



Facilitating Antiviral Drug Discovery Using Genetic and Evolutionary Knowledge

Xuan Xu⁺, Qing-Ye Zhang ^{*,†}, Xin-Yi Chu, Yuan Quan, Bo-Min Lv and Hong-Yu Zhang *

Hubei Key Laboratory of Agricultural Bioinformatics, College of Informatics, Huazhong Agricultural University, Wuhan 430070, China; xxpinky@webmail.hzau.edu.cn (X.X.); chuxy@webmail.hzau.edu.cn (X.-Y.C.); quanyuan@mail.hzau.edu.cn (Y.Q.); lbm612@webmail.hzau.edu.cn (B.-M.L.)

Correspondence: zqy@mail.hzau.edu.cn (Q.-Y.Z.); zhy630@mail.hzau.edu.cn (H.-Y.Z.);

Tel.: +86-27-8728-0877 (H.-Y.Z.)

+ These authors contributed equally.

Abstract: Over the course of human history, billions of people worldwide have been infected by various viruses. Despite rapid progress in the development of biomedical techniques, it is still a significant challenge to find promising new antiviral targets and drugs. In the past, antiviral drugs mainly targeted viral proteins when they were used as part of treatment strategies. Since the virus mutation rate is much faster than that of the host, such drugs feature drug resistance and narrow-spectrum antiviral problems. Therefore, the targeting of host molecules has gradually become an important area of research for the development of antiviral drugs. In recent years, rapid advances in high-throughput sequencing techniques have enabled numerous genetic studies (such as genome-wide association studies (GWAS), clustered regularly interspersed short palindromic repeats (CRISPR) screening, etc.) for human diseases, providing valuable genetic and evolutionary resources. Furthermore, it has been revealed that successful drug targets exhibit similar genetic and evolutionary features, which are of great value in identifying promising drug targets and discovering new drugs. Considering these developments, in this article the authors propose a hosttargeted antiviral drug discovery strategy based on knowledge of genetics and evolution. We first comprehensively summarized the genetic, subcellular location, and evolutionary features of the human genes that have been successfully used as antiviral targets. Next, the summarized features were used to screen novel druggable antiviral targets and to find potential antiviral drugs, in an attempt to promote the discovery of new antiviral drugs.

Keywords: genetics; evolution; host receptors

1. Introduction

Since the earliest yellow fever virus was discovered in 1901, the number of viruses has continued to increase. On average, three to four new species are found each year. So far, more than two hundred viruses that infect humans have been identified [1]. Human viral infection is the most unpredictable disease among infectious diseases. Currently, there are several viruses with high mortality rates in the world. For instance, Coronavirus disease (COVID-19) caused by SARS-CoV-2 has led to a global pandemic and can result in a series of respiratory diseases, such as pneumonia and lung failure [2–4]; Ebola virus causes a range of highly pathogenic symptoms, such as hemorrhagic fever, and is highly contagious, with a mortality rate of 57%–90% [5]; Severe Acute Respiratory Syndrome (SARS), has affected more than 8000 people and killed 774 people by July 2003 [6]; Epstein-Barr virus (EBV), a common human herpes DNA virus, can lead to lifelong infection in more than 90 percent of the population, and causes a variety of human malignancies [7]; Human Immunodeficiency Viruses (HIV/AIDS) causes severe defects in human cellular immune function. As of 2018, 77.3 million people worldwide have been infected with HIV, and 35.4 million people have died from AIDS-related diseases [8]. Thousands of human plagues



Citation: Xu, X.; Zhang, Q.-Y.; Chu, X.-Y.; Quan, Y.; Lv, B.-M.; Zhang, H.-Y. Facilitating Antiviral Drug Discovery Using Genetic and Evolutionary Knowledge. *Viruses* **2021**, *13*, 2117. https://doi.org/10.3390/v13112117

Academic Editor: Jamil S. Saad

Received: 26 September 2021 Accepted: 19 October 2021 Published: 20 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). are still found every year worldwide [9]. These viruses usually have the characteristics of wide spread and strong concealment. As a result, mortality from viral diseases remains high. Due to the limitation of virus resistance, effective preventive measures and treatment drugs are still lacking for many viral diseases. There is a pressing need for the development of antiviral drugs.

Antiviral drugs are usually classified according to their effects: anti-HIV drugs (such as Zindovudine, Abacavir, Nelfinavir, Delavirdine, Nevirapine, etc.) [10,11]; anticytomegalovirus (CMV) drugs (such as Ganciclovir, Valganciclovir, Cidofovir, Formivirsen, etc.) [12,13]; anti-hepatitis virus drugs (such as Telaprevir, Ribavirin, Simeprevir, Boceprevir, Sofosbuvir, etc.) [14,15]; anti-herpes virus drugs (such as Acyclovir, Valacilovir, Idoxuridine, Brivudin, etc.) [16,17]; and anti-influenza and respiratory virus drugs (such as Amantadine, Rimantadine, Osehamivir, Zanamivir, etc.) [18]. Although there are dozens of drugs available for the treatment of some important viral diseases, they only target a small number of viral pathogens [19–21]. In addition, the occasional appearance of more pathogenic strains of known or previously unknown viruses continues to raise public health concerns and reminds people of the need for effective treatments. The research statistics on antiviral drugs in the NCBI PubMed from 1945 to 2021 showed that the number of research papers dedicated to this topic has increased year by year, from only one in 1945 to a maximum of 17,198 in 2013, indicating a growing interest in the development of antiviral drugs.

Drug development has always been a very long and costly process, whose comprehensive cost estimates range from hundreds of thousands of dollars to approximately \$2.6 billion [22–24]. Among thousands of new chemical structures, only a few prove to be potential drugs [22]. It is estimated that only 10.4% of new phase I clinical development projects were finally approved by the US Food and Drug Administration (FDA) between 2003 and 2011 [25]. Some studies have shown that the efficiency of research and development (R&D) has been steadily declining, as measured by the number of new drugs that are put on the market per billion US dollars of Research and Development expenditure in the global biotechnology and pharmaceutical industries [23,26]. One main reason for failed drug development is the lack of effectiveness of drug targets [27,28]. Considering the fact that viruses lack a cell structure, their genetic material is easily perturbed by the external environment and the host cellular molecular environment. Traditional antiviral drugs mostly inhibit the enzymes that are essential for virus reproduction at different stages of virus replication, thereby blocking this replication [19,29]. However, more and more studies have shown that this pathogen-targeting strategy, though successful in many cases, is not effective enough to combat the emergence of drug resistance [30].

Over the past few decades, significant advances in high-throughput sequencing, analytical techniques, and molecular biology, have brought deeper understanding of the genetic and evolutionary mechanisms of viral infection and provided new opportunities for therapeutic interventions [31]. At present, a large number of gene loci related to susceptibility to viral diseases (such as HIV, hepatitis, dengue (DEN) etc.) have been discovered by genome-wide association studies (GWAS) [32]. These results further elucidate the genetic structure of susceptibility to infectious diseases. It is known that genetic diversity drives evolution and contributes to adaptation to new environments. Gelbar et al. used ultra-deep sequencing to examine 43 clinical samples taken from early human infections with HIV, respiratory syncytial virus pneumonia (RSV), and Cytomegalovirus (CMV). They identified the presence of multiple distinct genotypes in the HIV and CMV samples as the major drivers of increased diversity [33]. Pairo-Castineira et al. performed a GWAS comparing 2244 critically ill COVID-19 patients with heathy individuals in the UK. Subsequently, using the transcriptome-wide Mendelian randomization (MR) method to investigate the causal relationship between potentially modifiable risk factors and health outcomes, they found that the high expression of encoding tyrosine kinase 2 (TYK2) was significantly related to critical COVID-19. Therefore, the authors speculated that the TYK2-targeted drug (JAK inhibitor Baricitinib) may have the potential to combat SARS-CoV-2 [34]. Meanwhile, hosttargeted antiviral drug discovery strategies have become popular in recent years. Since the

mutation rate of the host protein is significantly lower than that of the virus, host-targeted antiviral drugs are thought to offer a higher genetic barrier to against mutations than directacting antiviral drugs [35]. Ullah et al. recently discovered that the inhibitor of human RACK1 protein shows similar efficacy to the approved anti-herpes drug acyclovir, and that RACK1 protein is a potential target for the development of broad-spectrum antiviral drugs [36]. Tyrrell et al. selected the endoplasmic reticulum a-glucosidases as a target based on the glycosylation mechanism, and iminosugars were shown to have potential efficacy in difficult-to-treat infections [37]. Up to now, this strategy has also been successful in the treatment of DEN [38] and other viral diseases [39–41]. Burmeister et al. proposed a treatment method based on ecological evolution systems. The virus microorganisms and the host's internal environment are regarded as an ecosystem, and host and microbial pathogens are part of this complex interaction system [42]. It can be seen that the strategy of targeting host factors circumvents the barrier created by the high-frequency mutation of viruses, and provides a unique opportunity for the development of more broad-spectrum antiviral drugs [42].

