Transcriptional Approach in the Identification of Drug Targets in *Candida* spp.

9

Mahnoor Patel, M. Amin-ul Mannan, and Banhishikha Datta

Abstract

High-throughput sequencing technologies have become essential in studies on genomics, epigenomics, and transcriptomics. While sequencing information has traditionally been elucidated using a low-throughput technique called Sanger sequencing, high-throughput sequencing (HTS) technologies are capable of sequencing multiple DNA molecules in parallel, enabling hundreds of millions of DNA molecules to be sequenced at a time. This advantage allows HTS to be used to create large data sets, generating more comprehensive insights into the cellular genomic and transcriptomic signatures of various diseases and developmental stages of disease-causing pathogens. The transcriptomics techniques like microarray and RNA sequencing (RNA-seq) can be used to compare differential expression of the genes and the underlying mechanism and regulatory pathways over diseased and normal states. In this chapter, we have elucidated the transcriptomics approach for the identification of the lead compounds for the diseases caused by *Candida* species. *Candida* spp. are commensal organisms and regarded as opportunist pathogens. It causes serious systemic infection with a mortality rate of ~50% in immunocompromised patients. The new clinical isolates are showing resistance to the existing drugs, and hence, new candidate molecules are required. The chapter enumerates the various technologies which can be deployed to identify the candidate drug molecules.

M. Patel · M. Amin-ul Mannan (✉) · B. Datta
Department of Molecular Biology & Genetic Engineering, School of Bioengineering & Biosciences, Lovely Professional University, Phagwara, Punjab, India
e-mail: mohammad.20597@lpu.co.in

S. Hameed, Z. Fatima (eds.), *Integrated Omics Approaches to Infectious Diseases*,
https://doi.org/10.1007/978-981-16-0691-5_9

Keywords

Big data analysis · Next-generation sequencing · Integrated genomics · *Candida* spp. · Antifungals

9.1 Introduction

High-throughput omics approaches like genomics, transcriptomics, proteomics, and metabolomics contribute a great deal in understanding the biological process including the identification of candidate molecules for therapy. It can generate a large number of data in a single day [1]. However, all of these omics technologies face challenges like cleaning of data, identification of biomolecules, reduction of data dimensionality, biological contextualization, statistical validation, handling and storage of data, sharing, and archiving. Large-scale omics data set access is important for biological processes improvement and in systems biology. Since the procedural costs to experiment with high-throughput sequencing are far more economical as compared to costs a decade ago, it generates enormous data sets. On one hand, it is challenging, but at the same, it also creates an exhilarating opportunity for the biologists, biostatisticians, and computational biologists to analyze those data [2]. Omics approaches based on the global analysis of biological samples with an aid of high-throughput analysis and bioinformatics provide novel insights into the biological processes [1].

This chapter is an attempt to comprehend one of the omics approaches, i.e., transcriptomics, for the identification of the candidate molecules for the therapy of fungal disease caused by *Candida* spp. Candidiasis is an opportunistic fungal infection caused by *Candida* spp. Recent trends suggest that the number of cases and deaths related to candidiasis is alarming and escalating. Antifungal and multidrug resistance is one of the major challenges in the management of candidiasis. Among all *Candida* spp., *Candida auris* has emerged as a multidrug-resistant strain [3, 53]. The Centers for Disease Control and Prevention (CDC) has recognized it as a global threat. The transmission of this fungus resembles methicillin-resistant *Staphylococcus aureus* [4]. The chapter revolves around describing the omics approach for the identification of lead compounds and thereof its therapeutics.

9.2 Omics Approach

After the introduction of omics technologies in the post-genomic era, biological studies are characterized wisely and rapidly developed [5]. These kinds of technologies include genomics, proteomics, metabolomics, transcriptomics, lipidomics, and phenomics [52]. The omics approach is generally based upon the global analyses of biological samples with the help of high-throughput technology including bioinformatics which provides novel insights of the biological samples and their phenomena [6]. HTS technologies, like whole-exome sequencing, can be

used to identify novel variants and other mutations that may underlie many genetic disorders. The current high-density arrays with multiplexed features permit a sample size of ~20,000 cells with automated features and permit high-sample handlings. Numerous research and clinical applications like pharmaceuticals, diagnostics, therapeutics, disease prevention and pharmacogenomics, evolutionary genetics, and developmental biology including comparative genomics use the latest approach of genomics, proteomics, transcriptomics, and metabolomics [7]. Some of the tools used in drug discovery are mentioned in Table 9.1. A major fundamental difference between transcriptomics with other omics techniques is the activity measurement of a single class of molecules. The traditional methods require a different assay to measure the gene function, mutational analysis, metabolite and enzyme activity, and ligand-receptor interaction. Transcriptomics bridge the gap between genomics and proteomics and can aid in new drug discovery of economical, affordable, and better-quality drugs.

9.2.1 Genomics

Genomics is nucleic acid-based technology that relies upon several steps, namely, sample collection, high-quality extraction of nucleic acid, preparation of library, clonal amplification, and sequencing. Every approach is based on the downstream application of the sample. After sequencing, the process workflow includes cleaning of data, filtering, assembling, alignment, variant calling, annotation, and functional prediction [1, 2].

