



Use of Bioinformatics in High-Throughput Drug Screening

12

Tanya Waseem, Mustafeez Mujtaba Babar, Gholamreza Abdi, and Jayakumar Rajadas

Abstract

Bioinformatics has emerged as a vital component of almost all the fields of biological sciences. Its ability to quickly generate, analyze, and interpret large amounts of data has enabled researchers to integrate it into drug discovery and development. The traditional process of drug discovery had its limitations including increased time consumption and cost, low success rates, inaccurate drug target selection, regulatory and ethical concerns, and lack of personalization. To overcome these challenges, bioinformatics has gained much interest in different stages of drug discovery. In this chapter, we summarize the role of bioinformatics in the high-throughput drug screening process involving both ligand-based and structure-based

The original version of this chapter was revised. A correction to this chapter can be found at https://doi.org/10.1007/978-981-99-8401-5_18

T. Waseem

Department of Pharmaceutical Chemistry, Shifa College of Pharmaceutical Sciences, Shifa Tameer-e-Millat University, Islamabad, Pakistan

M. M. Babar (✉)

Department of Basic Medical Sciences, Shifa College of Pharmaceutical Sciences, Shifa Tameer-e-Millat University, Islamabad, Pakistan

Advanced Drug Delivery and Regenerative Biomaterials, Stanford University School of Medicine, Stanford University, Palo Alto, CA, USA

e-mail: mustafeez.babar@fulbrightmail.org

G. Abdi

Department of Biotechnology, Persian Gulf Research Institute, Persian Gulf University, Bushehr, Iran

J. Rajadas

Advanced Drug Delivery and Regenerative Biomaterials Laboratory, Cardiovascular Institute and Pulmonary and Critical Care Medicine, Stanford University School of Medicine, Stanford University, Palo Alto, CA, USA

screening strategies. The most prominent advantage it offers is the ability to handle large amounts of data within seconds to minutes. This chapter also provides a brief account of various bioinformatics tools and databases which have a prominent role in the drug screening process. Although bioinformatics has been proven beneficial, it still has some limitations in terms of the complexity of data that is handled. With the advent of artificial intelligence and machine learning, it is expected that it would definitely strengthen the biomedical field.

Keywords

High-throughput screening · Virtual screening · Bioinformatics · Drug discovery · Drug development

12.1 Introduction

The continuous developments in the field of bioinformatics have paved ways for scientists to discover its applications in disease genetics, exploration of drug targets and drug design and discovery processes (Xia 2017). The complexity and high cost associated with the drug discovery process has led to the use of in-silico approaches to ease the drug development in connection with the experimental techniques. However, the process of drug development starting with the target discovery or identification and then designing and synthesizing drugs to modify the pathological processes is highly expensive due to the associated cost and time constraints of the drug development pipeline (Pereira et al. 2020).

Bioinformatics has been the mainstay of all scientific research for quite some time now. It is a collection of biological data that can be accessed and analyzed using computational tools and algorithms (Jawdat 2006). The major contributions towards the development of bioinformatics are the whole genome sequencing and the progress in the fields of proteomics, genomics and transcriptomics generating high-quality data which is retained in several databases and accessible for further developments (Jin et al. 2021). Deep learning has gained much attention due to its seemingly flawless performance in tasks of the machine learning including structure prediction, biological sequence analysis, protein interactions, biological diagnosis and image processing along with the prediction of biological properties and features (Li et al. 2019a). The data from the field of metabolomics has enabled researchers to study the biological processes to develop and understand the pathways and factors which are essential in physiological responses. Omics fields has influenced not only medical sciences but also other fields such as it provides considerable input in determining the factors necessary for plant growth and its molecular processes. All this is possible merely due to the availability of data which is a fruit of bioinformatics (Ambrosino et al. 2020). Era of omics has expanded the understanding of systems biology by providing valuable insights at all levels necessary for the understanding of biological systems such as proteins, transcripts, metabolites, and genes, requiring complex data processing and computations which can be done by bioinformatic approaches and tools (Waseem et al. 2020). Translational bioinformatic approaches have gained much interest for the advancements in precision medicine and they are

focused on the patient specific needs by researching on the pharmacogenomics with the aid of artificial intelligence techniques (Ritchie et al. 2019). The future of bioinformatics in health sciences is quiet promising as it aims to discover modern approaches for direct clinical practices.