On the other hand, recent studies have also found that successful drug targets often have common genetic and evolutionary features, which are of great help in screening potential drug targets and discovering new drugs [43–45]. Inspired by the above information, in this study we first summarized the genetic characteristics, cellular sublocation, and evolutionary features of existing approved host-targeted antiviral drug targets. Next, based on the summarized features, potential host-aimed antiviral targets were screened, in an attempt to promote the development of new antiviral drugs.

2. Feature Analysis for Host Targets of Approved Antiviral Drugs

2.1. Genetic Features

For viruses, the primary barrier to invasion is the cell's innate immunity. This antiviral response first detects viral components through host cells, then transduces the antiviral signals, transcription and translation of antiviral effectors, and finally establishes the body's antiviral state [46]. Combined with the host-targeted strategy, much progress has been made in identifying specific human genetic variations that help to increase sensitivity or resistance to viral diseases [47]. Currently, the most popular genetic testing methods focus on the discovery of a disease's potentially susceptible genes through GWAS [48–50] and the detection of genetic polymorphisms related to the phenotype of specific viral diseases through whole-exome sequencing (WES) [51,52]. For example, Akcay et al. recently reported that HLA-DP/DQ mutations contribute to HBV persistence and resistance to vaccines. Besides, susceptible alleles in the HLA-DP/DQ locus are more common in Asian populations than in Whites, which may explain why chronic HBV infection rates are higher in Asian populations [50]. Additionally, Hashemi et al. declared that host-targeted antiviral drugs may be subject to host gene polymorphisms that alter their ability to inhibit the function of target proteins [53]. Indeed, in the clinical setting, suboptimal response to Alisporivir has been reported in 10–15% of patients, which may be due to the potential influence of single nucleotide polymorphisms (SNPs) [54]. Another study tested whole exon mutations at the individual level and showed that the progression of HIV-1 disease is related to microbial infection and innate immune activation, and the role of microbes in the damaged intestinal mucosa is thought to be related to disease progression. Therefore, microbial translocation and the activation of innate pattern recognition receptors (PPR) may affect disease progression and non-AIDS complications [55]. Gene mutations in EC 004 (TAB2), LTNP 005 (PIK3C2B), LTNP 007 (FGD6), and LTNP 008 (PRKDC) are also particularly significant for the slow-progressing phenotype of HIV [55]. These genes, as well as the immunological and cellular biological mechanisms involved in them, could be the subject of further research on the pathogenesis and disease progression of HIV, and may also become potential new drug targets. In addition, based on a genome-wide loss-of-function CRISPR screen, Daniloski et al. obtained highly ranked genes in host. The dysfunction of these genes significantly reduces the infection rate of SARS-CoV-2. Moreover, nine of these genes are druggable genes corresponding to 26 small molecule inhibitors. The experimental results showed that seven of the inhibitors can reduce the SARS-CoV-2 viral load by more than 100-fold. Interestingly, among these seven effective antiviral agents, four target the same gene, *PI3K catalytic subunit type 3* (*PI3KC3*) [56]. Therefore, the authors contended that CRISPR is an effective genetic method for screening antiviral targets [56–58]. Furthermore, since GWAS can only identify genes that are significantly associated with the disease, there is an increasing number of studies incorporating MR to identify the causal relationship between gene mutations and disease. By combining GWAS and MR methods, some host genes, such as *IFNAR2*, *TYK2*, *JAK1*, *JAK2*, *TNF*, *IL6*, *TMPRSS2*, and *ILIR1* have been identified as potential therapeutic targets for COVID-19 [34,59]. These findings provide important clues for a better understanding of the mechanism of COVID-19. Taken together, genetics-identified disease genes are considered a promising source of drug targets to help increase the success rate of antivirus drug development [28].

In a previous study [28], we combined the information from eight disease gene databases (including Clinvar [60], Online Mendelian Inheritance in Man (OMIM) [61], The Human Gene Mutation Database (HGMD) [62], Orphanet [63], GWASdb [64], INtegrated TaRget gEne PredItion (INTREPID) [65], Genetic Association Database (GAD) [66], and DisGeNET [67]) and used the degree of clinical support to award these genes different quality scores (termed druggability score). Through a comprehensive analysis, it was confirmed that these disease genetics-derived druggability scores can be used to measure the association between genes and diseases, which offers important values for screening potential targets and predicting drug activity [28]. The higher the score, the stronger the association between the gene and the disease, and the easier it is to develop the gene as a drug target.

To analyze the genetic features of antiviral drug targets, firstly, 36 known host targets of approved antiviral drugs were downloaded from DrugBank [68] (Table S1). Based on the druggability scores derived from the SCG-Drug database (Quan et al. 2019), 32 targets had relevant score information (accounting for 88.9% of the total 36 targets). Next, the median score of these targets with 34 viral diseases was calculated to measure the strength of the association between genes and pathogenic viruses. The results showed that 13 genes received genetic scores equal to or greater than three. Besides, 27 targets were documented in association with more than one viral disease, as shown in the last column of Table 1.

Gene Symbol	Uniprot ID	Protein Name	SCG Virus Diseases Median Score	Subcellular Location	Origin	Associated Virus Diseases Name (Only List the Top 3) *
C1QA	P02745	Complement C1q subcomponent subunit A	1	Secreted	Vertebrata	Herpes Simplex
C1QB	P02746	Complement C1q subcomponent subunit B	1	Secreted	Vertebrata	Herpes Simplex; Carcinoma, Merkel Cell
C1QC	P02747	Complement C1q subcomponent subunit C	1	Secreted	Vertebrata	Herpes Simplex; Adenoviridae Infections
C1R	P00736	Complement C1r subcomponent, EC 3.4.21.41	3	Secreted	Vertebrata	AIDS Dementia Complex
CCR5	P51681	C-C chemokine receptor type 5, C-C CKR-5, CC-CKR-5, CCR-5, CCR5	3	Cell membrane	Eumetazoa	Hepatitis, Viral, Human; West Nile Fever; Hepatitis C
CD4	P01730	T-cell surface glycoprotein CD4	3	Cell membrane	Vertebrata	Measles; Picornaviridae Infections; HIV Infections
CHRNA3	P32297	Neuronal acetylcholine receptor subunit alpha-3	3	Cell membrane	Eumetazoa	AIDS Dementia Complex; Carcinoma, Merkel Cell; Influenza, Human
CHRNA4	P43681	Neuronal acetylcholine receptor subunit alpha-4	1	Cell membrane	Eumetazoa	AIDS Dementia Complex; Influenza, Human; Picornaviridae Infections
CHRNA7	P36544	Neuronal acetylcholine receptor subunit alpha-7	3	Cell membrane	Eumetazoa	AIDS Dementia Complex
CXCR4	P61073	C-X-C chemokine receptor type 4	3	Cell membrane	Eumetazoa	Herpes Simplex; Papillomavirus Infections; Epstein-Barr Virus Infections
DRD2	P14416	D(2) dopamine receptor	1	Cell membrane	Eumetazoa	AIDS Dementia Complex; Hemorrhagic Fevers, Viral; Hepatitis
ENPP1	P22413	Ectonucleotide pyrphosphatase/phosphodiesterase family member 1	3	Cell membrane	Eukaryota	West Nile Fever; Hepatitis; AIDS Dementia Complex
FCGR1A	P12314	High affinity immunoglobulin gamma Fc receptor I	1	Cell membrane	Eumetazoa	Leukoencephalopathy, Progressive Multifocal; Hemorrhagic Fever, Ebola; Hemorrhagic Fevers, Viral
FCGR2A	P12318	Low affinity immunoglobulin gamma Fc region receptor II-a	2	Cell membrane	Eumetazoa	Leukoencephalopathy, Progressive Multifocal; West Nile Fever; Picornaviridae Infections
FCGR2B	P31994	Low affinity immunoglobulin gamma Fc region receptor II-b	1	Cell membrane	Eumetazoa	Picornaviridae Infections; Hemorrhagic Fevers, Viral; Hepatitis C
FCGR2C	P31995	Low affinity immunoglobulin gamma Fc region receptor II-c	1	Cell membrane	Eumetazoa	HIV Infections; Herpes Simplex; Hepatitis C
FCGR3A	P08637	Low affinity immunoglobulin gamma Fc region receptor III-A	3	Cell membrane	Eumetazoa	Herpesviridae İnfections; Hepatitis, Viral, Human; Measles
FCGR3B	O75015	Low affinity immunoglobulin gamma Fc region receptor III-B	1	Cell membrane	Eumetazoa	Influenza, Human; Picornaviridae Infections; West Nile Fever

Table 1. Basic information of approved antiviral host targets.

