

Unraveling allosteric landscapes of allosterome with ASD

Xinyi Liu^{1,2,†}, Shaoyong Lu^{1,2,†}, Kun Song^{1,2,†}, Qiancheng Shen^{1,2,3,†}, Duan Ni^{1,†}, Qian Li^{2,3}, Xinheng He^{1,2}, Hao Zhang², Qi Wang⁴, Yingyi Chen^{1,2}, Xinyi Li^{1,2}, Jing Wu^{2,3}, Chunquan Sheng^{5,6}, Guoqiang Chen¹, Yaqin Liu¹, Xuefeng Lu^{3,*} and Jian Zhang^{1,2,6,*}

¹State Key Laboratory of Oncogenes and Related Genes, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200127, China, ²Medicinal Bioinformatics Center, Shanghai Jiao Tong University School of Medicine (SJTU-SM), Shanghai 200025, China, ³Department of Assisted Reproduction, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine (SJTU-SM), Shanghai 200011, China, ⁴China National Pharmaceutical Industry Information Center, Shanghai, 200040, China, ⁵School of Pharmacy, Second Military Medical University, Shanghai, 200433, China and ⁶School of Pharmaceutical Sciences, Zhengzhou University, Zhengzhou 450001, China

Received September 02, 2019; Revised September 30, 2019; Editorial Decision October 04, 2019; Accepted October 10, 2019

ABSTRACT

Allosteric regulation is one of the most direct and efficient ways to fine-tune protein function; it is induced by the binding of a ligand at an allosteric site that is topographically distinct from an orthosteric site. The Allosteric Database (ASD, available online at <http://mdl.shsmu.edu.cn/ASD>) was developed ten years ago to provide comprehensive information related to allosteric regulation. In recent years, allosteric regulation has received great attention in biological research, bioengineering, and drug discovery, leading to the emergence of entire allosteric landscapes as allosteromes. To facilitate research from the perspective of the allosterome, in ASD 2019, novel features were curated as follows: (i) >10 000 potential allosteric sites of human proteins were deposited for allosteric drug discovery; (ii) 7 human allosterome maps, including protease and ion channel maps, were built to reveal allosteric evolution within families; (iii) 1312 somatic missense mutations at allosteric sites were collected from patient samples from 33 cancer types and (iv) 1493 pharmacophores extracted from allosteric sites were provided for modulator screening. Over the past ten years, the ASD has become a central resource for studying allosteric regulation and will play more important roles in both

target identification and allosteric drug discovery in the future.

INTRODUCTION

The concept of allostery, or allosteric regulation, proposed in the 1960s, describes the existence of indirect coupling between two topographically and spatially distinct types of binding sites: allosteric and orthosteric sites (1,2). The binding of effectors (e.g. small molecules, ions, and DNA/RNA) to an allosteric site enables a protein to propagate perturbations from an allosteric site to an orthosteric site, ultimately achieving exquisite control of the protein's functional activities (3–7). Compared to the relatively conserved orthosteric sites in homologous protein families, the targeting of structurally less conserved allosteric sites by allosteric modulators confers the advantages of increased specificity and reduced side effects (8–14). In this context, the identification of allosteric modulators for a desired target is the subject of intense drug discovery efforts (15–20).

Despite the pharmacological merits of allosteric modulators, their discovery has faced a long-standing conundrum because of the difficulty in identifying allosteric sites via the experimental determination of allosteric protein-modulator complexes (21). Remarkably, recent computational advances in the detection of allosteric sites have significantly facilitated the identification of allosteric modulators for related targets (22–29), including sirtuin 6 (SIRT6) (15), signal transducer and activator of transcription 3 (STAT3) (30), glutathione peroxidase 4 (GPX4) (31) and

*To whom correspondence should be addressed. Tel: +86 21 63846590 776922; Fax: +86 21 64154900; Email: jian.zhang@sjtu.edu.cn
Correspondence may also be addressed to Xuefeng Lu. Tel: +86 21 23271699 5539; Fax: +86 21 53078108; Email: xuefenglu163@163.com

†The authors wish it to be known that, in their opinion, the first five authors should be regarded as Joint First Authors.

cathepsin K (32). Based on the many computational allosteric approaches contributing to practical applications, we have proposed that allosteric modulator discovery has evolved from the realm of serendipity to structure-based drug design (33). Therefore, it is necessary to comprehensively annotate potential allosteric sites in known human allosteric proteins with both computational predictions and experimental determinations. These sites are useful for virtual screening and subsequent medicinal chemistry optimization of allosteric modulators.

As for orthosteric drugs, somatically acquired mutations at allosteric sites, referred to as allosteric mutations, may confer resistance to allosteric therapeutics (34–36). For example, single point mutations (A337V, P465S, V468F, and I502L) at the allosteric myristoyl binding site in the C-lobe of BCR-ABL kinase cause insensitivity to the allosteric inhibitor ABL001 in chronic myeloid leukemia (37). Analogously, the gatekeeper V211D mutation in the allosteric site of the N-lobe of MEK1 confers resistance to the allosteric inhibitor binimetinib in BRAF K601E-associated colon cancer (38). As such, it is advisable to systematically annotate known allosteric mutations from clinical samples. A comprehensive understanding of allosteric mutations, particularly clinically related variations, thus broadens the scope of available drug targets.