12.2 High-Throughput Drug Screening

Drug discovery involves the detection or identification of drug candidates against a biological target. High-throughput screening (HTS) enables us to screen a large library of drug candidates against a selected target in multiple well plates which ultimately leads to the discovery of a novel lead compound. The major advantage of HTS is that the inactive compounds are eliminated at an early stage before pre-clinical or clinical testing of the drugs, thus saving the cost of analysis of inactive compounds. With the advancements in the area of computational chemistry and genetic biology, a number of new druggable targets have emerged and the library of synthetic and semi-synthetic drugs has grown exponentially (Kainkaryam and Woolf 2009; Hsu et al. 2021). The HTS allows us to examine the effect of a library of compounds on different targets using a single compound per well technique. In this way a library of molecules is screened in-vitro against a specific target and the inhibitors/stimulators are identified for further processing. However, the market output of the drug discovery still faces some challenges due to the unexpected and undesirable pharmacological and toxicological profile of the screened compounds in clinical trials (Scannell et al. 2012). The pharmacokinetic profile of the screened compounds also poses a limitation in the successful drug development process which is identified at a later stage and hence results in the loss of considerable time and resources. Cell-based assays although are relatively slow and expensive as compared to biochemical HTS methods but provide data related to the toxicity profile as well as the kinetic properties. These assays are used in the initial stages of drug development HTS of the library of compounds with the desirable characteristics. Cell-based HTS methods are employed while considering the quality control, and the automation techniques are carefully regulated to optimize the outcome and the development of a new chemical entity (NCE) (Schaduangrat et al. 2020).

Traditional HTS methods employ a single drug per well technique in which large number of chemical resources are wasted in exploring the active compounds as the library only have a few hits among the thousands of compounds (Volochnyuk et al. 2019). The inactive compounds are in large number and utilize great deal of time and resources even after automation. Moreover, sometimes the data results in false positive and negatives leading to the miscalculation and hence polluting the overall drug development process. Miniaturization techniques are used to overcome some of the limitations in which small amount of testing reagents and compounds are utilized and are also time efficient (Wilson et al. 2020; Wölcke and Ullmann 2001). Another strategy used for boosting the efficiency of the HTS process is the use of pooling strategy in which compounds are primarily screened as a mixture and then secondary

screening is done for the compounds with positive results from primary pooled mixtures (Kainkaryam and Woolf 2009). Hence it optimizes the resources, and reduces the error and cost. Nevertheless, the choice of pooling design and its development and implementation are some of the limitations. Combinatorial pooling strategy finds its applications in disease diagnostics as well such as one recent study identified the use of HTS for SARS-CoV-2 diagnostic testing for asymptomatic patients who are a carrier of the virus and pose a significant threat to the disease spread (Shental et al. 2020).

12.3 Bioinformatics in High-Throughput Drug Screening

With the advent of bioinformatic tools that complement the drug discovery process, we have overcome many of the limitations and problems encountered in late 90 s in the processes of HTS. Bioinformatics has emerged as a multidisciplinary field which is a vital part of drug discovery and development process such as screening of compound libraries, identification of biological targets, proteomics, genomics, biological, chemical and virtual screening of compounds. Cheminformatics along with the bioinformatics has resolved various problems encountered in the drug development and screening process (Parikh et al. 2023). Bioinformatic techniques are employed for the identification of novel drug targets, modelling of target proteins, designing of druggable compounds, determining their interactions with the target and prediction of physicochemical properties and toxicology profiling. Machine learning techniques and algorithms are being developed to aid these processes (Chavda et al. 2021).

Various data mining tools are being used which provide large datasets for the identification of potential targets as well as the compounds that will bind with those targets and produce a response. These tools also enable us to establish the effectiveness of the drug candidate and its binding interactions (Yang et al. 2012; Patel et al. 2020).

Molecular modelling tools are used to generate models of the target proteins and biological systems using different techniques which enable us to virtually analyze the target structure and predict the binding site as well as the binding interactions or the groups needed for potential binding (Haghighatlari and Hachmann 2019).