In every area of biological investigation, genomic technology is widely used. It includes genomics research which consists of functional as well as structural genomics. Three-dimensional structures of proteins that are encoded by a genome are also included in the structural genomics study [8]. It allows high-throughput methods for the structural analysis with the help of experimental and modeling approach combination. Today, the major branch of genomics is involved in sequencing the genome of the various organisms [9]. Describing genes and functions of proteins and their interaction with other proteins falls under preview of functional genomics. Bioinformatics and microarrays are significant tools for genomics. It includes metagenomics, epigenomics, and pharmacogenomics [1, 2, 10]. The genomics data of *Candida* spp. can be obtained from the Candida Genome Database which is maintained by the US National Institutes of Health [11]. The database has the represented genomes of *C. albicans*, *C. auris*, *C. dubliniensis*, *C. parapsilosis*, and *C. tropicalis*. Genomic sequences, including gene expression and protein information, can be retrieved from the database.

Table 9.1 Representative 'omics' approach used in drug discovery

Omics approach	Uses	Resource	Resource link
Genomics	• Mechanism of pathogenesis • Identification of virulent genes • Discovery of candidate molecules • Efficacy and toxicity of drugs	• GWAS Central	https://www.gwascentral.org/
		• PharmGKB	https://www.pharmgkb.org/
		• dbGaP	https://dbgap.ncbi.nlm.nih.gov/
			http://www.candidagenome.org/
Proteomics	• Drug target efficacy • Protein toxicology • Protein-protein network interaction • Mass spectral database	• ProteomicsDB	https://www.proteomicsdb.org/
		• The Human Protein Atlas	https://www.proteinatlas.org/
		• DITOP	http://bioinf.xmu.edu.cn/index.jsp
			https://massbank.eu/MassBank/
Transcriptomics	• High-throughput functional genomic data • Gene expression data linked to phenotype data • Minimum Information About a Microarray Experiment • Minimum Information About a High-throughput Nucleotide SEQuencing Experiment • Understanding of cell pathways	• GEO	https://www.ncbi.nlm.nih.gov/geo/
		• Open TG-GATEs	https://toxico.nibiohn.go.jp/english/index.html
		• MIAME	https://www.ncbi.nlm.nih.gov/geo/info/MIAME.html
		• MINSEQE	http://fged.org/projects/minseqe/
		• The LINCS Consortium	https://lincsproject.org/
Metabolomics	• Small-molecule metabolites found in the human body • GC/MS profiling studies of metabolites • Drug target efficacy and safety evaluation • Metabolic toxicity	• Human Metabolome	https://hmdb.ca/
		• Golm Metabolome	http://gmd.mpimp-golm.mpg.de/
		• MetabolomeExpress	https://www.metabolome-express.org/

9.2.2 Proteomics

The proteome is the whole set of proteins of any organism which is translated by every organism or any biological system. Proteins differ in the presence of genetic and environmental changes. Proteomics comprises a particular type of cells, body

fluids, and tissues including their functions and structures. Proteomics is very helpful in understanding the research at the translation level after the genomics; it also helps in understanding the post-translation modifications. It is an essential tool for understanding the genome at its expression level [1, 2].

For the quantification of proteins from the multiple samples, proteomics approach is generally used. It uses both shotgun and targeted approaches. Recent advancement in mass spectrometer (MS) radically increases the sensitivity and decreases the sample amount which is required in the high-throughput analysis. It also allows for minimal differences in protein profusions and posttranslational modification identification [12]. The major steps in proteomics include proper sample collection, extraction of protein and peptides, enzymatic digestions, and fractionation/separation by using liquid chromatography followed by MS, identification, and quantification of proteins and peptides. Proteomics moved from the traditional 2D-PAGE-based spot extraction of proteins followed by the LC-MS or matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) system [1, 2, 10].

9.2.3 Transcriptomics

All the set of RNA molecules are coming under the transcriptomics studies are known as transcriptome. It includes mRNA, tRNA, and rRNA molecules with the other noncoding RNA molecules present in the cells. Unlike the whole genome, it varies under the influence of external conditions of the environment. It examines changes that occur in the entire transcriptome under different biological surroundings [13]. The various transcriptomics technologies are shown in Fig. 9.1.

RNAs are the sequence which is generated from the DNA sequences, and that is why they are the mirrors of DNA sequences. In the transcription process, RNA synthesis is the initial step of expression of the gene. Although same genome exists in every cell of an organisms, every cell expresses different genes at different transcriptional control to generate the diverse repertoire of proteins [14]. Transcriptomics data helps researchers in the understanding of gene function and its comparison of different types of healthy cells transcriptome to the transcriptome of the diseased cell. These type of data help researchers in understanding the genes' misleading functions and its interpretation [1, 2].

9.2.4 Metabolomics

Complex biochemical cascades of end products are generally known as metabolites. It can link with the genome, transcriptome, and proteome to a phenotype which can provide an important detail for the metabolic variation and complement in the genetic basis of discovery [15]. Metabolomics is generally used for the determination of the absolute and relative amount of sugars, amino acids, lipids, nucleotides,