We have created a large, comprehensive database, the Allosteric Database (ASD), that has provided comprehensive information focusing on allosteric regulation since 2009 (39–41). Over the past decade, it has been visited >63 000 times from >100 countries around the world. ASD has garnered high acclaim from the scientific community (19,20,42) and has become a valuable resource for versatile allosteric applications. The fruitful results include the construction of benchmarking datasets for high-quality allosteric sites (AS-Bench (43)), the prediction of allosteric sites (e.g. Allosite (44) /AllositePro (45), AlloPred (46), PARS(47), CavityPlus (48), CryptoSite (49), pocketZebra (50), PocketDB (51), ALLO (52), ExProSE (53)), the assessment of allosteric interaction (Alloscore (54)), the exploration of allosteric communication (e.g. MCPATH (55), AlloSigMA (26)), the analysis of allosteric mutations and evolution (e.g. AlloDriver (27), AlloMAPS (25), Trans-Omics (56)), and the screening of allosteric modulators (AlloFinder (30)). Such knowledge has promoted an upsurge in allosteric research, leading to the emergence of the framework of the allosterome. To further understand the landscape of the allosterome, in ASD 2019, we provide new allosteric features, including (i) the introduction of computational prediction of the locations of potential allosteric sites for known human allosteric proteins, (ii) the addition of allosteric mutations at allosteric sites of human proteins through large-scale exploration of somatic mutations in clinical samples across 33 cancer types, (iii) the construction of 7 new human allosterome maps and (iv) the development of allosteric site-based pharmacophore models. In addition, allosteric data such as experimentally determined allosteric proteins, sites, and modulators as well as allosteric-related diseases and allosteric drugs in different phases of clinical trials have also been updated and curated. Overall, the current ASD is useful for achieving a comprehensive understanding of the allosterome, facilitating the investigation of allosteric regu-

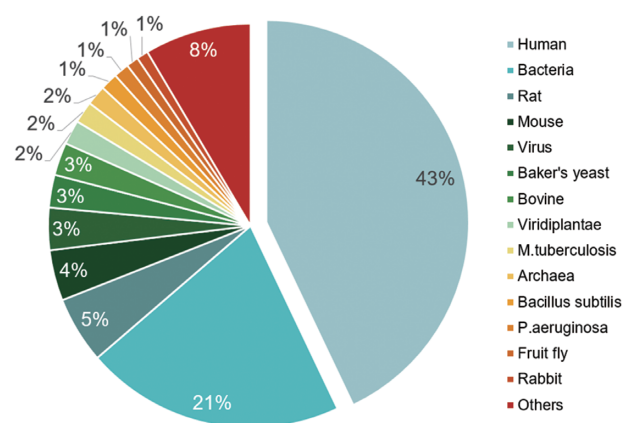


Figure 1. Statistics for allosteric proteins in ASD 2019.

latory mechanisms and evolution and the discovery of allosteric drugs.

DATABASE GROWTH AND STATISTICS

New allosteric molecules and features, including structures, affinities, sites, functions, related diseases and external links, were collected using the methods described in our previous publications (39–41). According to the statistical data for ASD 2019 provided in Table 1, the known allosteric proteins have experienced a rapid increase over the past three years. Currently, 1949 allosteric proteins from various species are archived in the ASD, representing an increase of more than 32% (476 new proteins) since the last version. The allosteric phenomenon has been widely observed in 22 different protein categories defined by UniProt keyword ontology (57), and the growth of the data on allosteric proteins has been highly concentrated in the following categories: ion channels (from 93 to 154), GPCRs (from 114 to 169), kinases (from 201 to 249), and transferases (kinases excluded, from 21 to 119), which play key roles in disease therapies and biological pathways (3,4). Findings of additional classes of proteins have also confirmed the intrinsic properties of allostery, such as 147 hydrolases (proteases and phosphatases excluded), 126 oxidoreductases, 122 transcription factors, 120 transporters (ion channels excluded) and 87 lyases. The allosteric proteins in ASD 2019 come from 274 different species, and nearly 75% of the records belong to humans (43%) and bacterial species (31%) (see Figure 1). Notably, more than half of the 476 new proteins are from humans (244), and the acceleration of human record accumulation has contributed greatly to allosterome development. In addition to proteins, ASD 2019 contains 82070 allosteric modulators, composed of 31376 activators, 37471 inhibitors and 14622 regulators. Among the 10532 new curated modulators, 94% (9891) are exogenous allosteric lead compounds for human targets developed in the last three years, indicating a significant acceleration of the progress of allosteric drug discovery.

In addition to these allosteric molecules, the data on their allosteric features and annotations are expanded in ASD 2019, including an ~19% increase in allosteric interactions (from 75462 to 89554), an ~32% increase in al-

Table 1. Data statistics for allosteric proteins and modulators in the updated ASD 2019

Data category	ASD 2019	ASD 2016
Number of all proteins^a	1949	1473
Number of kinases	249	201
Number of transferases	185	144
Number of GPCRs	169	114
Number of ion channels	154	93
Number of hydrolases	147	117
Number of oxidoreductases	126	97
Number of transcription factors	122	94
Number of transporters	120	103
Number of proteases	103	78
Number of lyases	87	69
Number of other proteins ^b	487	363
Number of all modulators	82 070	71 538
Number of protein-modulator interactions	89 554	75 462
Number of crystal/NMR protein structures	26 363	11 683
Number of protein-modulator complex structures	2542	1930
Number of potential allosteric sites	10 081	0
Number of allosterome maps	9	2
Number of allosteric mutations	1312	0
Statistics of allosteric drug dataset		
Number of allosteric drugs	538	49
Number of allosteric drug targets	95	20
Number of allosteric target-drug interactions	638	56
Number of allosteric-related diseases	5983	3350

^aIn the definition of the classifications for allosteric proteins in ASD 2019, several subtypes have been removed from the enzyme categories, including transferases (kinases excluded), hydrolases (proteases and phosphatases excluded), transporters (ion channels excluded), and receptors (nuclear receptors excluded).

^bProtein categories including >50 proteins are displayed, while others are included in the ‘other proteins’.

losteric protein-modulator complexes (from 1930 to 2542), an ~79% increase in related diseases (from 3350 to 5983), a 126% increase in protein structures (from 11683 to 26363), and a nearly 10-fold increase in allosteric chemicals in different phases of drug development (i.e. allosteric drugs, from 49 to 538). These allosteric drugs are well organized in the ‘ALLO-DRUG’ category under the FEATURES menu and classified into five phases: preclinical (451), phase I (23), phase II (36), phase III (9) and approved (19). Further analysis revealed that the targets associated with allosteric drugs include not only the traditional ion channels, GPCRs, and kinases but also many new types of proteins, including transporters, oxidoreductases, ligases, etc. These consistently rich data illustrate the ongoing expansion of the landscapes of the allosterome.

