation, migration, and integration into a cohesive circuitry. All of these processes are meticulously orchestrated and lead to a highly specialized human brain, capable of processing complex language, cognition, and emotion. On the other hand, any abnormality in these cellular processes results in neurodevelopmental disorder [1]. Intellectual Disability (ID) previously called mental retardation is a neurodevelopmental disorder and manifests as deficient cognitive functioning, adaptive behavior, and social skills compared to individual's age group. It mainly affects three different aspects of an individual's

* **Corresponding author Zafar Ali:** Centre for Biotechnology and Microbiology, University of Swat, Swat 19130, Pakistan; E-mail: zafaralibiotech@gmail.com

Syeda Marriam Bakhtiar and Erum Dilshad (Eds.)

personality; the conceptual aspects such as language, reading, writing, math, reasoning, knowledge, and memory; the social aspects that include empathy, social judgment, interpersonal communication skills, ability to make and retain friendships; the practical aspects concerning personal care, job, money handling, and organizing scholastic tasks.

The terminology used for ID is evolving; "Global Developmental Delay" is being used for children aged less than 5 years having a delay in more than one area of development namely acquisition of motor skills, speech and language, cognition, personal-social activities of daily living [2]. As these delays may be transient, around 2/3 of children diagnosed with global developmental delay may eventually be diagnosed with ID after 5 years of age [3].

The global prevalence of ID is 1-3% in the general population [4], with a higher incidence in males than females. The etiology of ID may be either genetic or non-genetic or environmental factors and both contribute comparably. The non-genetic factors include malnutrition of pregnant mothers or children, infections, toxin exposure, trauma and other obstetric complications, therefore. Genetic causes which account for almost 50% of cases include chromosome abnormalities such as aneuploidies and large deletion or duplications; point mutations and indels (small insertion and deletions) [5]. In either case, the diagnosis can be complicated because of heterogeneity of the disease both clinically and genetically.

2. DIAGNOSIS AND CLASSIFICATION OF ID

The clinical diagnosis of ID is based on an intelligence quotient (IQ) testing. It is a score derived from several different standardized tests designed to assess relative intelligence. These standardized tests include Stanford-Binet Intelligence Scale and the Wechsler Adult Intelligence Scale (WAIS) and are updated time to time. The former is used to access young children and includes both verbal and non-verbal subsets. It tests five factors namely knowledge, quantitative reasoning, visual-spatial processing, working memory, and fluid reasoning, on the other hand, is used for adults and older adolescents.

Its current version (WAIS-IV) consists of four core indexes further divided into five cores and ten supplemental subsets as shown in Fig. (**16.1**). In either case, the score of less than 70 established the diagnosis of ID [5, 7].

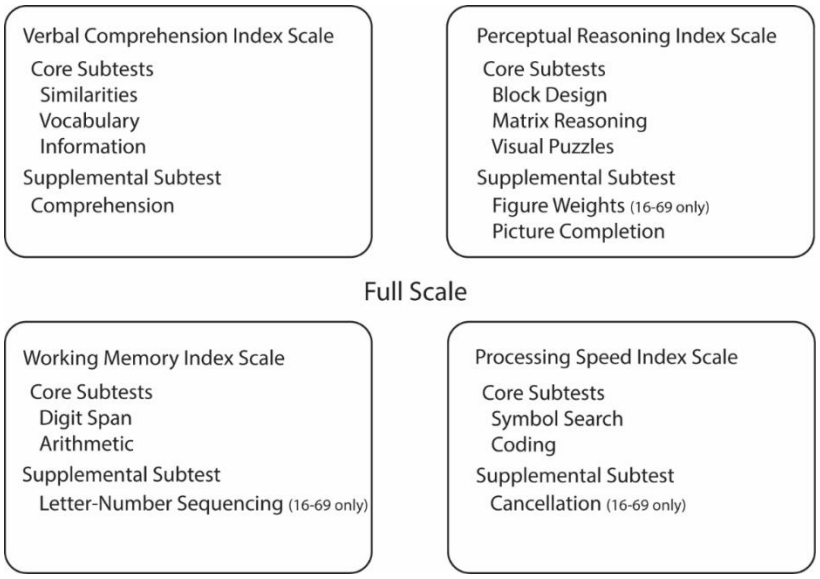


Fig. (16.1). Wechsler Adult Intelligence Scale (WAIS)- IV scale showing core and supplemental subsets of components used for assessment of Intelligence Quotient (Adopted and recreated from Wechsler 2008 [6].