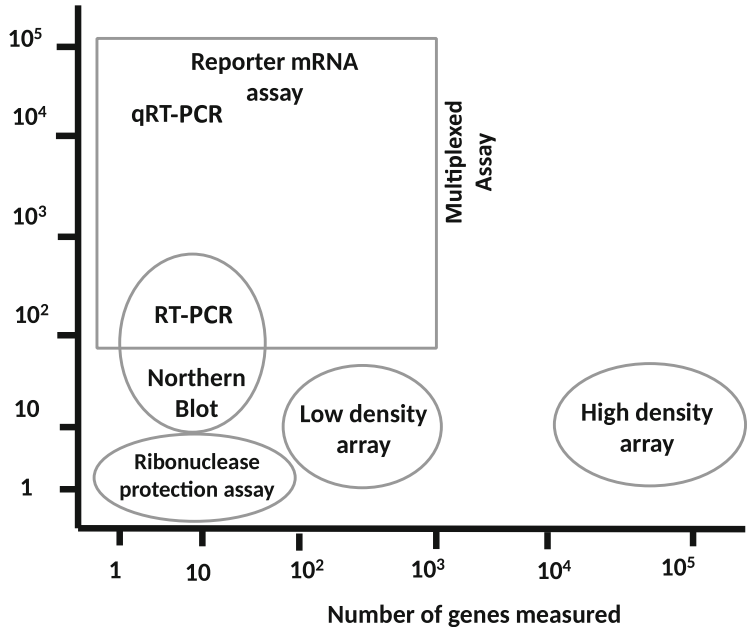

Fig. 9.1 Gene expression "transcriptomics" technologies (*the diagram is not drawn to the scale*)

steroids, and drug constituents [54]. Depending on instrumentation and applications in research, metabolomics can capture information of small molecules in liquid, solid, and capillary electrophoresis. Important steps in metabolomics research include experimental design, appropriate sample collection strategies, quenching of metabolism, extraction of optimized metabolites and its reconstitution from the sample, and data analysis of MS or NMR which includes data alignment, filtration, imputation, statistical analysis, annotation, and network or pathway analysis. These steps are extremely variable and depend on the sample analysis and platform used for the process [10, 16].

9.2.5 Pharmacogenomics

Every person possesses unique variants of the genome, which leads to an individual's diverse reactions to drugs. Pharmacogenomics gives an idea about how a person's genes affect drug responses so that safe and effective medications can be developed and the determination of its doses can be done. It helps in the discovery process of genes responsible for particular diseases including its investigation on the effects of genetic factors for medication for predicting a person's response [17]. Confrontational drug reactions are described as an important cause for hospitalizations including deaths in few countries. Pharmacogenomics empowers researchers in understanding how inherited gene variances affect responses toward

medications. Pharmacogenomics gives opportunities to researchers for understanding the mechanism of differences of inherited disease-related genes affecting the patient's body response to the medications, in which data is important for a prediction about the drug's effectiveness and it's response in the patient [1, 2].

9.2.6 Epigenomics

DNA sequence changes leads to the change in the gene expression, which are often heritable. Epigenetics studies the inheritance pattern in the gene expression without any changes in the DNA sequence per se. Epigenomics deals with the analysis of epigenetics at the global level and in the entire genome and genetic information in terms of DNA sequences as it is also able to affect the functions of particular genes [18]. There are five diverse mechanisms of epigenetic regulation: (1) methylation of DNA, (2) posttranslational modification of histone, (3) variants of histones, (4) RNA interference, and (5) nuclear organization. Genome function can be changed under the influence of exogenous factors which usually occurs in CpG islands, which is a GC-rich region of DNA based on methylation which is the most common genomic parameter (e.g., regions of the promoter, regulatory domains of genes, and intergenic regions of a genome) [1, 2].

9.2.7 Immunomics

Immunomics is the study of the regulation and response processes of the immune system against the pathogen. Immunomics deals with every molecule of the immune system including the targets of immune cells and their functions. There are many techniques related to genomics, bioinformatics, and proteomics which include immunomics [19]. After the advancement in genomics and proteomics, the immunomics approach uses bioinformatics, structural biology, high-throughput screening, and biochip for studying immune cells and their responses. Immunomics is generally used for discovering new susceptibility of genes and their correlation with the immune cells [20]. Every person's immune system possesses a great level of diversity as compared to the person's other body systems. For research on a highly complex system, traditionally developed methods are mostly limited. Immunomics may prove as an advanced newly developed approach. Generally, it is used for vaccine development including target identification and diagnosis of disease [1, 2, 51]. Immunological research became more effective with the help of the immunoinformatics approach which is also known as computational immunology. It applies in silico modeling including analyzing the problems and data of the immune system. This new branch of bioinformatics having several software and resources focused on immunology, helps in understanding complete immune system properties [21].

9.2.8 Cytomics

Cytomics comprises structural and functional research of the cellular systems. This kind of study involves databases at the genomic level. Cytomics studies also involve the use of many technologies at the genomic and proteomic levels [22]. Noninvasive, sensitive, fluorescence-based technologies are mostly involved in the studies of cytomics for conducting single-cell integrated analysis [23, 24]. Cell imaging and quantitative data of fluorescent technique which is performed on a single cell is helpful in a comprehensive analysis of cellular processes. Cytomics comprises current technologies including flow cytometry, confocal laser scanning microscopy, high-content screening, laser capture microdissection, bio-imaging, and laser scanning cytometry [25]. It provides approaches and strategies for the pharmaceutical research like a validation of target, development of drug, toxicological and pharmaceutical evaluation and validation, and efficiency at the clinical level for prognostic and personalized medicine [26].