NEW FEATURES AND FUNCTIONS

To further investigate the landscapes of the allosterome, novel allosteric features have been developed and organized in ASD 2019, including ‘ALLOSITE-POTENTIAL’ for potential allosteric sites in human proteins, ‘ALLOSTEROME’ for the evolution of allostery in a protein family, ‘ALLO-MUTATION’ for allosteric mutations in cancers, and ‘ALLO-DRUG’ for allosteric modulators in the pre-clinical and clinical phases. In addition to these data, two allosteric tools, ‘ALLO-PHARM’ and ‘ALLO-PATHWAY’, are integrated into the ‘TOOLS’ menu of the ASD, providing computational protocols for screening allosteric modulators and detecting allosteric pathways from orthosteric sites to allosteric sites. Figure 2 shows the workflow of the major ASD components or their combinations applied to allosteric drug design, target identification, mechanism re-

search and allosterome analysis. The details of each new feature are described below.

POTENTIAL ALLOSTERIC SITES

Allosteric sites tend to be under lower sequence-conservation pressure and exhibit more structural diversity than conserved orthosteric sites, endowing allosteric modulators with the potential to achieve higher specificity, fewer side effects and lower toxicity (43–45,47,58). However, the slow pace of experimental validation seriously hinders the disclosure of the entire allosterome; allosteric sites have been identified in only a few proteins. Therefore, a large-scale dataset containing the potential allosteric sites of proteins is highly desirable and was therefore built in the current release. In ASD 2019, 10081 potential allosteric sites predicted from 4013 human proteins were constructed using our AllositePro method (45) (see Supporting Information) and are available under ‘ALLOSITE-POTENTIAL’ in the ‘FEATURES’ menu. These sites may be selectively shown through a combination of checkbox filters and protein subclass tree maps. Clicking the selected human protein causes an interactive page to open, and this page provides the details of potential allosteric sites, including protein information, a 3D representation window, a site selection panel, and a grid showing site properties. The proteins in the ALLOSITE-POTENTIAL dataset account for 75% of human proteins with PDB structures (59) and are widely distributed in all protein categories. These potential allosteric sites could provide an effective resource leading to the identification of allosteric regulation throughout the human proteome and be useful for allosteric drug discovery.

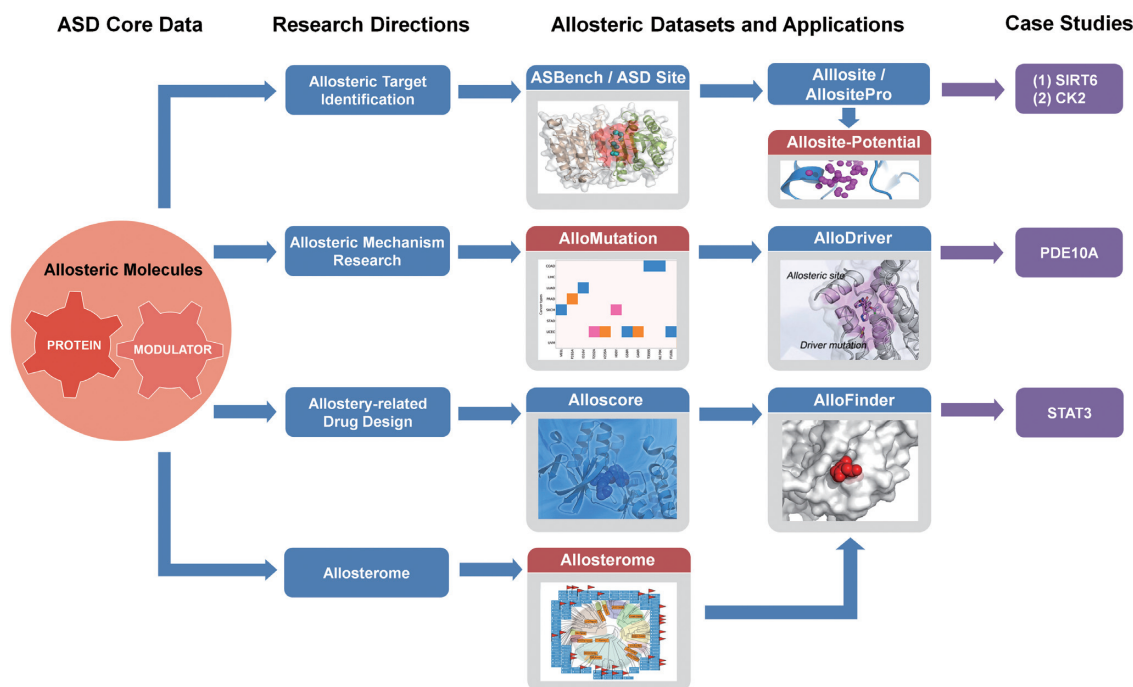


Figure 2. The workflow of the ASD major components or their combinations used for allosteric target identification, allosteric mechanism research, allostery-related drug design, and allosterome analysis. (i) The high-quality allosteric site benchmarking dataset (ASBench) derived from the ASD can be used to develop computational allosteric site prediction methods such as Allosite and AllositePro. The real cases include the identification of the SIRT6 allosteric activators MDL-800 and MDL-801 and a novel CK2 α allosteric site. The Allosite-Potential datasets constructed by AllositePro could also provide an effective way for allosteric site identification throughout the human proteome. (ii) The allosteric mutation dataset in the ASD can be used to develop methods to predict allosteric driver mutations. The real cases are the identification of the allosteric L1143F driver mutation in human protein tyrosine phosphatase receptor type K (PTPRK) and the allosteric P360A driver mutation in human phosphodiesterase 10A (PDE10A). (iii) The allosterome maps can be used for allosteric evolutionary analysis of an allosteric site (an allosteric modulator) within its protein family (known allosteric modulators). (iv) The AlloFinder-based platform that integrates allosteric site prediction (AllositePro), allosteric interaction evaluation (Alloscore), and allosteric evolutionary analysis (Allosterome) can be used to automatically screen for allosteric modulators for a given target. A real case is the identification of a STAT3 allosteric inhibitor, K116. The new features in ASD 2019 are highlighted with red color.