Gene Symbol	Uniprot ID	Protein Name	SCG Virus Diseases Median Score	Subcellular Location	Origin	Associated Virus Diseases Name (Only List the Top 3) *
GRIN3A	Q8TCU5	Glutamate receptor ionotropic, NMDA 3A, GluN3A	3	Cell membrane	Vertebrata	AIDS Dementia Complex
HLA-B	P18465	HLA class I histocompatibility antigen, B alpha chain	3	Cell membrane	Vertebrata	HIV Infections; Hepatitis; Picornaviridae Infections
IFNAR1	P17181	Interferon alpha/beta receptor 1	1	Cell membrane	Vertebrata	Leukoencephalopathy, Progressive Multifocal; Hemorrhagic Fevers, Viral; West Nile Fever
IFNAR2	P48551	Interferon alpha/beta receptor 2	3	Cell membrane	Vertebrata	Hepatitis; Picornaviridae Infections; Leukoencephalopathy, Progressive Multifocal
IMPDH1	P20839	Inosine-5'-monophosphate dehydrogenase 1		Cytoplasm	Cellular_ organisms	
IMPDH2	P12268	Inosine-5'-monophosphate dehydrogenase 2	1	Nucleus	Cellular_ organisms	Sarcoma, Kaposi
NEU1	Q99519	N-acetyl-alpha-neuraminidase 1	2	Cell membrane	Euk + Bac	Carcinoma, Merkel Cell; Leukoencephalopathy, Progressive Multifocal; Sarcoma, Kaposi
NEU2	Q9Y3R4	N-acetyl-alpha-neuraminidase 2		Cytoplasm	Euk + Bac	I
NR1I2	O75469	Nuclear receptor subfamily 1 group I member 2	1	Nucleus	Eumetazoa	Hepatitis; Measles; Paramyxoviridae Infections
NT5C2	P49902	Cytosolic purine 5'-nucleotidase, EC 3.1.3.5	1	Cytoplasm	Eukaryota	AIDS Dementia Complex; Cytomegalovirus Infections; Sarcoma, Kanosi
PNP	P00491	Purine nucleoside phosphorylase, PNP, EC 2.4.2.1	1	Cytoplasm	Euk + Bac	Cytomegalovirus Infections; AIDS Dementia Complex; Epstein-Barr Virus Infections
SLCO1B1	Q9Y6L6	Solute carrier organic anion transporter family member 1B1	3	Cell membrane	Mammalia	Fatigue Syndrome, Chronic; Picornaviridae Infections; HIV Infections
SLCO2B1	O94956	Solute carrier organic anion transporter family member 2B1		Cell membrane	Eumetazoa	
TERT	O14746	Telomerase reverse transcriptase, EC 2.7.7.49	1	Nucleolus	Eukaryota	Carcinoma, Merkel Cell; Leukoencephalopathy, Progressive Multifocal; Influenza, Human
TOP2A	P11388	DNA topoisomerase 2-alpha, EC 5.6.2.2	1	Nucleoplasm	Eukaryota	Leukoencephalopathy, Progressive Multifocal; Carcinoma, Merkel Cell; Hepatitis, Viral, Human

Table 1. Cont.

			Table 1. Cont.			
Gene Symbol	Uniprot ID	Protein Name	SCG Virus Diseases Median Score	Subcellular Location	Origin	Associated Virus Diseases Name (Only List the Top 3) *
TUBA4A	P68366	Tubulin alpha-4A chain	4	Cytoskeleton	Eukaryota	AIDS Dementia Complex; Leukoencephalopathy, Progressive Multifocal
TUBB	P07437	Tubulin beta chain		Cytoplasm, cytoskeleton	Eukaryota	
TYMS	P04818	Thymidylate synthase, TS, TSase, EC 2.1.1.45	1	Nucleus	Euk + Bac	Sarcoma, Kaposi; Influenza, Human; Hepatitis

Genes located on the cell membrane are shown in bold. * Complete candidate targets and diseases association information are listed in Table S1. Disease names are listed in descending order of SCG scores.

2.2. Subcellular Location Features

The coding capacity of the viral DNA genome is limited, so it is necessary to use host cell factors to promote the replication of its genome and to produce progeny viruses [69]. A critical initial step in the life cycle of an infectious virus is the identification of and interaction with the host proteins. Therefore, to initiate this cycle, the viral genome needs to be transported from viral particles to the cytoplasm through different viral attachment proteins on their shells [70]. The particles are first attached to the specific proteins on the surface of the host cell. These specific proteins are called receptors. The presence of receptors on the cell surface is a major determinant of viral host selectivity and plays a key regulatory role in host range and viral pathogenesis [71]. Until now, antiviral drug strategies targeting cell surface receptors have made progress in the treatment of different viral infectious diseases [72]. One of the main areas of focus for research is anti-HIV drugs. HIV entry into target cells is a complex process, consisting of multiple biological responses, which is triggered by the interaction between the HIV envelope protein (Env) glycoprotein gp120 and the host cell receptor CD4. This interaction causes a conformational change in gp120, exposing a co-receptor binding site [73]. Subsequently, after binding to the CCR5 or CXCR4 core receptor of the primary HIV isolate, Env undergoes further conformational changes, gp41-mediated membrane fusion, and virion core entry into the cytoplasm [74]. Therefore, the polymorphism of the CCR5 gene can affect the spread of HIV or the development of the disease. Targeting the process of HIV invasion into cells is considered to be an effective method to inhibit HIV replication [11]. The anti-HIV efficacy of small molecule CCR5 antagonists, such as Cenicriviroc [75] and Vicriviroc (SCH-D) [76], has proven this theory.

Given the above findings, a subcellular localization feature needs to be considered in drug development and protein function development. By analyzing the mechanism of virus invasion, it can be speculated that cell membrane receptors are more likely to become potential targets for antiviral drugs. To test this conjecture, we also began with the 36 known antiviral human targets. The Uniprot database was used to retrieve their subcellular localization information [77]. To improve the accuracy and credibility of the results, the manually annotated and reviewed Swiss-Prot database was chosen for the analysis [77]. A total of 20,233 pieces of human genetic information was collected from the Swiss-Prot database, of which 14,574 were labeled with subcellular localization information. Subcellular location information was available on the database for 32 out of the 36 known antiviral human targets, and 21 were located on the cell membrane (65.6%). It was observed that the antiviral targets were indeed significantly enriched on the cell membrane surface (p < 0.001, hypergeometric test). This characteristic can be used as a key indicator for subsequent antiviral host target screening. It is worth noting that among the 20 scored cell membrane receptors, 11 received a median SCG score of three, accounting for 84.6% (11/13) of the total number of targets with a score of three or more. Therefore, the cell membrane receptors demonstrated a high genetic association with viral diseases.

2.3. Evolutionary Features

Viruses initially interact with the host through their capsid or envelope proteins. There is an argument that the primitive capsid protein-encoding genes are recruited from cellular hosts, which implies that the coevolution between virus and host is deeply rooted in the origins of viruses [78]. Furthermore, some studies have shown that the evolution of the virus and parasite host range play an important role in the emergence of infectious diseases [78,79]. Coevolution between viruses and hosts has also led to larger genetic differences between replication populations, which are often related to the range of hosts that co-evolutionary viruses can infect [80,81]. For example, health disasters caused by influenza viruses have been common in human history. One of the causes of influenza virus epidemics is its ability to continuously perform antigen transformation through genetic diversity. At the same time, the host's immune defense mechanism is constantly evolving. This process might be compared with the evolutionary race between predators and prey [78,82]. By studying the mechanism of host resistance to infection, SNPs of the

interferon-inducible transmembrane proteins (IFITMs), especially *ifitm3* gene, were found to be able to increase or decrease susceptibility to influenza infection [83,84], showing the antiviral significance of the host protein's evolution. Synergistic benefits can also be produced by integrating the virus-host evolutionary mechanism and the interaction between pathogens and hosts to improve therapeutic effects [42]. So far, two effective treatment strategies, evolutionary model in which viruses in the host control or eliminate pathogens [42,85,86]. Furthermore, Phillips et al. found that the ability of host chaperone proteins to influence virus evolution is a key factor in determining viral protein mutation fitness, which provides new insights for the design of antiviral treatments [87].