Virtual high-throughput screening tools are developed to overcome the limitations of traditional HTS systems which results in loss of reagents and resources. When the HTS assays are complex and tedious, virtual drug screening approaches are used to complement the HTS. The drug libraries are screened using in-silico experimentation to determine their binding interactions with biological targets using molecular modelling tools (Mcintosh-Smith et al. 2015). Such techniques enable scientists to screen large libraries in an efficient manner without the expenditure of viable resources. However, these tools require considerable expertise to operate, operate using complex algorithms and are not always error-free. One of the major advantages of using virtual HTS is that it is economical and less time consuming as compared to the traditional experimentation; using large

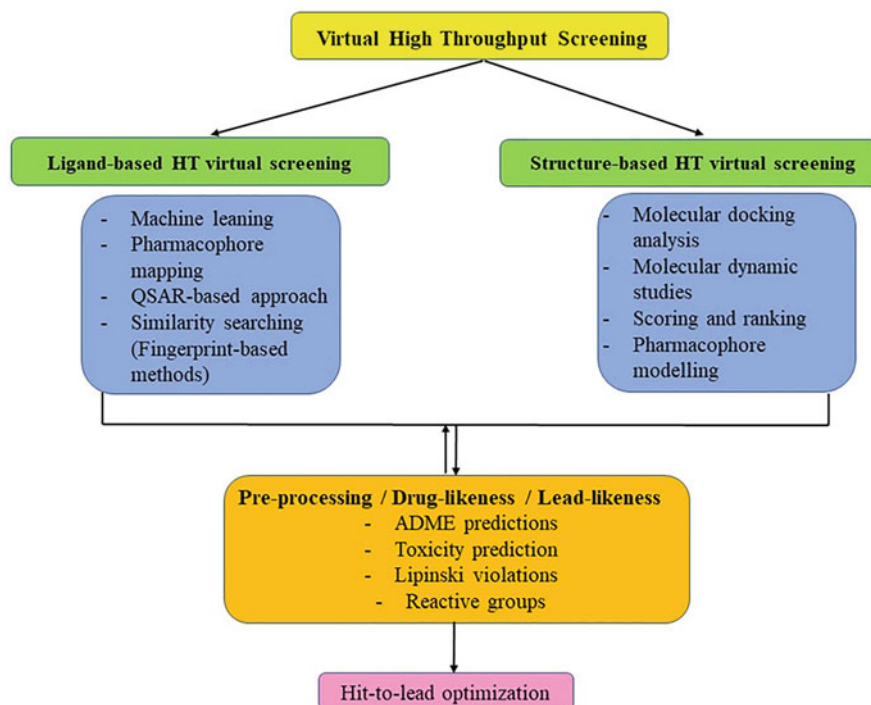


Fig. 12.1 Virtual high-throughput screening strategies for lead identification (Stumpfe and Bajorath 2020; Guterres and Im 2020; Da Silva Rocha et al. 2019)

volume of reagents to screen for a potential active agent from millions of compounds (Mohammad et al. 2021). Structure-based and ligand-based virtual screening are two strategies for hit-to-lead discovery and optimization (Fig. 12.1). These approaches are purely theoretical when compared to the HTS which is purely experimental technique but both aim at the generation of a lead compound for the successful drug discovery process. A combination of both approaches can help in the efficient delivery of successful drug candidates without incurring resource wastage and added costs of analyzing thousands of compounds in HTS experiments (Zhang et al. 2022).

One of the methods for virtual screening is molecular docking which involves the study of interaction between the drug molecule and the biological target followed by the analysis of binding energies and interacting amino acid residues of the binding pocket. The spatial arrangement of drug with its target is based on the induced fit theory and results in the identification of its mechanism of action (Lin et al. 2020). Pharmacophore modelling is another method which is used to design basic structural model of the drug candidate from which lead compound is generated, followed by the screening of databases. The structural features of a pharmacophore are based on its complementary target and by adjusting these features compounds with desirable activity can be designed. Another way is to proceed to the screening of small

molecules based on chemical similarity searching on various databases. ZINC is one such database (Seidel et al. 2017; Lin et al. 2020).