ALLOSTEROME MAPS

Owing to remarkable advances in structural biology and chemical biology in the past three years, an increasing number of allosteric sites have been determined by experimental crystallographic methods. These experimentally determined allosteric sites can be used to construct allosterome maps of protein families to illuminate the evolution of allosteric regulation (7). In ASD 2019, seven new human allosterome maps, including ‘Ion Channel’, ‘Protease’, ‘Phosphatase’, ‘Nuclear Receptor’, ‘Transferase’, ‘Transporter’, and ‘Hydrolase’ maps, were built to reveal the biological evolution of allosteric sites in a family; the ‘GPCR’ and ‘Kinase’ allosterome maps were also thoroughly updated by adding recently identified members (see Table 2). During the construction of the maps, all family annotations were collected from UniProt (57), GPCRDB (60), the Nuclear Receptor Resource (61), the Enzyme database (62), IUPHAR/BPS Pharmacology (63) and Wikipedia (<https://en.wikipedia.org/>), and the maps were constructed based on the multiple sequence alignment and dendrogram methods described previously (41).

The allosterome maps are organized in a grid under the ‘ALLOSTEROME’ feature in the ASD (Figure 2 and Table 2). Clicking on one map causes the details of the allosterome to be shown in a new tab, and the allosteric pro-

teins with modulators and complexes are well annotated with white circle and red flag symbols, respectively. In nearly all the maps (except the transporter map), the percentage of allosteric proteins with modulators reached >70%, with particularly high percentages being achieved for the ‘GPCR’ (95%) and ‘Ion Channel’ allosteromes (84%). In four maps, >10 allosteric proteins with allosteric complexes are marked, including 30 kinases, 16 hydrolases, 13 proteases and 12 GPCRs. For each family, these allosteric sites actually cluster into different locations throughout the sub-families. For example, 12 allosteric sites in human GPCRs show 3 different locations according to structural superimposition and 5 allosteric sites in ion channels are located in 4 different regions. The ‘ALLOSTEROME’ feature has systematically recorded allosteric protein accumulation from decades of allosterome development and offers the possibility of predicting the locations of novel sites based on known sites from the same family.

ALLOSTERIC MUTATIONS

Mutations around an allosteric site may trap a protein in either an active or inactive conformation, and this structural disruption of uncontrolled proteins can perturb downstream signaling pathways and elicit multiple physiological and pathological conditions, including cancer (7). The

Table 2. Summary and statistics of the allosterome module in ASD 2019

Allosterome category ^a	Number of human allosteric proteins		
	All	Has modulators	Has complexes
Protein kinase allosterome	113	85	30
GPCR allosterome	114	108	12
Ion channel allosterome	94	79	4
Hydrolase allosterome	66	52	16
Transferase allosterome	54	40	9
Transporter allosterome	53	34	7
Protease allosterome	45	34	13
Phosphatase allosterome	19	14	5
Nuclear receptor allosterome	17	12	4

^aIn the definition of the classifications for allosteric proteins in ASD 2019, several subtypes have been removed from the enzyme categories, including transferases (kinases excluded), hydrolases (proteases and phosphatases excluded) and transporters (ion channels excluded).

increasing number of clinical mutations discovered from next-generation sequencing (NGS) data provides us with an unprecedented opportunity to systematically examine allosteric regulation related to disease conditions and promotes the development of small datasets and novel resources in this new field (24–27). To help biologists and chemists to easily explore allosteric mutations validated from NGS samples, in ASD 2019, the high-quality ‘Allo-Mutation’ dataset was established under the ‘FEATURES’ menu. Applying our previous workflow (24,27) to the most recent data on allosteric complexes and NGS samples, 1312 somatic missense mutations located in the allosteric sites of 133 human proteins were extracted from >100 000 patient samples from TCGA (64) and COSMIC (65) across 33 cancer types. Each allosteric mutation in a human protein from one kind of TCGA cancer is defined as a mutation record. The frequency of variants at allosteric sites for the allosteric proteins in each cancer type is depicted as a heatmap in Figure 3. For instance, allosteric mutations in BRAF present the highest frequency in THCA, SKCM and COAD samples, and allosteric mutations in PIK3CA have also been identified in numerous kinds of cancers. This module could be of great help for not only understanding the allosteric mechanism but also identifying allosteric targets in precision medicine.

ALLOSTERIC TOOLS

To facilitate studies of allosteric mechanisms and drug discovery, six specified allosteric computational tools are integrated under the ‘TOOLS’ menu of the ASD. ‘AlloPharm’ was built for allosteric target prediction by reverse searching of query molecules against allosteric pharmacophore datasets generated from allosteric complexes and ligand-free receptors. This structure-based pharmacophore tool provides an ensemble of spatially distributed chemical features that are essential for specific target-ligand binding (see Supporting Information). In total, 1493 pharmacophore models were extracted from 1374 allosteric protein-modulator complex structures, and 267 pharmacophore models were generated from unique allosteric pockets in 226 proteins via a pocket-based method (66–68). These pharmacophores of allosteric sites will be valuable for the high-throughput virtual screening of allosteric modulators. ‘Allo-Pathway’ is a tool for analyzing the allosteric communi-

cation pathway through revised dynamic network analysis (see Supporting Information). The allosteric pathways predicted by this tool could support an understanding of the roles of allostery in biological processes and signal transduction networks. In addition, AlloDriver (27) for allosteric mutagenesis prediction, AlloFinder (30) for the investigation of allosteric mechanisms, AllositePro (45) for allosteric site identification, and AlloScore (54) for the evaluation of allosteric interactions are integrated together in ASD 2019. Collectively, all of the above methods are expected to expedite the discovery of allosteric drugs and research on allosteric mechanisms.