Tomislav et al. showed that the pathogenicity of disease genes is closely related to their evolutionary origin [88]. The results of research by Wang et al. indicated that successful targets share some common evolutionary characteristics, and suggested that evolutionary information can help to identify the drug targets with the greatest potential for therapeutic development [43,45]. An in-depth understanding of the antiviral activity and action mechanism of host target genes will help to develop broad-spectrum antiviral drugs that enhance their activity or have similar effects [89].

There are more than 30 popular algorithms through which to infer the origins of gene families [90–92]. It is a challenge to obtain accurate gene ages because the age of the gene families depends on the accuracy of the genetic algorithm. To solve this problem, Liebeskind et. al. provided a relatively consistent genetic age data set and mapped the species with orthologous genomes to a reference tree from SwissTree to infer the age of each gene [93]. Based on the rules, human genes are divided into the following eight classes by origin: (1) cellular organism, (2) the common ancestors of eukaryotes and archaea (Euk_Archaea), (3) horizontal gene transfer of bacteria (Euk + Bac), (4) Eukaryota, (5) Opisthokonta, (6) Eumetazoa, (7) vertebrata, (8) mammals. It is believed that this set of data includes minor errors.

Based on the above knowledge, the evolutionary origin of 36 human antiviral targets were analyzed. The UniprotKB data corresponding to 36 targets were obtained by querying the UniprotKB database. The evolution information statistics results indicated that the 36 UniProtKB were mainly distributed in genes originating in class 6 (Eumetazoa) (14/36, p = 0.04, hypergeometric test) (shown in Figure 1a). As for the 21 cell membrane receptors, 13 of them originated in the Eumetazoan period, and the trend of enrichment was more obvious (p = 0.00034, hypergeometric test) (shown as in Figure 1b). The results of the gene functional annotation using DAVID showed that these receptors are significantly related to immune response ($p = 1.4 \times 10^{-8}$).



Evolutionary biology feature analysis

Figure 1. The statistical histogram of evolutionary features of approved antiviral receptors and cell membrane receptors. The results of the evolutionary information statistics showed that the 36 approved targets were distributed mainly in genes belonging to class 6 (Eumetazoa) (14/36, p = 0.04, hypergeometric assay) (**a**). As for the 21 cell membrane receptors, 13 of them were from the Eumetazoan period, and the tendency for enrichment was more pronounced (p = 0.00034, hypergeometric test) (**b**).

A previous evolutionary study of 658 successful targets by Quan et al. indicated that these genes are significantly enriched in cellular organism, euk + bac, and Eumetazoa [45]. Our results were partially consistent with their study, suggesting that human genes originating in the Eumetazoan period have greater potential to become antiviral targets.

3. Rational Screening of Potential Antiviral Host Targets

It is often inadequate to screen targets based solely on a single feature summarized by omics data because they usually depend on large-scale experiments and cannot comprehensively explain the causal relationship between human diseases [45,94]. Besides, biotechnology continues to identify an increasing number of disease-related genes. Not all of them are potential drug targets. Identifying new and effective targets remains a major priority for modern drug discovery [95]. Based on the analysis of the genetic features, subcellular localization, and evolutionary features of approved antiviral host targets, it is summarized that an effective antiviral target generally has the following three characteristics: (1) a median of genetic score(s) associated with viral disease(s) equal to or greater than three; (2) subcellular localization on the cell membrane; (3) origins in the Eumetazoan period. By using these criteria, we screened out some potential antiviral host targets (Figure 2).



Figure 2. Antiviral host targets screening by genetic, subcellular location, and evolutionary biology features (different shapes and colors of symbols represent various genes related to viral diseases). Based on the analysis of the genetic features, subcellular localization, and evolutionary features of the approved antiviral host targets, it is summarized that an effective antiviral target generally exhibits the following three characteristics: (1) a median of genetic scores associated with viral disease (s) equal to or greater than three; (2) subcellular localization on the cell membrane; (3) origins in the Eumetazoan period.

3.1. Target Screening by Summarized Features

As mentioned above, the SCG-Drug database integrates information from eight common genetic disease databases and provides different druggability scores for these genes. It provides us with a huge data set for screening potential antiviral drug targets. We started with 914,190 gene-disease associations in the SCG-Drug database and obtained 19,232 genes by removing repetitions. These genes were used for screening according to the three features described above. First, through genetic screening, 7214 genes were found to be associated with 34 viral diseases. For each gene, the median score associated with viral disease(s) was taken as its genetic score. Next, these genes were divided into two categories based on whether there was relevant information on the three well-known drug databases (Therapeutic Target Database (TTD) [96], DGIdb [97], and DrugBank [68]). Of the 2277 genes with relevant drug information on the databases, we found that 687 are located on the cell membrane and 248 originated in the Eumetazoan period. For the 4937 genes with no drug information, 744 were found to be located on the cell membrane, and 239 originated in the Eumetazoan period. To find candidate genes that are strongly associated with viral disease(s), genes with genetic scores equal to or greater than three were selected. Furthermore, through searching the studies available on NCBI PubMed database, we selected genes with virus-related reports as candidate targets. Finally, 35 candidate targets that met all of the above requirements were obtained (Table 3). For the 13 candidate targets with drug information, their corresponding drugs may be able to treat viral diseases through drug repositioning. As for the remaining 22 candidate targets without drug information, novel antiviral drugs may be derived from them through structure-based drug design.

					C			
Gene Symbol	Uniprot ID	Protein Name	Drug Information	Recognized Target	SCG Virus Disease(S) Median Score	Associated Virus Disease(s) Name (Only List the Top 3) *	Evolutionary Age	Pubmed Number ^a
ADRA2A	P08913	Alpha-2A adrenergic receptor	known	Yes	3	AIDS Dementia Complex; Hepatitis	Eumetazoa	13
ADRB2	P07550	Beta-2 adrenergic receptor	known	Yes	3	Leukoencephalopathy, Progressive Multifocal; Paramyxoviridae Infections; Picornaviridae Infections	Eumetazoa	52
ADRB3	P13945	Beta-3 adrenergic receptor	known	Yes	3	Hepatitis; AIDS Dementia Complex; Picornaviridae Infections	Eumetazoa	1
AGER	Q15109	Advanced glycosylation end product-specific receptor	unknown	No	3	Leukoencephalopathy, Progressive Multifocal; Hepatitis; AIDS Dementia Complex	Eumetazoa	27
AGTR1	P30556	Type-1 angiotensin II receptor	known	Yes	3	Leukoencephalopathy, Progressive Multifocal; Paramyxoviridae Infections; Hapatitis	Eumetazoa	3
ASIC1	P78348	Acid-sensing ion channel 1, ASIC1	known	Yes	3	Leukoencephalopathy, Progressive Multifocal	Eumetazoa	3
BIN1	O00499	Myc box-dependent-interacting protein 1	unknown	No	6	Fatigue Syndrome, Chronic; AIDS Dementia Complex	Eumetazoa	14
BSG	P35613	Basigin	unknown	No	3	Hepatitis; West Nile Fever; Epstein–Barr Virus Infections	Eumetazoa	19
CALHM1	Q8IU99	Calcium homeostasis modulator protein 1	unknown	No	9	AIDS Dementia Complex	Eumetazoa	1
CD209	Q9NNX6	CD209 antigen	unknown	Yes	3	West Nile Fever; Hepatitis; Hemorrhagic Fever, Ebola	Eumetazoa	65
CD86	P42081	T-lymphocyte activation antigen CD86	known	Yes	3	Respiratory Syncytial Virus Infections; Picornaviridae Infections: Hepatitis D	Eumetazoa	758
CRHR2	Q13324	Corticotropin-releasing factor receptor 2	unknown	Yes	3	Hepatitis; Fatigue Syndrome, Chronic; West Nile Fever	Eumetazoa	4
DRD4	P21917	D(4) dopamine	known	Yes	3	AIDS Dementia Complex	Eumetazoa	5
DRP2	Q13474	Dystrophin-related protein 2	unknown	No	3	AIDS Dementia Complex	Eumetazoa	1
FAS	P25445	Tumor necrosis factor receptor superfamily member 6	unknown	Yes	3	Dementia Complex; Carcinoma, Merkel Cell	Eumetazoa	1588
GHSR	Q92847	Growth hormone secretagogue	known	Yes	3	Carcinoma, Merkel Cell	Eumetazoa	7

receptor type 1

Table 2. Basic information of candidate targets.