Based on this IQ level, ID is classified into four subtypes. The types are mild (IQ: 55–70), moderate (IQ: 40–55), severe (IQ: 25–40) and profound (IQ< 25). Mild ID is most prevalent, accounting for almost 85% of cases, followed by the moderate ID having 10% of the patients. Severe and profound cases are the least prevalent with only 5% of the patients [8]. Based on inheritance pattern, the monogenic form of ID is divided into autosomal dominant and recessive, x-linked dominant and recessive. On the other hand, ID is grouped into non-syndromic ID and syndromic ID. In case of non-syndromic ID, ID is the sole clinical feature among the patients. While in syndromic ID, affected individuals manifest one or more clinical comorbidities along with ID. The associated comorbidities may be of dysmorphic, neurological, or systemic nature. ID may also co-occur with autism spectrum disorders (ASD), microcephaly [9 - 11].

3. AUTISM SPECTRUM DISORDERS (ASD)

ASD is a neurodevelopmental disease and is characterized by a significant impairment in social and communicative skills and abilities. It is one of the many disorders that co-occur with ID. The prevalence of comorbidity of ID with ASD ranges from 10% to 40% [12, 13]. The clinical phenotype of ASD is extremely

CHAPTER 17

Primary Microcephaly and Schizophrenia: Genetics, Diagnostics and Current Therapeutics

Iram Anjum[1,*], Aysha Saeed[1], Komal Aslam[1,2], Bibi Nazia Murtaza[3] and **Shahid Mahmood Baig[4]**

[1] *Department of Biotechnology, Kinnaird College for Women, Lahore, Pakistan*

[2] *Department of Biotechnology, Lahore College for Women University, Lahore, Pakistan*

[3] *Department of Zoology, Abbottabad University of Science and Technology, Abbottabad, Pakistan*

[4] *Pakistan Science Foundation, Islamabad, Pakistan*

Abstract: Intellectual disabilities (ID) are among the most common genetic disabilities worldwide. Over the last two decades, ID has especially drawn special scientific interest being the key to understanding normal brain development, growth, and functioning. Here, we discuss two intellectual disabilities to better understand the emerging trends in disease diagnosis as well as the therapies available for their management. Primary microcephaly (MCPH) is a monogenic genetic disorder with twenty-eight loci (MCPH1-MCPH28) mapped so far with all the causative genes being elucidated as well. The role of these genes in disease prognosis along with their association with various MCPH-linked phenotypes plays an important role in the molecular diagnosis of the disease. As there is no cure/treatment yet available to enlarge a congenitally small brain, management modalities in use include physical, speech and occupational therapies as well as psychological and genetic counselling to not only reduce the incidence of the disorder but also to help families cope better. The second intellectual disability being discussed here is schizophrenia which is a multifactorial disorder owing to its complex and extremely heterogeneous etiology. Although various environmental factors play an important role, the genetic factors have been identified to play the most pivotal role in disease presentation as to date, 19 loci (SCZD1-SCZD19) have been linked to schizophrenia. However, underlying genes for only six of these loci have been mapped along with 10 other genes that are either linked to schizophrenia or show susceptibility to it. Diagnosis of schizophrenia needs careful consideration and various tests and tools currently employed for complete diagnosis have been discussed here. The management options for schizophrenia include pharmacological, non-pharmacological and intracranial therapies. These disorders shed light on the important role omics technologies have played not only in better understanding of the disease prognosis but also assisting in disease diagnosis and treatment modalities too.

[*] **Corresponding author Iram Anjum**: Department of Biotechnology, Kinnaird College for Women, Lahore, Pakistan; E-mail: iram.anjum@kinnaird.edu.pk

Syeda Marriam Bakhtiar and Erum Dilshad (Eds.)

Keywords: Cortical development, Dopamine pathways, Genetic heterogeneity, Genotype-phenotype associations, Occipitofrontal head circumference (OFC), Primary microcephaly (MCPH), Schizophrenia.