DISCUSSION AND CONCLUSION

The discovery and development of first-in-class drugs is emerging as a focal point of the pharmaceutical industry (69,70). The development of an efficient strategy for discovering first-in-class drugs is therefore an area of intensive research. Allosteric drugs that exhibit a unique mechanism of action by targeting structurally diverse allosteric sites provide an innovative and promising opportunity for treating human diseases (11,12).

However, allosteric drug discovery is fraught with obstacles, and the greatest challenge is the effectiveness of using experimental methods to identify bona fide allosteric sites for a therapeutic target. As a supplement to experimental methods, computational predictions of allosteric sites in proteins made using the ASD data can expedite allosteric drug discovery. In particular, many of the computational predictions have been validated by experimental observations. For example, based on the Allosite-based allosteric site prediction (15), we recently discovered the first-in-class cellularly active SIRT6 allosteric activators MDL-800 and MDL-801 bound to the computationally predicted allosteric site and performed further crystallographic determinations. Through the AlloFinder-based platform (30) that integrates allosteric site prediction (AllositePro, an updated version of Allosite) (45), allosteric interaction evaluation (AlloScore) (54), and allosteric evolutionary analysis (Allosterome), we also recently identified a STAT3 allosteric inhibitor, K116. These findings strongly support the notion that computational allosteric methods have played an important role in structure-based drug design (23,33). In this study, we predicted a total of 10 081 potential allosteric

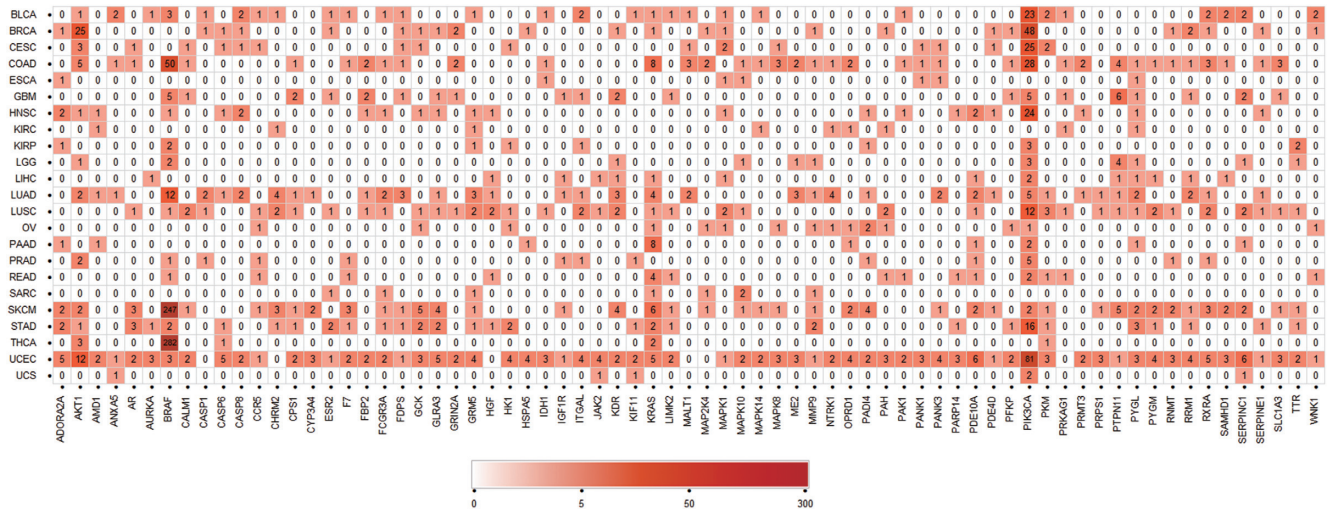


Figure 3. Heatmap displaying the frequency of mutations at the allosteric sites of human proteins in each cancer type. For each gene or each cancer type, records with a total frequency >6 are displayed in this heatmap.

sites for 4013 human allosteric proteins using AllositePro (45). This novel dataset of potential allosteric sites represents a breakthrough attributable to ASD 2019, because it is a valuable resource providing an attractive starting point for structure-based drug design at previously unexplored allosteric sites.

Allosteric mutations cause another challenge for allosteric drug discovery because of the notorious difficulty of identifying the locations of these mutation sites in proteins, as they may, for example, occur in allosteric ligand binding sites or allosteric communication pathways. The ‘Allo-Mutation’ dataset in ASD 2019, consisting of 1312 somatic missense mutations located in the allosteric sites of 133 human proteins, is useful for not only developing computational predictions of allosteric mutations but also uncovering new clinical disease targets. Such is the case for AlloDriver, which was recently developed based on this dataset (27). AlloDriver identified an L1143F driver mutation at the allosteric site of human protein tyrosine phosphatase receptor type K (PTPRK) from samples obtained from head and neck squamous cell carcinoma (HNSC) patients, which may expand the potential of PTPRK as a therapeutic target for HNSC (27). Additionally, the AlloMAPS database was built based on the theoretical prediction of allosteric effects of mutations on protein structure and function (25). Thus, this novel ‘Allo-Mutation’ dataset represents another breakthrough attributable to ASD 2019, because it is important for the identification of clinically relevant allosteric driver mutations that can expand the available therapeutic targets as well as for the investigation of allosteric communication pathways.

It has been suggested that allosteric sites are under lower evolutionary pressure than orthosteric sites. Notably, we have newly constructed human allosterome maps of seven protein families in ASD 2019, in addition to the update of two previous allosterome (GPCR and protein kinase) maps in ASD 2016. This ‘ALLOSTEROME’ dataset in ASD 2019 represents a useful improvement, which can reveal the specificity of an allosteric site in its protein family as well as trace

the origin and evolution of allosteric regulation in protein families. Moreover, this information can also aid in the development of allosteric modulators that exploit differences in allosteric sites to achieve improved selectivity.