9.3 Omics Application in Pharmaceutical Research

There are a few limitations in the application of the omics approach till now. Such omics data may create false-negative or false-positive results because of the large number of massive complicated data. Because of this limited accuracy and sensitivity of the methods, many times, few important functional biological molecules which are present in trace amount cannot be detected [27, 28]. Furthermore, the assessment of omics data lacks proper specificity. Nowadays, research in the pharmaceutical industry generally relies on the omics approach which includes genomics, proteomics, transcriptomics, and metabolomics. It also uses multiple combinations of omics technologies [29].

Every phase of pharmaceutical research which includes drug development, evaluation of efficacy, validation of target and discovery, safety assessment, and development of personalized medicine uses many kinds of omics approach. It is the most powerful and efficient tool in pharmaceutical discovery. It is becoming the most vital part of the network and systems biology which makes it possible for understanding in-depth concepts easily including the pathological processes and simulation performance of pathogen's interactions with the host immune system for all diseases [30]. The omics approach reveals all the possible key pathways and their mechanism for enhancing pharmaceutical research and drug development. Additionally, studying omics also highlights all the probable targets for new drug development, which allows safety assessment efficiently and personalized medicine development [31, 32].

9.3.1 Target Discovery

Target discovery is very essential for developing new drugs. In the past, new drug development process for any disease was dependent on the 500 early known drug

targets. After completion of the Human Genome Project back in 2003, the genomic studies indicate that there roughly 22,000 protein-encoding genes [33]. Till now, around 10% of genes have been explored for drug target identification, and still, there is much needed to be done. Generally, developing a drug based on a single chemical with a single target group is not efficient. In recent years, omics technologies with systems biology applied widely also provide an idea about the identification of target and novel drug development. In present times, there are many new omics technologies applied for designing new drugs including the discovery of targets, microbial genomics and proteomics, nuclear magnetic resonance, RNAi, gene transfection, and gene knockout modeling. These omics approach produces a vast number of data and many databases which have been constructed, like Online Mendelian Inheritance in Man (OMIM), Therapeutic Target Database (TTD), Cancer Gene Census, and Gene Expression Omnibus (GEO) [34, 35, 55] (Fig. 9.2).

9.3.2 Toxicity and Toxicogenomics

The toxicology of drugs plays an essential role in drug development and pharmaceutical research. Toxicity is the greatest reason for the terminating process of drug development. Toxicology of drugs can guide the clinical medication for reducing adverse reactions of drugs. From the last 20 years, many more omics technologies are applied for toxicology analysis in drug development and also promoted its discovery in the different fields of research in toxicology [36, 37].

The genomics application in the toxicology field is known as toxicogenomics. For clarifying the relationship between the changes in gene expression and toxicity, a toxicogenomics study is applied. It also helps in identifying probable genetic toxicants, and after that, their mechanism of action is understood. For understanding toxicogenomics, the microarray is generally used. It is reflected that practically, all toxic reactions depend on changes that occur in profiles of gene expression [56]. As compared with the traditional drug toxicity research, the new field of toxicogenomics delivers a more comprehensive and sensitive platform for the safety assessment of drugs. Measuring the expression of a gene at a larger scale, the most sensitive and relevant changes in genetics can be found which can be used for risk management as biomarkers. For example, gene expression which is involved in the repair of DNA damage may be a genotoxicity sign.

Transient changes at the earliest expression of genes are related to the stress response of the body, whereas long-term changes in the profile of genes are related to chronic toxicity [38]. It may become an adaptive response in the body. For the determination of chronic toxicity, carcinogenicity, and the drug's secondary toxic effects, this technology is very much essential. Furthermore, in the early stages of the new drug development process, precisely expressed genes or proteins specific to toxicants also developed like a biomarker for understanding and predicting drugs' potential toxicity. This generally helps in a lead compound generation for producing and evaluating toxicity with high efficiency and sensitivity. This mode of toxicity evaluation mode delivers more valuable and relevant information about the toxicity