Gene Symbol	Uniprot ID	Protein Name	Drug Information	Recognized Target	SCG Virus Disease(S) Median Score	Associated Virus Disease(s) Name (Only List the Top 3) *	Evolutionary Age	Pubmed Number ^a
GPC5	P78333	Glypican-5	unknown	No	3	Leukoencephalopathy, Progressive Multifocal;	Eumetazoa	1
GPR65	Q8IYL9	Psychosine receptor	unknown	No	6	Leukoencephalopathy, Progressive Multifocal	Eumetazoa	1
GRM4	Q14833	Metabotropic glutamate receptor 4	known	Yes	3	Leukoencephalopathy, Progressive Multifocal; Hepatitis; Picornaviridae Infections	Eumetazoa	1
GRPR	P30550	Gastrin-releasing peptide receptor	unknown	Yes	3	AIDS Dementia Complex	Eumetazoa	1
IL1R1	P14778	Interleukin-1 receptor type 1	known	Yes	3	Leukoencephalopathy, Progressive Multifocal; Picornaviridae Infections; HIV Infections	Eumetazoa	17
IL1RL1	Q01638	Interleukin-1 receptor-like 1	unknown	No	3	Influenza, Human; Hepatitis D; West Nile Fever	Eumetazoa	31
KCNQ1	P51787	channel subfamily KQT member 1	known	Yes	3	Picornaviridae Infections; AIDS Dementia Complex; Measles	Eumetazoa	5
LGR5	075473	Leucine-rich repeat-containing G-protein coupled receptor 5	unknown	Yes	3	HIV Infections; Influenza, Human; Paramyxoviridae Infections	Eumetazoa	25
MAG	P20916	Myelin-associated glycoprotein	unknown	No	3	Leukoencephalopathy, Progressive Multifocal	Eumetazoa	100
NPSR1	Q6W5P4	Neuropeptide S receptor	known	Yes	3	Hepatitis; Respiratory Syncytial Virus Infections; Influenza, Human	Eumetazoa	1
PARD6A	Q9NPB6	Partitioning defective 6 homolog alpha	unknown	No	3	AIDS Dementia Complex	Eumetazoa	2
PERP	Q96FX8	p53 apoptosis effector related to PMP-22	unknown	No	3	Carcinoma, Merkel Cell; Leukoencephalopathy, Progressive Multifocal	Eumetazoa	5
RGS7	P49802	Regulator of G-protein signaling 7	unknown	No	3	Leukoencephalopathy, Progressive Multifocal	Eumetazoa	1
RHOU	Q7L0Q8	Rho-related GTP-binding protein RhoU	unknown	No	3	Carcinoma, Merkel Cell	Eumetazoa	4
SGCA	Q16586	Alpha-sarcoglycan, Alpha-SG	unknown	No	14	Fatigue Syndrome, Chronic	Eumetazoa	7

Table 2. Cont.

Table 3. Cont.									
Gene Symbol	Uniprot ID	Protein Name	Drug Information	Recognized Target	SCG Virus Disease(S) Median Score	Associated Virus Disease(s) Name (Only List the Top 3) *	Evolutionary Age	Pubmed Number ^a	
SHB	Q15464	SH2 domain-containing adapter protein B	unknown	No	3	AIDS Dementia Complex	Eumetazoa	19	
SYNE2	Q8WXH0	Nesprin-2	unknown	No	7	Fatigue Syndrome, Chronic;	Eumetazoa	4	
TGFBR1	P36897	TGF-beta receptor type-1	known	Yes	3	Hepatitis; Hepatitis B; Carcinoma, Merkel Cell	Eumetazoa	29	
TNFSF10	P50591	Tumor necrosis factor ligand superfamily member 10	unknown	No	3	Hepatitis D; HIV Infections; Leukoencephalopathy, Progressive Multifocal	Eumetazoa	224	

^a Refers to the number of NCBI Pubmed (https://www.ncbi.nlm.nih.gov/, accessed on 20 October 2021) studies related to this gene and virus. * Complete candidate targets and diseases association information are listed in Table S2. Disease names are listed in descending order of SCG scores.

Table 3. Cont.

3.2. Target Screening by Protein Structures

Although we can use various technologies to obtain a large number of potential therapeutic targets, it remains a great challenge to bring these targets into clinical trials. The reported statistics results showed that about 10% of genes in the entire human genome are involved in the development of disease, that is, about 3000 potential targets are suitable for drug treatment [98]. Therefore, predicting whether a protein features a pocket that can bind drug-like molecules with high affinity (druggability [99]) is of great significance in the target screening stage of drug discovery [100]. In order to improve the efficiency of target discovery, many computational methods have been developed to predict target druggability [101–104]. PockDrug [101] is a reliable model for predicting the druggability of pockets, which is based on a linear discriminant analysis of 52 molecular descriptors. It selects the most stable and effective model according to different prediction objects. In addition, PockDrug maintained the best combination of molecular descriptors that affect the pharmacological properties of the pocket. With the best combination, PockDrug can achieve an average accuracy of about 87% [101].

For those targets about which there is no drug information, we need to perform druggability analysis based on protein structures to determine whether it has the potential to be a small molecule drug target. The PDB structure of candidate targets were queried in Swiss-Prot. However, not all of these candidate genes feature known protein structures. To improve the accuracy of the results, we only performed the druggability prediction on the potential targets with reported crystal structures. Therefore, 12 candidate genes with known protein crystal structures were selected for druggability analysis. All the predicted information by PockDrug is summarized in Table 4 and Table S3. Pockets with druggability scores of 0.5 or more were considered druggable.

Gene Symbol	PDB ID	Protein Name	Resolution (Å)	Ligand(s) ^a	Number of Pockets	Number of Druggable Pockets (Score \geq 0.5)	Best Druggable Pocket Score
AGER	303U	Advanced glycosylation end product-specific receptor Myc	1.50	MLR, SO4	24	13	1.00
BIN1	2FIC	box-dependent-interacting protein 1	1.99	XE	11	7	0.91
BSG	3I84	Basigin	2.00	CL	2	0	
CD209	2XR6	CD209 antigen	1.35	07B, MAN, AE9, CA, CL	2	2	0.78
CRHR2	3N93-AB	Corticotropin-releasing factor receptor 2	2.50	MAL, GOL	41	18	1.00
FAS	3TJE-F	Tumor necrosis factor receptor superfamily member 6	1.93	CD, EDO, CL	4	1	0.84
IL1RL1	4KC3-B	Interleukin-1 receptor-like 1	3.27	NAG, MSE	11	7	0.98
LGR5	4UFR-AC	Leucine-rich repeat-containing G-protein coupled receptor 5	2.20	NAG, CL	34	21	0.99
PARD6A	1WMH-B	Partitioning defective 6 homolog alpha	1.50		4	3	0.99
RGS7	2A72	Regulator of G-protein signaling 7	2.00	CL	7	2	0.60
RHOU	2Q3H	Rho-related GTP-binding protein RhoU	1.73	GDP, MG	6	1	0.93
TNFSF10	1DG6-A	Tumor necrosis factor ligand superfamily member 10	1.30	ZN, CL	2	1	0.65

Table 4. Druggability information of candidate targets.

^a The ligand information is derived from the RCSB PDB (https://www.rcsb.org/, accessed on 20 October 2021) [105].

We used the "predict druggability with proteins" function. Next, we uploaded the corresponding PDB file, selected the Fpocket estimation method (not guided by ligand information), and set the Ligand proximity threshold to the default, 5.5 Å. The complete and detailed prediction results are shown in Table S3.

4. Antiviral Drug Discovery Based on Screened Host Targets

4.1. Drug Repositioning

Drug repositioning has always been an important way to reduce drug development costs and speed up the research process [106]. At present, the number of approved antiviral drugs targeting host genes is significantly higher than that of drugs directly targeting viral genes [107]. Therefore, it is feasible to adopt a host-centered strategy and to integrate genetic and evolutionary knowledge to increase the success rate of drug repositioning. Among the 35 candidate targets listed in Table 3, 13 targets can be found with relevant approved drug information in the DrugBank database. It is believed that the existing approved drugs targeting these genes may be used in antiviral therapies through drug repositioning.