In summary, we anticipate that the new features in ASD 2019 will become a useful tool for allosteric modulator screening, allosteric mutation and evolutionary analysis, and allosteric mechanism research.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

National Natural Science Foundation of China [91753117, 21778037, 81721004, 81603023, in part]; Shanghai Sailing Program [16YF1406500]; Shanghai Health and Family Planning Commission [20184Y0268]; Subgrant [2017YFC0908502] of the Chinese National Precise Medical Research key project [2017YFC0908500]; Innovation Program of Shanghai Municipal Education Commission [2019-01-07-00-01-E00036]; Shanghai Science and Technology Innovation [19431901600]; Shanghai Health and Family Planning System Excellent Subject Leader and Excellent Young Medical Talents Training Program [2018BR12], Subgrant [2018ZX09711002-001-006] of the National Science & Technology Major Project ‘Key New Drug Creation and Manufacturing Program’ [2018ZX09711002]; Natural Science Foundation of Shanghai Municipal Commission of Health and Family Planning [20164Y0033]; Shanghai Natural Science Foundation [16ZR1449300]. Funding for open access charge: National Natural Science Foundation of China.
Conflict of interest statement. None declared.

REFERENCES

1. Changeux, J.P. (2013) The concept of allosteric modulation: an overview. *Drug Discov. Today Technol.*, **10**, e223–e228.