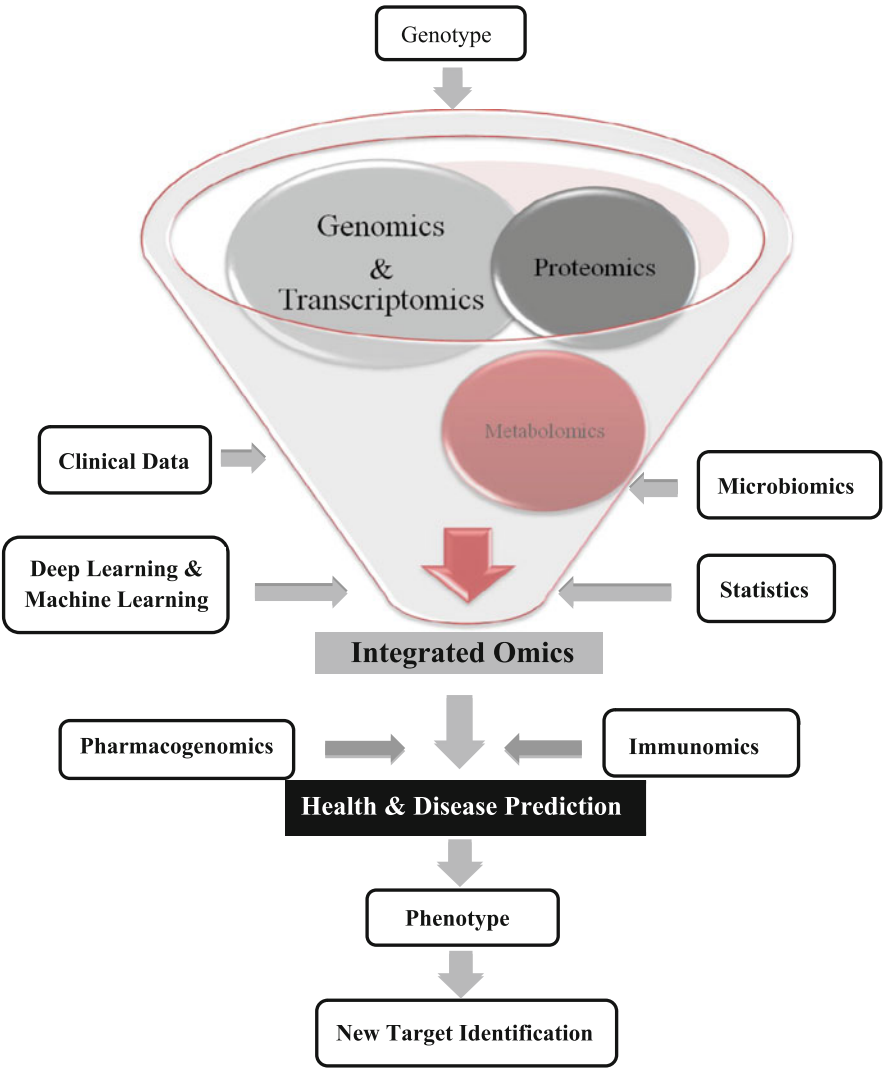

Fig. 9.2 The omics approach in the identification of new drug target

mechanism in a short time relatively. As compared to the toxicity study in a traditional manner, this newly developed omics technology which is known as toxicogenomics brings a revolution in drug toxicology studies [39].

9.3.3 Toxicoproteomics

As toxicogenomics is a larger field of study, toxicoproteomics is just a part of this vast field. Toxicoproteomics helps in identifying critical proteins and their pathways in biological systems that affected and respond to environmental exposure and adverse chemical reactions with the help of global expression technologies of proteins [40]. Traditional toxicology, expression analysis of differential proteins, and pathology are three major integrated areas of toxicoproteomics. Recently, this technology can reveal the expression of toxicant-reduced proteins; also, it can help study posttranslational modification with protein-protein interaction [41]. By doing a comparison of specific cells, organs, and tissues' protein expression profiles with those profiles which are generated by toxicants, toxicoproteomics can highlight in a very short period the specificity of the toxic protein expression which can execute functional molecule deficiency caused by toxicants. Consequently, with the help of the antibody analysis method, new markers of toxic proteins can also be discovered. These kinds of toxic markers can be applied for studying the mechanism of the human body at a safer dose [42].

9.4 Emerging Disease Causative Agents of *Candida* spp.

9.4.1 Prevalence of *Candida* Species

Candida spp. cause superficial skin infection to mucosal and deep tissue infections. It contributes to high mortality and morbidity [3]. The emergence of multidrug-resistant clinical isolates and a limited number of antifungal agents adds to the seriousness of the problem. One of the major problems with candidiasis or candidemia is biofilm formation [43]. Severely ill or immunocompromised patients are generally more prone to developing both superficial and life-threatening infections [44]. It is also a very common infection in AIDS patients which leads to malnutrition and causes interference in the absorption of the medications which was proved in epidemiological studies [43].

9.4.2 Prevalence of Non-albicans *Candida* Species

Generally, *Candida albicans* is the predominant species involved in invasive fungal infections. However, recent literature suggested there are increasing cases of emergence of non-albicans *Candida*. These epidemiological changes are associated with severe immunosuppression or critical illness and broad-spectrum antibiotic exposure with increasing age [44]. An investigation exhibited that more than half of the cases of infections caused by *Candida* species were reportedly by *Candida albicans*, and the other incidence rate for the non-albicans infection rates was reportedly 14% by *Candida parapsilosis* and *Candida glabrata* including 7% by the *Candida tropicalis* and 2% by *Candida krusei* [43, 45].

9.4.3 Targets of Antifungal Candidates

When compared to antibacterial research, slight advancements have occurred in the development of a new antifungal agent. As fungi are eukaryotic organism having a close evolutionary relationship with the human host, this creates complications in the search for the antifungal agents. The help of new approaches about antifungal therapy including target identification and rational drug design technologies provides imperative acceleration in the development process of an antifungal agent by reducing the time for the cure and improving the quality of patient's life. Nowadays, commercially available antifungal agents have targets that are restricted to the plasma membrane and cell wall. Some examples of antifungal candidate's new as well as old targets are mentioned in Table 9.2 [43, 57].

The evolution of drug resistance is a rapid phenomenon in *Candida* spp. Some of the online database tools that can be used for drug target identification are mentioned in Table 9.3. For developing a new drug, the first step is to identify the drug target and its validation. It is also very much essential for the elucidation of disease pathology mechanism identification and the effects of drugs [46]. Using the in silico approach, our lab also contributed to the identification of target molecules in multidrug-resistant *C. auris* (unpublished data). Even though these tools help in the identification of the drug molecule, however, it needs to be validated in wet lab experiments.