Quantitative Structure-activity relationship (QSAR) is a technique in which some quantifiable property of a compound is correlated with its biological activity based on experimental data. Quantitative descriptors are used to identify the active agents against a target of interest by comparing their features such as lipid solubility, permeability, electronic features, size and shape of the molecule and ADME properties. 3D QSAR modelling is still used in pharmaceutical industry due to its ability of accurate structural predictions using minimal calculations (Vucicevic et al. 2019). In-silico screening methods are used for the development of drugs for a wide range of diseases such as tuberculosis (Macalino et al. 2020), CVDs (Savoji et al. 2019), COVID-19 (Gupta et al. 2023), hepatitis (Hdoufane et al. 2022), diabetes (Akhtar et al. 2019), neurodegenerative diseases (Aldewachi et al. 2021) and cancer therapy (Vougas et al. 2019).

12.4 Applications of Bioinformatics in High-Throughput Drug Screening

Omics technology has emerged as a turning point in the health sciences which provides data related to the biological systems and includes proteomics, genomics, metabolomics and transcriptomics. The first step in the HTS is the target identification and various drug targets have been discovered and identified with the help of bioinformatic approaches (Martis et al. 2011). Data mining approaches include high-throughput chemogenomic and proteomics. A wide range of data mining sources are available which have all the necessary information needed for the identification of a biological target such as structural databases (UniPort, PubMed, InterPro), text mining tools (GeneWays, Texpresso, BioRat), microarray databases (SMD, Oncomine, caArray), clustering database (GenePattern, ArrayMiner, Genecluster), supervised analysis platform (SAM) and interactome and pathway databases (KEGG, PathwayExplorer, Pathguide) (Yang et al. 2012; Agamah et al. 2020). One study reported the use of various bioinformatic tools and databases to develop a human-virus interactome for ZIKA virus using an algorithm OralInt, potentially highlighting various druggable targets against ZIKA virus (Fig. 12.2) (Esteves et al. 2017).

Assay development is a crucial step for the success of screening process. Specificity and sensitivity of assay is the basis of the whole experimentation and bioinformatic techniques have been utilized to develop highly sensitive screening assays. Virtual screening assays are developed as a complementary approach to the HTS and can be regarded as a basic simulation of the HTS assays using the knowledge of biophysics and computer sciences. These simulations are also conducted to optimize the conditions needed to run an assay. In the simulation models, various parameters such as temperature, reagents and time duration can be adjusted leading to the highly sensitive assay. MolMind is one such tool which combines the laboratory based assays and in-silico methods (Szymański et al. 2012). In-silico toxicological analysis

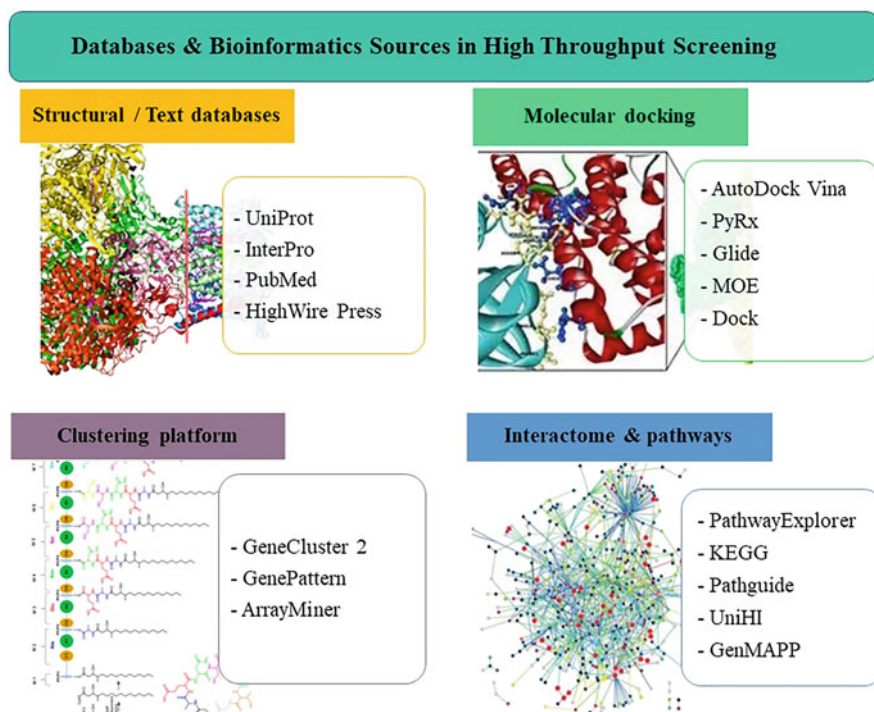


Fig. 12.2 Databases and bioinformatic sources in high-throughput drug screening (Yang et al. 2012)