By searching the ClinicalTrials databases, the efficacy of some repositioned drugs was supported by the clinical evidence. For instance, *IL1R1* (Interleukin-1 receptor type 1) encodes a cytokine receptor belonging to the interleukin-1 receptor family, which participates in the regulation of innate immune and inflammatory processes [108]. Anakinra is a human interleukin-1 receptor antagonist approved for the treatment of adult rheumatoid arthritis (RA) and neonatal multiple inflammatory diseases (NOMID). It was predicted by disease genetics that drugs targeting *IL1R1* would be used for Paramyxoviridae infections, such as respiratory syncytial virus infections, influenza, and so on (Table 3, Table S2). According to the records of the ClinicalTrials database, Anakinra is currently being studied for the treatment of COVID-19. Related ClinicalTrials.gov identifiers are NCT04362943, NCT04364009, NCT04330638, NCT02735707, NCT04362111, NCT04339712, NCT04366232, NCT04341584, etc. Some of these studies are already in phase 3 or 4 (NCT02735707, NCT04330638, NCT04364009, NCT04362111).

The gene *AGTR1* (Type-1 angiotensin II receptor) is one of the important effect factors in mediating the volume and blood pressure of the cardiovascular system [109]. The genetic data suggest that drugs targeting *AGTR1* can be used in the treatment of viral diseases, such as Paramyxoviridae infections, AIDS dementia complex, etc. (Table 3, Table S2). Losartan is an angiotensin II receptor blocker used to treat high blood pressure and diabetic nephropathy and to reduce the risk of stroke. Losartan is indeed being investigated for the treatment of respiratory syndrome / coronavirus infection (NCT04312009, NCT04311177, NCT04340557) and HIV (NCT01852942, NCT01529749), which partially coincides with our predictions.

Likewise, Formoterol, targeting *ADRB2* (Beta-2 adrenergic receptor), *etc.*, is an inhaled beta-2 agonist that was approved by the FDA in 2001 for the treatment of chronic obstructive pulmonary disease (COPD) and asthma [110]. It was inferred through genetic data that it can be used in the treatment of paramyxoviridae infections, respiratory syncytial virus infections, etc. (Table 3, Table S2). Indeed, there is a phase 3 study that uses it for COVID-19 treatment (NCT04331054).

4.2. De Novo Drug Discovery

It can be seen from the results of the druggability prediction (Table 4, Table S3) that candidate targets including *FAS*, *CD209*, *LGR5*, *CRHR2*, *AGER*, *BIN1*, *BSG*, *IL1RL1*, *PARD6A*, *RGS7*, *RHOU*, and *TNFSF10* have one or more druggable pockets. These twelve potential antiviral host target genes are divided into two types, according to whether they are known targets or not. One is the recognized target type, including *CD209*, *FAS*, *LGR5*, and *CRHR2* genes (however, without drug information); the other is the non-recognized type, including *AGER*, *BIN1*, *BSG*, *IL1RL1*, *PARD6A*, *RGS7*, *RHOU*, and *TNFSF10* genes. The twelve potential antiviral host targets have no approved drug information, therefore they may be used for de novo drug discovery.

The *CD209* gene encodes the pathogen-recognition receptor, *CD209* antigen, which is expressed on the surface of immature dendritic cells and is involved in the initiation of the primary immune response [111]. Many reports have shown that the *CD209* antigen could act as an attachment receptor for Ebolavirus [112], HIV-1 and HIV-2 [112], Dengue

virus [113], and others. The encoding protein's crystal structure has been released (PDB id: 2XR6), and the predicted best druggable pocket is shown in Figure 3a. Based on the identified cavity, more novel drug screening work is worth carrying out.

The *LGR5* gene encodes the Leucine-rich repeat-containing G-protein coupled receptor 5 for R-spondins that potentiates the canonical Wnt signaling pathway [114]. The *LGR5* gene is a proven target and has been used as a marker of adult tissue stem cells in the intestine, stomach, hair follicle, and mammary epithelium [115]. The receptor protein's crystal structure was released and deposited in the Protein Data Bank (PDB id: 4ufr), and the predicted best druggable pocket was the cyan surface, as shown in Figure 3b. The pocket could be used as an important cavity for drug screening.

The *CRHR2* gene encodes the corticotropin-releasing factor receptor 2. The protein's crystal structure was released and deposited in the Protein Data Bank (PDB id: 3n93). Its molecular functions include corticotropin-releasing factor receptor activity, corticotropin-releasing hormone receptor activity, and peptide hormone binding [116]. The predicted best druggable pocket was the cyan surface, as shown in Figure 3c.

The *FAS* gene belongs to the tumor necrosis factor receptor superfamily, and encodes the receptor for *TNFSF6/FASLG* [117]. The molecular functions of *FAS* include calmodulin binding, identical protein binding, kinase binding, signaling receptor activity, and transmembrane signaling receptor activity [118,119]. The *FAS* gene is a proven target and is involved in autoimmune lymphoproliferative syndrome 1A disease. The encoding protein's crystal structure has been released (PDB id: 3tje), and the predicted best druggable pocket is shown in Figure 3d. Follow-up drug screening can be carried out based on this identified cavity.

The *AGER* gene encodes the advanced glycosylation end-product-specific receptor that mediates interactions between advanced glycosylation end products (AGE) and acts as a mediator of both acute and chronic vascular inflammation [120]. *AGER* is a member of the immunoglobulin superfamily of cell surface molecules, and *AGE/AGER* interaction has been linked to the regulation of the production/expression of TNF-alpha, oxidative stress, cancer, and endothelial dysfunction in type 2 diabetes [121]. The encoding protein's crystal structure has been released (PDB id: 3o3u), and the predicted best druggable pocket is shown in Figure 3e. Based on the identified cavity, more structure-based virtual screening could be performed in the future.

The *BIN1* gene encodes the Myc box-dependent-interacting protein 1 that is a membrane deforming protein. It was found to be a tumor repressor due to its interaction with MYC oncoproteins [122]. The *BIN1* gene is also involved in DNA repair, cell cycle progression, cytoskeleton regulation, apoptosis, and regulating cellular *BACE1* levels [123]. The protein of the *BIN1* gene's crystal structure was released and deposited in the Protein Data Bank (PDB id 2fic). Based on the crystal structure, the best druggable pocket was predicted and appeared in the cyan surface (Figure 3f).

The *IL1RL1* gene belongs to the interleukin-1 receptor family of cytokines and encodes the interleukin-1 receptor-like 1 protein, which is the receptor for interleukin-33 (IL-33) and whose signaling requires the association of the coreceptor *IL1RAP* [124]. The protein's alternative name is the ST2 protein. The signaling of *IL-33* is generated by its ligandbinding primary receptor, ST2, and the subsequent recruitment of accessory receptor IL-1RAP, which results in the juxtaposition of intracellular toll/interleukin-1 receptor domains of both receptors, which are necessary and sufficient for the activation of the NF- κ B and MAPK pathways in the target cells [125]. The ST2 protein's crystal structure has been released (PDB id: 4kc3), and the predicted best druggable pocket is shown in Figure 3g. This receptor was also described as a negative regulator of Toll-like receptor-IL-1 receptor signaling, and it is worthy of further study as a potential antiviral target.



Figure 3. The best druggable pocket (shown in cyan surface mode) of each target predicted by PockDrug. A total of 11 candidate genes with known protein crystal structures and with one or more druggable pockets are presented, including *CD209* (**a**), *LGR5* (**b**), *CRHR2* (**c**), *FAS* (**d**), *AGER* (**e**), *BIN1* (**f**), *IL1RL1* (**g**), *PARD6A* (**h**), *RGS7* (**i**), *RHOU* (**j**) and *TNFSF10* (**k**). All the predicted information by PockDrug is summarized in Table 4 and Table S3. Pockets with druggability scores of 0.5 or more were considered druggable.

The *PARD6A* gene encodes the partitioning defective 6 homolog alpha protein. Its short name is PAR-6, PAR6-alpha, or PAR-6A; the protein's alternative name is PAR6C, or Tax interaction protein 40 (TIP-40). PAR-6 protein is involved in asymmetrical cell division, cell polarization, and cell transformation processes [126]. The crystal structure of PAR-6 protein has been released and deposited in the database (PDB id: 1wmh), and the predicted best druggable pocket is shown in Figure 3h.