9.4.4 NGS and Fungal Diagnosis

The next generation sequencing offer valuable tools in understanding the molecular mechanisms of antifungals compounds. It can also be depolyed for the detection of new mycobiota and species specific identification. In mycological diagnosis and research, sequencing technology including an enhanced capability provides a

Table 9.2 Antifungal candidates with targets

Antifungal candidates	Target
Echinocandins	Inhibition of β-glucan synthesis
Nikkomycin and polyoxins	Inhibition of chitin synthase
Azoles	Inhibition of 14α-demethylase
Terbinafine and naftifine	Inhibition of squalene epoxidase
Polyenes, naphthoquinones, eugenol analogues, isoquercitrin	Bind to ergosterol
5-Flucytosine	Inhibition of DNA synthesis
Amphotericin B, miconazole, ciclopirox	Production of reactive oxygen species (ROS), leading to cell death
Sordarins	Inhibition of protein synthesis
Griseofulvin	Microtubule assembly
Triphenylethylenes	Inhibition of calcineurin signaling

Table 9.3 Software/tool/databases used for new drug target identification

Software/tool/database	Purpose
UniProt	The whole proteome of *Candida* can be retrieved
BLASTP	Using this tool, the retrieved proteins can be compared with the Human Protein database, and their foreignness can be determined
PVS and EMBOSS	Using these tools, antigenicity of the predicted protein can be identified. The score can be analyze using EMBOSS, and an antigenic propensity graph can be generated using PVS
ArgusLab	Used for the generation of antigenic peptide model from antigenic protein including energy and geometry calculation for each peptide
Swiss-MODEL	Used for the 3D model generation of antigenic protein
ProtParam	Used for physicochemical properties analysis of antigenic protein
SOPMA	Secondary structure prediction is done by using SOPMA
PROCHECK	Used for stereochemical quality analysis of immunogenic protein
ProSA-web	Used for calculation of Z-score
Kolaskar and Tongaonkar Method	Used for categorizing linear and conformation B-cell epitopes
Emini Surface Accessibility	Used for analyzing surface accessibility of the protein
Karplus and Schulz Flexibility	Used for analyzing the flexibility of the protein
Parker Hydrophilicity	Used for analyzing hydrophilicity of the protein
ElliPro	Used for analyzing conformational B-cell epitope
NetMHCIIpan 4.0	Used for analyzing helper T cell from immunogenic protein
NetCTL 1.2	Used for analyzing cytotoxic T cell from immunogenic protein
IEDB Immunogenicity	Used for analyzing strong and weak immunogenicity of the predicted CTL
PEP-FOLD3	Used for the 3D model generation of predicted CTL molecule
PatchDock	Used for performing molecular docking studies between HLA-A*0201 molecules and predicted CTL molecule
FireDock	Used for refining molecular docking results
RasMol	Used for visualization of molecular docking results

powerful tool. Next-generation sequencing (NGS) functionality is applied in the public health microbiology laboratories for the studies of metagenomics and outbreak monitoring. Speed and sensitivity of diagnosis of infectious fungal diseases including determination of the mycobiome can be increased with the help of advancement in molecular tools and techniques. NGS can enhance the creation of data at the molecular level because the mycobiota which is dependent on culture for identification method is the major limitation of many fungal species that cannot culture in vitro, and at the same time, NGS technology is very much valuable in the diagnosis. As it is true that the capability of the sequencer is limited, it means an application of the whole-genome sequencing for complex microbiome determining and diagnosis in the medical sample is far beyond the possibility which exists in today's sequencers. Nowadays, the increasing capacity of the sequencing platform with the continuous decrease in sequencing cost makes it a striking tool in mycology

Table 9.4 Comparative identification methods for the *Candida* spp.

Sr. no.	Method	Advantages	Disadvantages
1	Microbial	Cost-effective	Time-consuming
2	Molecular	Highly accurate	High cost
3	MALDI-TOF	Rapid and accurate identification	Cost is high
4	VITEK	Automatic identification	Cost is high
5	NGS	Accuracy is high	Cost is high, can't afford by every lab

for microbiome analysis [47], for example, *Candida glabrata*, *Candida auris*, *Candida tropicalis*, *Candida albicans*, and *Candida parapsilosis* (Table 9.4).

9.5 Current Challenges with Future Directions

Due to rapid advancement in high-throughput technologies and computational data analysis, the omics technology is getting wider publicity and acceptance. The transcriptional analysis is rapidly implemented in drug discovery owing to its sensitivity, large-scale quantitative data, reproducibility, and robust assay method. In a single day with the automated microplate reader or RNA sequencing, trillions of data sets can be captured. However, handling large data requires a specialist or data scientist with a background in biology, and often, it is difficult to obtain [48]. After 20 years of exploration of omics technologies, it became routine work to generate and deal with omics data which is not so much tedious about the generation of data sets with the help of high-throughput data by an analytical approach.

In the nearest future, for describing biology processes, multiple omics technology combinations are the general approach for assessment. It produces a complex large number of data at various levels which include DNA, RNA, and proteins [49]. By experimenting or with the help of exploration of Internet databases, omics data can be acquired. But it is more difficult for processing because of many reasons, like data-type diversity, redundancy of database, and uniformity lacking in description of standard data. How to deal with this kind of a large amount of data especially taken from different sources of multi-omics approach is the most difficult challenge in omics research. Network biology may become a probable solution to this problem for its efficient solution, as it describes biochemical systems like a network system for multi-omics data [50].

Acknowledgment Dr. Mannan would like to acknowledge lab funding from DST-SERB for a core research grant (EMR/2017/002299) vide sanction order SERB/F/10997/2018-2019.