is a preliminary assay which results in the filtering of potential toxicological compound while virtual screening. ADME-T methods are being used along with computational toxicology methods (i-drug discovery, ToxScope, OncoLogic, MetaDrug, HazardExpert, and e-TOX) for determining the toxicological profile of the drug candidates (Szymański et al. 2012). In one study, imaging techniques and fluorescent-based methods were combined to create a high-throughput drug screening assay using 3D organoids to assess the organoid growth and the effects of drugs (Li et al. 2022).

Data mining approaches and microarray techniques are utilized for HTS. One study reported the use of microarray analysis for the identification of micro RNAs and genes as biomarkers for the treatment and diagnosis of atrial fibrillation using different databases (Li et al. 2019b). The study of biological pathways utilizing bioinformatic tools has made it possible to identify the disease biomarkers and drug targets. One study reported the involvement of multiple RNAs expression in the regulation and progression of preeclampsia using different bioinformatics tools and databases. It also reported that the activation of JAK-STAT signaling pathway is related to the progression of preeclampsia (Liu et al. 2019).

12.5 Challenges and Limitations of Bioinformatics in High-Throughput Drug Screening

Virtual HTS is an efficient, robust and cost-effective technique for the screening of biologically active molecules from large datasets but it does not replace the traditional HTS methods, it simply complements it by narrowing down the possible hits and leads. Computational analysis, although, enables us to screen for a library of thousands of compounds in a day but it faces some challenges due to the complexity of data and sometimes generates erroneous results. The main focus, however, remains on the generation of efficient leads for subsequent optimization in drug development pipelines. The complexity of data generated by computational analysis is also challenging for the effective interpretation and requires highly skilled analysis. Machine learning techniques such as decision tree models and artificial neural networks are developed to overcome the complexity of data available for the computational analysis (Han et al. 2008; Butkiewicz et al. 2012). The computations and equations of QSAR models are highly complex and requires careful analysis, validations and may sometimes be impractical (Spiegel and Senderowitz 2020). The computational complexity of various multilayered techniques is a hindrance. The method development requires validation of data and sometimes the results are not reproducible raising a question on the validity of the data obtained (Stumpfe and Bajorath 2020). One of the major challenges highlighted and mentioned by multiple researchers in the field is the accuracy of data obtained from virtual screening methods. Sometimes its impractical to translate the outcome in human patients although significant evidence of activity is obtained from computational analysis. In structure-based drug screening, the binding energies of actives and inactive are closely related showing inaccuracy and sometimes putative interactions are generated for inactive leads which like HTS results in the generation of ineffective lead compounds identified at a later stage of testing (Jasial et al. 2016).

Current debate is on the ligand promiscuity of the biological targets which may points towards the inaccuracy of the binding interactions generated through the virtual screening. With the prior knowledge of drug-target binding interactions, virtual screening methods also faces a certain bias in the selection of screening library which leads to the high hit rates confused with the accuracy of the prediction (Stumpfe and Bajorath 2020).

12.6 Future of Bioinformatics in High-Throughput Drug Screening

Bioinformatics has emerged as an indispensable field in the drug discovery and screening processes. The traditional high-throughput screening requires the experimentation of large library of compounds having millions of drugs comprising of large number of inactive candidates. This resulted in wastage of resources, time and money. Bioinformatic tools and techniques enable us to shrink down the chemical library before high-throughput screening assays by ruling out the possible inactive

agents in in-silico or virtual screening steps. These virtual screening methods enable us to identify and select only those compounds which show promising results in virtual screening assays (Stumpfe et al. 2012; Stumpfe and Bajorath 2020). Hence these techniques save cost, resources and time by providing highly specific and nearly accurate predictions. Virtual screening era is promising and is predicted to progress further mainly due to its screening efficiency and enormous data handling capacity. Nevertheless, the virtual screening problems need to be encountered in the future to continue an integrative approach towards drug screening. The number one problem which requires attention is the generation of inaccurate binding energies and similarity hits; which require rigorous post-analysis to interpret the accuracy of results. Scientists are working to overcome this problem and have made some progress. In this post-genomic era, the field of molecular and chemical biology remain potential areas of growth that will enhance our understanding as well as the applications of virtual drug screening (Heikamp and Bajorath 2012; Sabe et al. 2021).