2. Liu, J. and Nussinov, R. (2016) Allosteric: an overview of its history, concepts, methods, and applications. *PLoS Comput. Biol.*, **12**, e1004966.
3. Goodey, N.M. and Benkovic, S.J. (2008) Allosteric regulation and catalysis emerge via a common route. *Nat. Chem. Biol.*, **4**, 474–482.
4. Nussinov, R., Tsai, C.J. and Liu, J. (2014) Principles of allosteric interactions in cell signaling. *J. Am. Chem. Soc.*, **136**, 17692–17701.
5. Good, M., Tang, G., Singleton, J., Remenyi, A. and Lim, W.A. (2009) The Ste5 scaffold directs mating signaling by catalytically unlocking the Fus3 MAP kinase for activation. *Cell*, **136**, 1085–1097.
6. Good, M.C., Zalatan, J.G. and Lim, W.A. (2011) Scaffold proteins: hubs for controlling the flow of cellular information. *Science*, **332**, 680–686.
7. Nussinov, R. and Tsai, C.J. (2013) Allosteric in disease and in drug discovery. *Cell*, **153**, 293–305.
8. Lu, S., Ji, M., Ni, D. and Zhang, J. (2018) Discovery of hidden allosteric sites as novel targets for allosteric drug design. *Drug Discov. Today*, **23**, 359–365.
9. Nussinov, R., Tsai, C.J. and Csermely, P. (2011) Allo-network drugs: harnessing allostery in cellular networks. *Trends Pharmacol. Sci.*, **32**, 686–693.
10. Nussinov, R. and Tsai, C.J. (2012) The different ways through which specificity works in orthosteric and allosteric drugs. *Curr. Pharm. Des.*, **18**, 1311–1316.
11. Lu, S., Li, S. and Zhang, J. (2014) Harnessing allostery: a novel approach to drug discovery. *Med. Res. Rev.*, **34**, 1242–1285.
12. Wu, P., Clausen, M.H. and Nielsen, T.E. (2015) Allosteric small-molecule kinase inhibitors. *Pharmacol. Ther.*, **156**, 59–68.
13. Clark, A.C. (2016) Caspase allostery and conformational selection. *Chem Rev.*, **116**, 6666–6706.
14. Yu, Z.H. and Zhang, Z.Y. (2018) Regulatory mechanisms and novel therapeutic targeting strategies for protein tyrosine phosphatases. *Chem Rev.*, **118**, 1069–1091.
15. Huang, Z., Zhao, J., Deng, W., Chen, Y., Shang, J., Song, K., Zhang, L., Wang, C., Lu, S., Yang, X. *et al.* (2018) Identification of a cellularly active SIRT6 allosteric activator. *Nat. Chem. Biol.*, **14**, 1118–1126.
16. Lu, S. and Zhang, J. (2019) Small molecule allosteric modulators of G-protein-coupled receptors: drug-target interactions. *J. Med. Chem.*, **62**, 24–45.
17. Chai, H., Cheng, X., Zhou, B., Zhao, L., Lin, X., Huang, D., Lu, W., Lv, H., Tang, F., Zhang, Q. *et al.* (2019) Structure-based discovery of a subtype-selective inhibitor targeting a transient receptor potential vanilloid channel. *J. Med. Chem.*, **62**, 1373–1384.
18. Ni, D., Li, X., He, X., Zhang, H., Zhang, J. and Lu, S. (2019) Drugging K-Ras(G12C) through covalent inhibitors: Mission possible? *Pharmacol. Ther.*, **202**, 1–17.
19. Hauser, A.S., Attwood, M.M., Rask-Andersen, M., Schioth, H.B. and Gloriam, D.E. (2017) Trends in GPCR drug discovery: new agents, targets and indications. *Nat. Rev. Drug Discov.*, **16**, 829–842.
20. Changeux, J.P. and Christopoulos, A. (2016) Allosteric modulation as a unifying mechanism for receptor function and regulation. *Cell*, **166**, 1084–1102.
21. Hardy, J.A. and Wells, J.A. (2004) Searching for new allosteric sites in enzymes. *Curr. Opin. Struct. Biol.*, **14**, 706–715.
22. Lu, S., Huang, W. and Zhang, J. (2014) Recent computational advances in the identification of allosteric sites in proteins. *Drug Discov. Today*, **19**, 1595–1600.
23. Lu, S., Shen, Q. and Zhang, J. (2019) Allosteric methods and their applications: facilitating the discovery of allosteric drugs and the investigation of allosteric mechanisms. *Acc. Chem. Res.*, **52**, 492–500.
24. Shen, Q., Cheng, F., Song, H., Lu, W., Zhao, J., An, X., Liu, M., Chen, G., Zhao, Z. and Zhang, J. (2017) Proteome-Scale investigation of protein allosteric regulation perturbed by somatic mutations in 7,000 cancer genomes. *Am. J. Hum. Genet.*, **100**, 5–20.
25. Tan, Z.W., Tee, W.V., Guarnera, E., Booth, L. and Berezovsky, I.N. (2019) AlloMAPS: allosteric mutation analysis and polymorphism of signaling database. *Nucleic Acids Res.*, **47**, D265–D270.
26. Guarnera, E., Tan, Z.W., Zheng, Z. and Berezovsky, I.N. (2017) AlloSigMA: allosteric signaling and mutation analysis server. *Bioinformatics*, **33**, 3996–3998.
27. Song, K., Li, Q., Gao, W., Lu, S., Shen, Q., Liu, X., Wu, Y., Wang, B., Lin, H., Chen, G. *et al.* (2019) AlloDriver: a method for the identification and analysis of cancer driver targets. *Nucleic Acids Res.*, **47**, W315–W321.
28. Wodak, S.J., Paci, E., Dokholyan, N.V., Berezovsky, I.N., Horowitz, A., Li, J., Hilser, V.J., Bahar, I., Karanicolas, J., Stock, G. *et al.* (2019) Allostery in its many disguises: from theory to applications. *Structure*, **27**, 566–578.
29. Schueler-Furman, O. and Wodak, S.J. (2016) Computational approaches to investigating allostery. *Curr. Opin. Struct. Biol.*, **41**, 159–171.
30. Huang, M., Song, K., Liu, X., Lu, S., Shen, Q., Wang, R., Gao, J., Hong, Y., Li, Q., Ni, D. *et al.* (2018) AlloFinder: a strategy for allosteric modulator discovery and allosterome analyses. *Nucleic Acids Res.*, **46**, W451–W458.
31. Li, C., Deng, X., Zhang, W., Xie, X., Conrad, M., Liu, Y., Angeli, J.P.F. and Lai, L. (2019) Novel allosteric activators for ferroptosis regulator glutathione peroxidase 4. *J. Med. Chem.*, **62**, 266–275.
32. Novinec, M., Korenc, M., Cafilisch, A., Ranganathan, R., Lenarcic, B. and Baici, A. (2014) A novel allosteric mechanism in the cysteine peptidase cathepsin K discovered by computational methods. *Nat. Commun.*, **5**, 3287.
33. Lu, S., He, X., Ni, D. and Zhang, J. (2019) Allosteric modulator discovery: from serendipity to structure-based design. *J. Med. Chem.*, **62**, 6405–6421.
34. Agnello, S., Brand, M., Chellat, M.F., Gazzola, S. and Riedl, R. (2019) A structural view on medicinal chemistry strategies against drug resistance. *Angew. Chem. Int. Ed. Engl.*, **58**, 3300–3345.
35. Shah, N.H. and Kuriyan, J. (2019) Understanding molecular mechanisms in cell signaling through natural and artificial sequence variation. *Nat. Struct. Mol. Biol.*, **26**, 25–34.
36. Kolch, W. and Kiel, C. (2018) From oncogenic mutation to dynamic code. *Science*, **361**, 844–845.
37. Wylie, A.A., Schoepfer, J., Jahnke, W., Cowan-Jacob, S.W., Loo, A., Furet, P., Marzink, A.L., Pelle, X., Donovan, J., Zhu, W. *et al.* (2017) The allosteric inhibitor ABL001 enables dual targeting of BCR-ABL1. *Nature*, **543**, 733–737.
38. Gao, Y., Maria, A., Na, N., Paula, A.D.C., Gorelick, A.N., Hechtman, J.F., Carson, J., Lefkowitz, R.A., Weigelt, B., Taylor, B.S. *et al.* (2019) V211D mutation in MEK1 causes resistance to MEK inhibitors in colon cancer. *Cancer Discov.*, **9**, 1182–1191.
39. Huang, Z., Zhu, L., Cao, Y., Wu, G., Liu, X., Chen, Y., Wang, Q., Shi, T., Zhao, Y., Wang, Y. *et al.* (2011) ASD: a comprehensive database of allosteric proteins and modulators. *Nucleic Acids Res.*, **39**, D663–D669.
40. Gao, Y., Mou, L., Shen, Q., Lu, S., Li, C., Liu, X., Wang, G., Li, S., Geng, L., Liu, Y. *et al.* (2014) ASD v2.0: updated content and novel features focusing on allosteric regulation. *Nucleic Acids Res.*, **42**, D510–D516.
41. Shen, Q., Wang, G., Li, S., Liu, X., Lu, S., Chen, Z., Song, K., Yan, J., Geng, L., Huang, Z. *et al.* (2016) ASD v3.0: unraveling allosteric regulation with structural mechanisms and biological networks. *Nucleic Acids Res.*, **44**, D527–D535.
42. Chang, R.L., Andrews, K., Kim, D., Li, Z., Godzik, A. and Palsson, B.O. (2013) Structural systems biology evaluation of metabolic thermotolerance in *Escherichia coli*. *Science*, **340**, 1220–1223.
43. Huang, W., Wang, G., Shen, Q., Liu, X., Lu, S., Geng, L., Huang, Z. and Zhang, J. (2015) ASBench: benchmarking sets for allosteric discovery. *Bioinformatics*, **31**, 2598–2600.
44. Huang, W., Lu, S., Huang, Z., Liu, X., Mou, L., Luo, Y., Zhao, Y., Liu, Y., Chen, Z., Hou, T. *et al.* (2013) AlloSite: a method for predicting allosteric sites. *Bioinformatics*, **29**, 2357–2359.
45. Song, K., Liu, X., Huang, W., Lu, S., Shen, Q., Zhang, L. and Zhang, J. (2017) Improved method for the identification and validation of allosteric sites. *J. Chem. Inf. Model.*, **57**, 2358–2363.
46. Greener, J.G. and Sternberg, M.J. (2015) AlloPred: prediction of allosteric pockets on proteins using normal mode perturbation analysis. *BMC Bioinformatics*, **16**, 335.
47. Panjkovich, A. and Daura, X. (2014) PARS: a web server for the prediction of protein allosteric and regulatory sites. *Bioinformatics*, **30**, 1314–1315.
48. Xu, Y., Wang, S., Hu, Q., Gao, S., Ma, X., Zhang, W., Shen, Y., Chen, F., Lai, L. and Pei, J. (2018) CavityPlus: a web server for protein cavity detection with pharmacophore modelling, allosteric site identification and covalent ligand binding ability prediction. *Nucleic Acids Res.*, **46**, W374–W379.