Funding The work was not directly funded from any research grant.

References

1. Mishra BB, Langefeld C, Olivier M, Cox LA (2020) Integrated Omics: tools, advances and future approaches. J Mol Endocrinol 62:21–45
2. Kai YS, Hui LR, Zi JH, Ru LX, Ji Y, Lei S, Dong ZW (2015) Omics in pharmaceutical research: overview, applications, challenges, and future perspectives. Chin J Nat Med 13:3–21
3. Park JY, Bradley N, Brooks S, Burney S, Wassner C (2019) Management of patients with *Candida auris* fungemia at the community hospital, Brooklyn, New York, USA, 2016-2018. Emerg Infect Dis 25:3
4. Torres SR, Kim HC, Leach L, Chaturvedi S, Bennett CJ, Hill DJ, Jesus MD (2019) Assessment of environmental and occupational exposure while working with multidrug-resistant (MDR) fungus *Candida auris* in an animal facility. J Occup Environ Hyg 14:7
5. Hume HKC, Vidigal J, Carrondo MJT, Middelberg APJ, Roldao A, Lua LHL (2019) Synthetic biology for bioengineering virus-like particle vaccines. Biotechnol Bioeng 116:919–935
6. Paananen J, Fortino V (2019) An omics perspective on drug target discovery platforms. Brief Bioinform:1–17
7. Doolan DL, Apte SH, Proietti C (2014) Genome-based vaccine design: the promise for malaria and other infectious diseases. Int J Parasitol 44:901–913
8. Bredel M, Jacoby E (2004) Chemogenomics: an emerging strategy for rapid target and drug discovery. Nat Rev Genet 5:262–275
9. Mishra NK, Shukla M (2014) Application of computational proteomics and lipidomics in drug discovery. J Theor Comput Sci 1:105
10. Wolfender JL, Litaudon M, Touboul D, Queiroz EF (2019) Innovative omics-based approaches for prioritization and targeted isolation of natural products – new strategies for drug discovery. Nat Prod Rep 36:855–868
11. Skrzypek MS, Binkley J, Binkley G, Miyasato SR, Simison M, Sherlock G (2017) The Candida Genome Database (CGD): incorporation of Assembly 22, systematic identifiers and visualization of high throughput sequencing data. Nucleic Acids Res 45(D1):D592–D596
12. Nami S, Mohammadi R, Vakili M, Khezripour K, Mirzaei H, Morovati H (2019) Fungal vaccines, mechanism of actions and immunology: a comprehensive review. Biomed Pharmacother 109:333–344
13. Ravinarayanan H, Coico R, Sundar K (2015) Identification of putative therapeutic targets in *Candida tropicalis*: an *in-silico* approach. Trends Bioinform 8:52–62
14. Medici NP, Poeta MD (2015) New insights on the development of fungal vaccines: from immunity to recent challenges. Mem Inst Oswaldo Cruz 110:966–973
15. Khan SR, Baghdasarian A, Fahlman RP, Michail K, Siraki AG (2014) Current status and prospects of toxicogenomics in drug discovery. Drug Discov Today 19:562–578
16. Jorge S, Dellagostin OA (2017) The development of veterinary vaccines: a review of traditional methods and modern biotechnology approaches. Biotechnol Res Innov 1:6–13
17. Perlin DS, Richardson RR, Izquierdo AA (2017) The global problem of antifungal resistance: prevalence, mechanisms, and management. Lancet 17(12):e383–e392
18. Mochon B, Cutler JE (2005) Is a vaccine needed against *Candida albicans*? Med Mycol 43:97–115
19. Cole ST (2002) Comparative mycobacterial genomics as a tool for drug target and antigen discovery. Eur Respir J 20:78–86
20. Taki T (2013) Bio-recognition and functional lipidomics by glycosphingolipid transfer technology. Proc Jpn Acad 89(7):302–320
21. Raghuwanshi R, Singh M, Shukla V (2018) Immunoinformatic approaches in epitope prediction for vaccine designing against viral infections. Virol Immunol J 2:2
22. Redi D, Raffaelli CS, Rossetti B, Luca AD, Montagnani F (2018) *Staphylococcus aureus* vaccine preclinical and clinical development: current state of the art. New Microbiol 41:208–213