The advancements in the field of artificial intelligence are a turning point for the pharmaceutical and health sciences as it is a step forward towards overcoming the limitations encountered in drug discovery (Zhong et al. 2018). Virtual screening is indispensable in drug discovery and development process. Sequential screening which is a widely known concept; computational screening integrated with experimental screening, should be practically incorporated in order to avoid problems at a later stage and overcome the limitations of both techniques (Achary 2020).

12.7 Conclusions and Future Perspectives

The field of bioinformatics has significantly contributed to the drug discovery and development process by providing an avenue though virtual high-throughput screening. However, the vast amount of unverified data available on genetic and protein repositories makes it essential for the bioinformaticians to pre-process it before its integration and interpretation can actually begin. In addition, biological complexity of available data hinders its wider usage. Experimental limitations and lack of availability and accessibility to a variety of user-friendly computer applications also appears to slow-down the HTS process. The advent of publicly available machine learning and artificial intelligence platforms can address some of the identified limitations. Moreover, the inter-disciplinary collaborative research can facilitate the drug development process.

References

- Achary PGR (2020) Applications of quantitative structure-activity relationships (QSAR) based virtual screening in drug design: a review. *Mini Rev Med Chem* 20:1375–1388

- Agamah FE, Mazandu GK, Hassan R, Bope CD, Thomford NE, Ghansah A, Chimusa ER (2020) Computational/in silico methods in drug target and lead prediction. *Brief Bioinform* 21:1663–1675
- Akhtar A, Amir A, Hussain W, Ghaffar A, Rasool N (2019) In silico computations of selective phytochemicals as potential inhibitors against major biological targets of diabetes mellitus. *Curr Comput Aided Drug Des* 15:401–408
- Aldewachi H, Al-Zidan RN, Conner MT, Salman MM (2021) High-throughput screening platforms in the discovery of novel drugs for neurodegenerative diseases. *Bioengineering* 8:30
- Ambrosino L, Colantuono C, Diretto G, Fiore A, Chiusano ML (2020) Bioinformatics resources for plant abiotic stress responses: state of the art and opportunities in the fast evolving-omics era. *Plan Theory* 9:591
- Butkiewicz M, Lowe EW, Mueller R, Mendenhall JL, Teixeira PL, Weaver CD, Meiler J (2012) Benchmarking ligand-based virtual high-throughput screening with the PubChem database. *Molecules* 18:735–756
- Chavda V, Sheta S, Changani D, Chavda D (2021) New bioinformatics platform-based approach for drug design. In: Balamurugan S, Krishnan A, Goyal D, Chandrasekaran B, Pandi B (eds) *Computation in bioinformatics: multidisciplinary applications*. Wiley, pp 101–120
- Da Silva Rocha SFL, Olanda CG, Fokoue HH, Sant'anna, C. M. R. (2019) Virtual screening techniques in drug discovery: review and recent applications. *Curr Top Med Chem* 19:1751–1767
- Esteves E, Rosa N, Correia MJ, Arrais JP, Barros M (2017) New targets for *Zika* virus determined by human-viral interactomic: a bioinformatics approach. *Biomed Res Int* 2017:1734151
- Gupta Y, Savvitskiy OV, Coban M, Venugopal A, Pleqi V, Weber CA, Chitale R, Durvasula R, Hopkins C, Kempaiah P (2023) Protein structure-based in-silico approaches to drug discovery: guide to COVID-19 therapeutics. *Mol Asp Med* 91:101151
- Guterres H, Im W (2020) Improving protein-ligand docking results with high-throughput molecular dynamics simulations. *J Chem Inf Model* 60:2189–2198
- Haghighatlari M, Hachmann J (2019) Advances of machine learning in molecular modeling and simulation. *Curr Opin Chem Eng* 23:51–57
- Han L, Wang Y, Bryant SH (2008) Developing and validating predictive decision tree models from mining chemical structural fingerprints and high-throughput screening data in PubChem. *BMC Bioinform* 9:401
- Hdoufane I, Bjjj I, Oubahmane M, Soliman ME, Villemin D, Cherqaoui D (2022) In silico design and analysis of NS4B inhibitors against hepatitis C virus. *J Biomol Struct Dyn* 40:1915–1929
- Heikamp K, Bajorath J (2012) The future of virtual compound screening. *Chem Biol Drug Des* 81:33–40
- Hsu MN, Tay ZM, Lin WN, Wei S-C (2021) Screening of antigen-specific antibody-secreting cells. *Handbook of single-cell technologies*. Springer
- Jasial S, Hu Y, Vogt M, Bajorath J (2016) Activity-relevant similarity values for fingerprints and implications for similarity searching [version 2; peer review: 3 approved]. *F1000Res* 5
- Jawdat D (2006) The era of bioinformatics. In 2006 2nd international conference on information & communication technologies. *IEEE*, p 1860–1865
- Jin S, Zeng X, Xia F, Huang W, Liu X (2021) Application of deep learning methods in biological networks. *Brief Bioinform* 22:1902–1917
- Kaikaryam RM, Woolf PJ (2009) Pooling in high-throughput drug screening. *Curr Opin Drug Discov Devel* 12:339
- Li Y, Huang C, Ding L, Li Z, Pan Y, Gao X (2019a) Deep learning in bioinformatics: introduction, application, and perspective in the big data era. *Methods* 166:4–21
- Li Y, Tan W, Ye F, Xue F, Gao S, Huang W, Wang Z (2019b) Identification of microRNAs and genes as biomarkers of atrial fibrillation using a bioinformatics approach. *J Int Med Res* 47:3580–3589