49. Cimermancic,P, Weinkam,P, Rettenmaier,T.J., Bichmann,L., Keedy,D.A., Woldeyes,R.A., Schneidman-Duhovny,D., Demerdash,O.N., Mitchell,J.C., Wells,J.A. *et al.* (2016) CryptoSite: Expanding the druggable proteome by characterization and prediction of cryptic binding sites. *J. Mol. Biol.*, **428**, 709–719.
50. Suplatov,D., Kirilin,E., Arbatsky,M., Takhaveev,V. and Svedas,V. (2014) pocketZebra: a web-server for automated selection and classification of subfamily-specific binding sites by bioinformatic analysis of diverse protein families. *Nucleic Acids Res.*, **42**, W344–W349.
51. Bhagavat,R., Sankar,S., Srinivasan,N. and Chandra,N. (2018) An augmented pocketome: detection and analysis of small-molecule binding pockets in proteins of known 3D structure. *Structure*, **26**, 499–512.
52. Akbar,R. and Helms,V. (2018) ALLO: A tool to discriminate and prioritize allosteric pockets. *Chem. Biol. Drug Des.*, **91**, 845–853.
53. Greener,J.G., Filippis,I. and Sternberg,M.J.E. (2017) Predicting protein dynamics and allostery using multi-protein atomic distance constraints. *Structure*, **25**, 546–558.
54. Li,S., Shen,Q., Su,M., Liu,X., Lu,S., Chen,Z., Wang,R. and Zhang,J. (2016) Alloscore: a method for predicting allosteric ligand-protein interactions. *Bioinformatics*, **32**, 1574–1576.
55. Kaya,C., Armutlulu,A., Ekesan,S. and Haliloglu,T. (2013) MCPATH: Monte Carlo path generation approach to predict likely allosteric pathways and functional residues. *Nucleic Acids Res.*, **41**, W249–W255.
56. Yugi,K., Kubota,H., Hatano,A. and Kuroda,S. (2016) Trans-omics: how to reconstruct biochemical networks across multiple ‘Omic’ layers. *Trends Biotechnol.*, **34**, 276–290.
57. UniProt, Consortium (2019) UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.*, **47**, D506–D515.
58. Panjkovich,A. and Daura,X. (2012) Exploiting protein flexibility to predict the location of allosteric sites. *BMC Bioinformatics*, **13**, 273.
59. Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) The Protein Data Bank. *Nucleic Acids Res.*, **28**, 235–242.
60. Isberg,V., Mordalski,S., Munk,C., Rataj,K., Harpsoe,K., Hauser,A.S., Vroiling,B., Bojarski,A.J., Vriend,G. and Gloriam,D.E. (2016) GPCRdb: an information system for G protein-coupled receptors. *Nucleic Acids Res.*, **44**, D356–D364.
61. Martinez,E., Moore,D.D., Keller,E., Pearce,D., Vanden Heuvel,J.P., Robinson,V., Gottlieb,B., MacDonald,P., Simons,S. Jr, Sanchez,E. *et al.* (1998) The Nuclear Receptor Resource: a growing family. *Nucleic Acids Res.*, **26**, 239–241.
62. Bairoch,A. (2000) The ENZYME database in 2000. *Nucleic Acids Res.*, **28**, 304–305.
63. Harding,S.D., Sharman,J.L., Faccenda,E., Southan,C., Pawson,A.J., Ireland,S., Gray,A.J.G., Bruce,L., Alexander,S.P.H., Anderton,S. *et al.* (2018) The IUPHAR/BPS Guide to PHARMACOLOGY in 2018: updates and expansion to encompass the new guide to IMMUNOPHARMACOLOGY. *Nucleic Acids Res.*, **46**, D1091–D1106.
64. Tomczak,K., Czerwinska,P. and Wiznerowicz,M. (2015) The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp. Oncol. (Pozn.)*, **19**, A68–A77.
65. Tate,J.G., Bamford,S., Jubb,H.C., Sondka,Z., Beare,D.M., Bindal,N., Boutselakis,H., Cole,C.G., Creatore,C., Dawson,E. *et al.* (2019) COSMIC: the catalogue of somatic mutations in cancer. *Nucleic Acids Res.*, **47**, D941–D947.
66. Koes,D.R. and Camacho,C.J. (2012) ZINCPharmer: pharmacophore search of the ZINC database. *Nucleic Acids Res.*, **40**, W409–W414.
67. Sunseri,J. and Koes,D.R. (2016) Pharmit: interactive exploration of chemical space. *Nucleic Acids Res.*, **44**, W442–W448.
68. Wang,X., Shen,Y., Wang,S., Li,S., Zhang,W., Liu,X., Lai,L., Pei,J. and Li,H. (2017) PharmMapper 2017 update: a web server for potential drug target identification with a comprehensive target pharmacophore database. *Nucleic Acids Res.*, **45**, W356–W360.
69. Eder,J., Sedrani,R. and Wiesmann,C. (2014) The discovery of first-in-class drugs: origins and evolution. *Nat. Rev. Drug Discov.*, **13**, 577–587.
70. Lu,S., Ni,D., Wang,C., He,X., Lin,H., Wang,Z. and Zhang,J. (2019) Deactivation pathway of Ras GTPase underlies conformational substates as targets for drug design. *ACS Catal.*, **9**, 7188–7196.