23. Csermely P, Korcsmaros T et al (2013) Structure and dynamics of molecular networks: a novel paradigm of drug discovery A comprehensive review. Pharmacol Ther 138:333–408
24. Zhang W, Li F, Nie L (2010) Integrating multiple 'omics' analysis for microbial biology: application and methodologies. Microbiology 156:287–301
25. Kumar G, Chaudhary KK, Misra K, Tripathi A (2017) Next-generation sequencing for drug designing and development: an omics approach for cancer treatment. Int J Pharm 13:709–723
26. Wheelock CE, Goss VM, Balgoma D et al (2013) Application of omics technologies to biomarker discovery in inflammatory lung diseases. Eur Respir J 42:802–825
27. Chassey B, Meyniel-Schicklin L, Aublin-Gex A, Andre P, Lotteau V (2012) New horizons for antiviral drug discovery from virus-host protein interaction networks. Curr Opin Virol 2:606–613
28. Cui T, Zeng J, He ZG (2018) Anti-tuberculosis drug target discovery by targeting the higher in-degree proteins (HidPs) of the pathogen's transcriptional network. J Tuberc 1
29. Carvalho A, Duarte-Oliveira C et al (2017) Fungal vaccines and immunotherapeutics: current concepts and future challenges. Curr Fungal Infect Rep 11:16–24
30. Pais P, Galocha M et al (2019) Microevolution of the pathogenic yeasts *Candida glabrata* during antifungal therapy and host infection. Microb Cell 6:142–159
31. Bencurova E, Gupta SK, Sarukhanyan E, Dandekar T (2018) Identification of antifungal targets based on computer modelling. J Fungi 4:81
32. Li X, Hou Y, Yue L, Liu S, Du J, Sun S (2015) Potential targets for antifungal drug discovery based on growth and virulence in *Candida albicans*. Antimicrob Agents Chemother 59:5885–5891
33. Kumar S, Kushwaha PP, Gupta S (2019) Emerging targets in cancer drug resistance. Cancer Drug Resist 2:61–77
34. Bar E, Gladiator A et al (2020) A novel Th cell epitope of *Candida albicans* mediates protection from fungal infection. J Immunol 188:5636–5646
35. Lattif AA, Mukherjee P et al (2011) Lipidomics of Candida albicans biofilms reveals phase-dependent production of phospholipid molecular classes and role for lipid rafts in biofilm formation. Microbiology 157:3232–3242
36. Cowell AN, Winzeler EA (2019) Advances in omics-based methods to identify novel targets for malaria and other parasitic protozoan infections. Genome Med 11:63
37. Sexton AE, Doerig C, Creek DJ, Carvalho TG (2019) Post-genomic approaches to understanding malaria parasite biology: linking genes to biological functions. Infect Dis 5:1269–1278
38. Becker JM, Kauffman SJ et al (2010) Pathway analysis of *Candida albicans* survival and virulence determinants in a murine infection model. PNAS 107:22044–22049
39. Bagnoli F, Baudner B et al (2011) Designing the next generation of vaccines for global public health. J Integr Biol 15:545–566
40. Toth R, Cabral V et al (2018) Investigation of *Candida parapsilosis* virulence regulatory factors during host-pathogen interaction. Sci Rep 8:1346
41. Lazo JS, McQueeney KE, Sharlow ER (2017) New approaches to difficult drug targets: the phosphatase story. SLAS Discov 22:1071–1083
42. Van Vleet TR, Liguori MJ et al (2019) Screening strategies and methods for better off-target liability prediction and identification of small-molecule pharmaceuticals. SLAS Discov 24:1–24
43. Sardi JCO, Scorzoni L, Bernardi T, Fusco-Almeida AM, Mendes Giannini MJS (2013) *Candida* species: current epidemiology, pathogenicity, biofilm formation, natural antifungal products and new therapeutic options. J Med Microbiol 62:10–24
44. Cortegiani A, Misseri G, Fasciana T, Giammanco A, Giarratano A, Chowdhary A (2018) Epidemiology, clinical characteristics, resistance, and treatment of infections by *Candida auris*. J Intensive Care 6:69
45. Maheshwari M, Kaur R, Chadha S (2016) *Candida* species prevalence profile in HIV seropositive patients from a major tertiary care hospital in New Delhi, India. J Pathog. https://doi.org/10.1155/2016/6204804

46. Malule HR, Lopez-Agudelo VA, Gomez-Rois D (2020) *Candida auris*: a bibliometric analysis of the first ten years of research (2008-2018). J Appl Pharm Sci 10:12–21
47. Zoll J, Snelders E, Verweij PE, Melchers WJE (2016) Next-Generation sequencing in the mycology lab. Curr Fungal Infect Rep 10:37–42
48. Opathy C, Gabaldon T (2019) Recent trends in molecular diagnostics of yeast infections: from PCR to NGS. FEMS Microbiol Rev 43:517–547
49. Nandikolla SK, Shaik M, Varali S, Seelam R (2011) Emerging trends in various fields with systems biology approach. J Comput Sci Syst Biol 13. https://doi.org/10.4172/0974-7230.S13-004
50. Korcsmaros T, Szalay MS, Bode S, Kovacs IA, Csermely P (2007) How to design multi-target drugs: target search options in cellular networks. Expert Opin Drug Discov 2:1–10
51. Cotugno N, Ruggiero A et al (2019) OMIC technologies and vaccine development: from the identification of vulnerable individuals to the formulation of invulnerable vaccines. J Immunol Res. https://doi.org/10.1155/2019/8732191
52. Kandpal RP, Saviola B, Felton J (2009) The era of omics unlimited. Biotechniques 46:351–355
53. Parente-Rocha J, Bailao AM et al (2017) Antifungal resistance, metabolic routes as drug targets, and new antifungal agents: an overview of endemic dimorphic fungi. Mediators Inflamm. https://doi.org/10.1155/2017/9870679
54. Roessner U, Bowne J (2009) What is metabolomics all about? Biotechniques 36:363–365
55. Roti G, Stegmaier K (2012) Genetic and proteomic approaches to identify cancer drug targets. Br J Cancer 106:254–261
56. Subhashini R, Jeyam M (2017) Computational identification of putative drug targets in *Malassezia globosa* by subtractive genomics and protein cluster network approach. Int J Pharm Pharm Sci 9:215–221
57. Xiao G, Zhang X, Gao Q (2017) Bioinformatic approaches for fungal omics. BioMed Res Int. https://doi.org/10.1155/2017/7270485