- Li X, Fu G, Zhang L, Guan R, Tang P, Zhang J, Rao X, Chen S, Xu X, Zhou Y, Deng Y, Lv T, He X, Mo S, Mu P, Gao J, Hua G (2022) Assay establishment and validation of a high-throughput organoid-based drug screening platform. *Stem Cell Res Ther* 13:219
- Lin X, Li X, Lin X (2020) A review on applications of computational methods in drug screening and design. *Molecules* 25:1375
- Liu S, Xie X, Lei H, Zou B, Xie L (2019) Identification of key circRNAs/lncRNAs/miRNAs/mRNAs and pathways in preeclampsia using bioinformatics analysis. *Med Sci Monit* 25:1679–1693
- Macalino SJY, Billones JB, Organo VG, Carrillo MCO (2020) In silico strategies in tuberculosis drug discovery. *Molecules* 25:665
- Martis E, Radhakrishnan R, Badve R (2011) High-throughput screening: the hits and leads of drug discovery—an overview. *J Appl Pharm Sci*:02–10
- Mcintosh-Smith S, Price J, Sessions RB, Ibarra AA (2015) High performance in silico virtual drug screening on many-core processors. *Int J High Perform Comput Appl* 29:119–134
- Mohammad T, Mathur Y, Hassan MI (2021) InstaDock: a single-click graphical user interface for molecular docking-based virtual high-throughput screening. *Brief Bioinform* 22:bbaa279
- Parikh PK, Savjani JK, Gajjar AK, Chhabria MT (2023) Bioinformatics and cheminformatics tools in early drug discovery. In: *Bioinformatics tools for pharmaceutical drug product development*, pp 147–181
- Patel L, Shukla T, Huang X, Ussery DW, Wang S (2020) Machine learning methods in drug discovery. *Molecules* 25:5277
- Pereira SA, Dyson PJ, Saraiva MLM (2020) Miniaturized technologies for high-throughput drug screening enzymatic assays and diagnostics—a review. *TrAC Trends Anal Chem* 126:115862
- Ritchie MD, Moore JH, Kim JH (2019) Translational bioinformatics: biobanks in the precision medicine era. *Pacific symposium on biocomputing 2020*. World Scientific, p 743–747
- Sabe VT, Ntombela T, Jhamba LA, Maguire GEM, Govender T, Naicker T, Kruger HG (2021) Current trends in computer aided drug design and a highlight of drugs discovered via computational techniques: a review. *Eur J Med Chem* 224:113705
- Savoji H, Mohammadi MH, Rafatian N, Toroghi MK, Wang EY, Zhao Y, Korolj A, Ahadian S, Radisic M (2019) Cardiovascular disease models: a game changing paradigm in drug discovery and screening. *Biomaterials* 198:3–26
- Scannell JW, Blanckley A, Boldon H, Warrington B (2012) Diagnosing the decline in pharmaceutical R&D efficiency. *Nat Rev Drug Discov* 11:191–200
- Schaduangrat N, Lampa S, Simeon S, Gleeson MP, Spjuth O, Nantasenamat C (2020) Towards reproducible computational drug discovery. *J Chem* 12:1–30
- Seidel T, Bryant SD, Ibis G, Poli G, Langer T (2017) 3D pharmacophore modeling techniques in computer-aided molecular design using LigandScout. In: *Tutorials in cheminformatics*, pp 279–309
- Shental N, Levy S, Wuvshet V, Skorniakov S, Shalem B, Ottolenghi A, Greenshpan Y, Steinberg R, Edri A, Gillis R (2020) Efficient high-throughput SARS-CoV-2 testing to detect asymptomatic carriers. *Sci Adv* 6:eabc5961
- Spiegel J, Senderowitz H (2020) Evaluation of QSAR equations for virtual screening. *Int J Mol Sci* 21:7828
- Stumpfe D, Bajorath J (2020) Current trends, overlooked issues, and unmet challenges in virtual screening. *J Chem Inf Model* 60:4112–4115
- Stumpfe D, Ripphausen P, Bajorath J (2012) Virtual compound screening in drug discovery. *Future Med Chem* 4:593–602
- Szymański P, Markowicz M, Mikiciuk-Olasik E (2012) Adaptation of high-throughput screening in drug discovery—toxicological screening tests. *Int J Mol Sci* 13:427–452
- Volochnyuk DM, Ryabukhin SV, Moroz YS, Savych O, Chuprina A, Horvath D, Zabolotna Y, Varnek A, Judd DB (2019) Evolution of commercially available compounds for HTS. *Drug Discov Today* 24:390–402