1. INTRODUCTION

Intellectual disabilities have also attracted a lot of scientific inquisitiveness since ancient times, as Hippocrates in the fifth century BC, proposed a physiological insight and identified intellectual disability as a neurodevelopmental defect and explained it as a defect caused by irregularities in the four humors of the cerebrum.

People with intellectual disabilities (ID) have since long been a target of public isolation and humiliation. Not only the early influential Greek and Roman scholars used to belittle such people, but till very recently, individuals with such disabilities have been tagged as mentally retarded. However, the term mental retardation has now been replaced by intellectual disability, owing to the offensive nature of the term as well as the negativity surrounding it.

However, the degree of intellectual disability is variable ranging from mild to severe and people with intellectual disabilities can learn new skills, although slower than their normal counterparts, depending on the severity of the condition. The underlying mechanisms causing intellectual diseases are complex and though many factors may contribute, their etiology can be mainly divided into two categories, *genetic abnormalities* (*e.g.* gene mutations, copy number variations and chromosomal aberrations) and *environmental* factors, which include maternal exposure to toxins, delivery and postnatal complications as well as traumas after birth [1].

To better understand the intellectual diseases, their etiology, diagnoses as well as the current therapeutics, here we discuss two intellectual diseases, primary microcephaly having a predominantly genetic etiology and schizophrenia being a multifactorial complex disorder.

2. PRIMARY MICROCEPHALY

Primary microcephaly or *microcephaly vera* or *microcephaly primary hereditary* (MCPH; OMIM 251200) is a congenital neurodevelopmental condition, characterized by a smaller head (at least -2 standard deviations at birth below the mean accounting for the age, sex and ethnicity) accompanied with non-progressive mild to severe intellectual disability. The smaller head characteristic

of this disorder explains the Greek origins of the term, *i.e.*, *μικρό* and *κεφάλι* meaning small and head respectively. It is essentially a disorder of the neocortex involving not only abnormal neuronal migration but also a reduction in the number of neurons in the developing neocortex [2, 3].

2.1. Clinical Attributes

Although the head circumference (HC) of MCPH individuals is at least -2 standard deviations (SD) at birth, it has been observed that most MCPH affected have HC > -4 SD, six months postnatally below the individuals of the same age and sex, meanwhile their femur length remains within 2 SD of the normal range. The HC measurement, which is also termed occipitofrontal circumference (OFC), is considered a direct indication of brain growth, and a key step in the assessment of childhood development and growth. If the head circumference is less than 4 SD, severe mental retardation is observed. Apart from a smaller head size, the symptoms of microcephaly may include seizures, developmental delays like a delayed speech, sitting, walking, and standing as well as cognitive impairments and intellectual delay. Other symptoms include high-pitched cry, hearing defects and vision problems. The severity of these symptoms varies and could be life-long and sometimes even be life threatening. However, it has also been observed that some affected persons may even show no symptoms at all except a smaller than average head circumference [2 - 4].

2.2. Types of Microcephaly

Primary microcephaly is one of the two types of microcephaly categorized based on the time and factors contributing to the smaller head [5]. Primary microcephaly is present at birth which can either be an inherited condition or the result of a malfunction during pregnancy. Generally, genetic abnormalities interfering with the expansion of the cerebral cortex during the development of the fetus cause primary microcephaly. Reasons for PM other than genetic may include any infection which may be passed from the mother to the fetus during pregnancy such as toxoplasmosis, German measles (rubella), cytomegalovirus, chickenpox (varicella), and zika virus. Exposure of a baby in the womb to alcohol, drugs or any toxic chemical can also lead to brain abnormalities.

Secondary microcephaly as opposed to its counterpart develops postnatally from an insult to an otherwise well-formed central nervous system. It is characterized by a remarkably low number of neuronal dendrites as well as synaptic connections. Environmental factors which most likely contribute are disruptive brain injuries, teratogen exposure, ischemic stroke, *etc.* affecting the postnatal development of the brain and thus secondary microcephaly.

SUBJECT INDEX

A

D

E

F

G

H

I

K

L

M

N

O

P

R

S

T

U

V

W

X

Y

Z