The *RGS7* gene encodes the regulator of G-protein signaling 7, which belongs to the regulator of the G-protein signaling superfamily. The combination of RGS7 protein and Gbeta5 constitutes a crucial regulator of G protein-coupled receptor signaling in the visual and nervous systems [127]. *RGS7* also plays an important role in synaptic vesicle exocytosis through its interaction with snapin [128]. The RGS7 protein's crystal structure (PDB id:2a72) and predicted best druggable pocket is shown in Figure 3i. RGS7 protein also modulates the activity of potassium channels [127]. As a potential antiviral target, it offers significant research potential.

The *RHOU* gene encodes the Rho-related GTP-binding protein, RhoU. The protein's alternative name is CDC42-like GTPase 1, GTP-binding protein-like 1, Wnt-1 responsive Cdc42 homolog 1 (Wrch-1), or Ryu GTPase. Active Wrch-1 can stimulate quiescent cells to reenter the cell cycle. Wrch-1 could regulate the actin cytoskeleton, cell morphology, cell proliferation, and migration [129]. The Wrch-1 protein's crystal structure has been determined and deposited in the database (PDB id: 2q3h). The predicted best druggable pocket is shown in Figure 3j. Antiviral drug screening could be performed based on this potential antiviral host target.

The *TNFSF10* gene encodes the protein that belongs to tumor necrosis factor ligand superfamily member 10. The protein's crystal structure was released and deposited in the Protein Data Bank (PDB id: 1dg6). The molecular functions include cytokine activity, identical protein binding, signaling receptor binding, tumor necrosis factor receptor binding, and others [130,131]. The predicted best druggable pocket is the cyan surface shown in Figure 3k. The pocket could be used as a valid cavity for drug screening.

5. Discussion and Conclusions

In recent years, new antiviral drugs have continued to appear, but the incidence of viral infections has continued to increase. Genetic and evolutionary mechanisms underlie the outbreak and epidemic of viruses. In particular, the continuous outbreak of highmortality viruses, such as COVID-19, HIV, Ebola virus, DEN, and influenza virus, has created an urgent need for research into and development of antiviral drugs. In general, the challenge for antiviral drug discovery arises from the mutation and drug-resistance of viruses. For instance, most antibodies can be evaded by single mutations of the influenza virus, which leads to simultaneous outbreaks of multiple influenza virus subtypes [132]. Traditional antiviral drugs are prone to developing drug resistance after long-term use due to this characteristic of the virus. Host-targeting antivirals not only provide a higher genetic barrier to drug resistance, but also exhibit broad-spectrum antiviral activities and are likely to be effective against newly emerging viruses [11]. Currently, there are multiple avenues under investigation as therapeutic strategies for host-targeted antivirals, such as certain innate or adaptive immune host activities that respond to viral infections (i.e., TLR agonists [133], CCR5 antagonists [75,134,135], etc.) and the regulation of inflammatory pathways (i.e., TNF- α -mediated antiviral activity [136] etc.). Meanwhile, the appropriate combinations of host-targeted antivirals may improve efficacy, expand antiviral activity, and reduce the likelihood of drug resistance. For instance, McHutchison et al. reported that in HCV-infected patients who failed to respond to treatment with peginterferon alfa and Ribavirin previously, re-treatment with Telaprevir combined with peginterferon alfa and Ribavirin was proven to be more effective than re-treatment with either drug alone [137,138].

Although host-targeting antiviral drug discovery strategies have become a major focus of research, very few human host proteins were explored as antiviral targets. In addition,

there are challenges in developing host-directed therapies. For example, emerging evidence suggests that viral resistance against host antivirals does occur [136]. Meanwhile, host-targeted therapies may produce some expected or unexpected side effects, since they may target host cell functions that are critical for cell survival [139]. Therefore, a more in-depth analysis of possible cytotoxicity or the potential to worsen the infection is needed when treating with host-targeted antivirals.

Genetic and evolutionary knowledge provides new insights into the pathogenesis of diseases by identifying specific genes or pathways associated with diseases, and thus provides opportunities to discover new drug targets. By summarizing the properties of 36 approved antiviral targets in genetics, cell biology and evolutionary biology, we found that the targets feature certain characteristics: (1) median genetic score(s) associated with viral disease(s) equal to or greater than three; (2) subcellular localization on the cell membrane; (3) origins in the Eumetazoan period. Based on these three principles, we used the gene set from the SCG-Drug database to screen new antiviral targets and obtained 35 host targets. By analyzing these candidate targets, we predicted some potential repositioned drugs and identified some new targets that may be used for de novo drug discovery.

Through the continuous development of biotechnology, significant advances have been made in high-throughput sequencing, omics analysis techniques, and molecular biology. Consequently, an increasing number of genetic, subcellular location, and evolutionary information about various genes will be accumulated. Especially in the context of the COVID-19 pandemic, the rapid screening of drug targets based on the above information will facilitate the discovery of druggable antiviral host targets, so as to promote the development of antiviral drugs.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/v13112117/s1. Table S1: 38 human targets for approved antiviral drugs downloaded from Drugbank. Table S2: Supplementary information for Table 3. Table S3: Pockdrug druggability prediction results.

Author Contributions: Conceptualization, X.X., Q.-Y.Z. and H.-Y.Z.; methodology, X.X.; data curation, X.X. and Y.Q.; writing—original draft preparation, X.X. and Q.-Y.Z.; writing—review and editing, X.X., Q.-Y.Z., X.-Y.C., B.-M.L. and H.-Y.Z.; visualization, X.X.; supervision, X.X.; project administration, Q.-Y.Z. and H.-Y.Z.; funding acquisition, H.-Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Special Projects for Technological Innovation in Hubei (Grant 2018ABA107) and the National Key R&D Project (Grant 2017YFD0200506).

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: DrugBank (https://go.drugbank.com/, accessed on 20 October 2021) and the SCG-Drug database (Quan et al. 2019).

Conflicts of Interest: The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript, apart from those disclosed.

References

- 1. Woolhouse, M.; Scott, F.; Hudson, Z.; Howey, R.; Chase-Topping, M. Human Viruses: Discovery and Emergence. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 2012, 367, 2864–2871. [CrossRef]
- Chan, J.F.-W.; Yuan, S.; Kok, K.-H.; To, K.K.-W.; Chu, H.; Yang, J.; Xing, F.; Liu, J.; Yip, C.C.-Y.; Poon, R.W.-S.; et al. A Familial Cluster of Pneumonia Associated with the 2019 Novel Coronavirus Indicating Person-to-Person Transmission: A Study of a Family Cluster. *Lancet* 2020, 395, 514–523. [CrossRef]
- 3. Huang, C.; Wang, Y.; Li, X.; Ren, L.; Zhao, J.; Hu, Y.; Zhang, L.; Fan, G.; Xu, J.; Gu, X.; et al. Clinical Features of Patients Infected with 2019 Novel Coronavirus in Wuhan, China. *Lancet* 2020, 395, 497–506. [CrossRef]
- 4. Li, Q.; Guan, X.; Wu, P.; Wang, X.; Zhou, L.; Tong, Y.; Ren, R.; Leung, K.S.M.; Lau, E.H.Y.; Wong, J.Y.; et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *N. Engl. J. Med.* **2020**, *382*, 1199–1207. [CrossRef]
- 5. Hasan, S.; Ahmad, S.A.; Masood, R.; Saeed, S. Ebola Virus: A Global Public Health Menace: A Narrative Review. *J. Fam. Med. Prim. Care* **2019**, *8*, 2189–2201. [CrossRef]