- Vougas K, Sakellaropoulos T, Kotsinas A, Foukas G-RP, Ntargaras A, Koinis F, Polyzos A, Myriantopoulos V, Zhou H, Narang S, Georgoulis V, Alexopoulos L, Aifantis I, Townsend PA, Sfikakis P, Fitzgerald R, Thanos D, Bartek J, Petty R, Tsirigos A, Gorgoulis VG (2019) Machine learning and data mining frameworks for predicting drug response in cancer: an overview and a novel in silico screening process based on association rule mining. *Pharmacol Ther* 203:107395
- Vucicevic J, Nikolic K, Mitchell JB (2019) Rational drug design of antineoplastic agents using 3D-Qsar, cheminformatic, and virtual screening approaches. *Curr Med Chem* 26:3874–3889
- Waseem T, Zargaham MK, Shahid F, Rajput TA, Ibrahim B, Babar MM (2020) New approaches to antimicrobial discovery: current development and future prospects. In: *New and future developments in microbial biotechnology and bioengineering*. Elsevier, pp 67–77
- Wilson BA, Thornburg CC, Henrich CJ, Grkovic T, O'keefe, B. R. (2020) Creating and screening natural product libraries. *Nat Prod Rep* 37:893–918
- Wölcke J, Ullmann D (2001) Miniaturized HTS technologies—uHTS. *Drug Discov Today* 6:637–646
- Xia X (2017) Bioinformatics and drug discovery. *Curr Top Med Chem* 17:1709–1726
- Yang Y, Adelstein SJ, Kassis AI (2012) Target discovery from data mining approaches. *Drug Discov Today* 17:S16–S23
- Zhang Y, Luo M, Wu P, Wu S, Lee T-Y, Bai C (2022) Application of computational biology and artificial intelligence in drug design. *Int J Mol Sci* 23:13568
- Zhong F, Xing J, Li X, Liu X, Fu Z, Xiong Z, Lu D, Wu X, Zhao J, Tan X, Li F, Luo X, Li Z, Chen K, Zheng M, Jiang H (2018) Artificial intelligence in drug design. *Sci China Life Sci* 61: 1191–1204