- 6. Marshall, A.H.; Rachlis, A.; Chen, J. Severe Acute Respiratory Syndrome: Responses of the Healthcare System to a Global Epidemic. *Curr. Opin. Otolaryngol. Head Neck Surg.* 2005, *13*, 161–164. [CrossRef] [PubMed]
- 7. Farrell, P.J. Epstein–Barr Virus and Cancer. Annu. Rev. Pathol. Mech. Dis. 2019, 14, 29–53. [CrossRef] [PubMed]
- 8. Gao, D.; Zou, Z.; Dong, B.; Zhang, W.; Chen, T.; Cui, W.; Ma, Y. Secular Trends in HIV/AIDS Mortality in China from 1990 to 2016: Gender Disparities. *PLoS ONE* **2019**, *14*, e0219689. [CrossRef] [PubMed]
- 9. Sun, W.; Singh, A.K. Plague Vaccine: Recent Progress and Prospects. npj Vaccines 2019, 4, 1–9. [CrossRef] [PubMed]
- 10. Maartens, G.; Celum, C.; Lewin, S.R. HIV Infection: Epidemiology, Pathogenesis, Treatment, and Prevention. *Lancet* 2014, 384, 258–271. [CrossRef]
- Ji, X.; Li, Z. Medicinal Chemistry Strategies toward Host Targeting Antiviral Agents. Med. Res. Rev. 2020, 40, 1519–1557. [CrossRef]
- 12. Pillet, S.; Pozzetto, B.; Roblin, X. Cytomegalovirus and Ulcerative Colitis: Place of Antiviral Therapy. *World J. Gastroenterol.* **2016**, 22, 2030–2045. [CrossRef]
- 13. Mareri, A.; Lasorella, S.; Iapadre, G.; Maresca, M.; Tambucci, R.; Nigro, G. Anti-Viral Therapy for Congenital Cytomegalovirus Infection: Pharmacokinetics, Efficacy and Side Effects. *J. Matern. Fetal. Neonatal. Med.* **2016**, *29*, 1657–1664. [CrossRef]
- 14. Jackson, W.E.; Everson, G.T. Sofosbuvir and Velpatasvir for the Treatment of Hepatitis C. *Expert Rev. Gastroenterol. Hepatol.* 2017, 11, 501–505. [CrossRef]
- Schaefer, E.A.K.; Chung, R.T. Anti-Hepatitis C Virus Drugs in Development. *Gastroenterology* 2012, 142, 1340–1350. [CrossRef] [PubMed]
- 16. De Clercq, E. Selective Anti-Herpesvirus Agents. Antivir. Chem. Chemother. 2013, 23, 93–101. [CrossRef] [PubMed]
- 17. Shiraki, K. Antiviral Drugs Against Alphaherpesvirus. Adv. Exp. Med. Biol. 2018, 1045, 103–122. [CrossRef]
- Amarelle, L.; Lecuona, E.; Sznajder, J.I. Anti-Influenza Treatment: Drugs Currently Used and Under Development. Arch. Bronconeumol. 2017, 53, 19–26. [CrossRef] [PubMed]
- 19. De Clercq, E.; Li, G. Approved Antiviral Drugs over the Past 50 Years. Clin. Microbiol. Rev. 2016, 29, 695–747. [CrossRef]
- 20. Menéndez-Arias, L.; Richman, D.D. Editorial Overview: Antivirals and Resistance: Advances and Challenges Ahead. *Curr. Opin. Virol.* 2014, *8*, iv–vii. [CrossRef]
- Mohammadi Pour, P.; Fakhri, S.; Asgary, S.; Farzaei, M.H.; Echeverría, J. The Signaling Pathways, and Therapeutic Targets of Antiviral Agents: Focusing on the Antiviral Approaches and Clinical Perspectives of Anthocyanins in the Management of Viral Diseases. Front. Pharmacol. 2019, 10, 1207. [CrossRef]
- 22. Brodniewicz, T.; Grynkiewicz, G. Preclinical Drug Development. Acta Pol. Pharm. 2010, 67, 578–585. [PubMed]
- 23. Scannell, J.W.; Blanckley, A.; Boldon, H.; Warrington, B. Diagnosing the Decline in Pharmaceutical R&D Efficiency. *Nat. Rev. Drug. Discov.* 2012, *11*, 191–200. [CrossRef] [PubMed]
- 24. Tufts Study Pegs Drug Development, Approval Cost at \$2.6B. GEN—Genetic Engineering Biotechnology News, 18 November 2014.
- Hay, M.; Thomas, D.W.; Craighead, J.L.; Economides, C.; Rosenthal, J. Clinical Development Success Rates for Investigational Drugs. *Nat. Biotechnol.* 2014, 32, 40–51. [CrossRef] [PubMed]
- 26. Munos, B. Lessons from 60 Years of Pharmaceutical Innovation. Nat. Rev. Drug. Discov. 2009, 8, 959–968. [CrossRef]
- 27. Vandamme, D.; Minke, B.A.; Fitzmaurice, W.; Kholodenko, B.N.; Kolch, W. Systems Biology-Embedded Target Validation: Improving Efficacy in Drug Discovery. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **2014**, *6*, 1–11. [CrossRef]
- Quan, Y.; Luo, Z.-H.; Yang, Q.-Y.; Li, J.; Zhu, Q.; Liu, Y.-M.; Cui, Z.-J.; Qin, X.; Lv, B.-M.; Xu, Y.-H.; et al. Systems Chemical Genetics-Based Drug Discovery: Prioritizing Agents Targeting Multiple/Reliable Disease-Associated Genes as Drug Candidates. *Front. Genet.* 2019, 10, 474. [CrossRef]
- 29. Nikonov, O.S.; Chernykh, E.S.; Garber, M.B.; Nikonova, E.Y. Enteroviruses: Classification, Diseases They Cause, and Approaches to Development of Antiviral Drugs. *Biochem. Mosc.* 2017, *82*, 1615–1631. [CrossRef]
- Prussia, A.; Thepchatri, P.; Snyder, J.P.; Plemper, R.K. Systematic Approaches towards the Development of Host-Directed Antiviral Therapeutics. *Int. J. Mol. Sci.* 2011, 12, 4027–4052. [CrossRef]
- 31. Münk, C.; Sommer, A.F.R.; König, R. Systems-Biology Approaches to Discover Anti-Viral Effectors of the Human Innate Immune Response. *Viruses* **2011**, *3*, 1112–1130. [CrossRef]
- 32. Chapman, S.J.; Hill, A.V.S. Human Genetic Susceptibility to Infectious Disease. Nat. Rev. Genet. 2012, 13, 175–188. [CrossRef]
- 33. Gelbart, M.; Harari, S.; Ben-Ari, Y.; Kustin, T.; Wolf, D.; Mandelboim, M.; Mor, O.; Pennings, P.S.; Stern, A. Drivers of Within-Host Genetic Diversity in Acute Infections of Viruses. *PLoS Pathog.* **2020**, *16*, e1009029. [CrossRef] [PubMed]
- 34. Pairo-Castineira, E.; Clohisey, S.; Klaric, L.; Bretherick, A.D.; Rawlik, K.; Pasko, D.; Walker, S.; Parkinson, N.; Fourman, M.H.; Russell, C.D.; et al. Genetic Mechanisms of Critical Illness in COVID-19. *Nature* **2021**, *591*, 92–98. [CrossRef] [PubMed]
- Prasad, M.; Ranjan, K.; Brar, B.; Shah, I.; Lalmbe, U.; Manimegalai, J.; Vashisht, B.; Gaury, M.; Kumar, P.; Khurana, S.K.; et al. Virus-Host Interactions: New Insights and Advances in Drug Development Against Viral Pathogens. *Curr. Drug Metab.* 2017, 18, 942–970. [CrossRef] [PubMed]
- 36. Ullah, H.; Hou, W.; Dakshanamurthy, S.; Tang, Q. Host Targeted Antiviral (HTA): Functional Inhibitor Compounds of Scaffold Protein RACK1 Inhibit Herpes Simplex Virus Proliferation. *Oncotarget* **2019**, *10*, 3209–3226. [CrossRef] [PubMed]
- Tyrrell, B.E.; Sayce, A.C.; Warfield, K.L.; Miller, J.L.; Zitzmann, N. Iminosugars: Promising Therapeutics for Influenza Infection. *Crit. Rev. Microbiol.* 2017, 43, 521–545. [CrossRef] [PubMed]

- 38. Plummer, E.; Buck, M.D.; Sanchez, M.; Greenbaum, J.A.; Turner, J.; Grewal, R.; Klose, B.; Sampath, A.; Warfield, K.L.; Peters, B.; et al. Dengue Virus Evolution under a Host-Targeted Antiviral. *J. Virol.* **2015**, *89*, 5592–5601. [CrossRef] [PubMed]
- Villalón-Letelier, F.; Brooks, A.G.; Saunders, P.M.; Londrigan, S.L.; Reading, P.C. Host Cell Restriction Factors That Limit Influenza A Infection. Viruses 2017, 9, 376. [CrossRef]
- Glennon, E.K.K.; Dankwa, S.; Smith, J.D.; Kaushansky, A. Opportunities for Host-Targeted Therapies for Malaria. *Trends Parasitol.* 2018, 34, 843–860. [CrossRef]
- 41. Varikuti, S.; Jha, B.K.; Volpedo, G.; Ryan, N.M.; Halsey, G.; Hamza, O.M.; McGwire, B.S.; Satoskar, A.R. Host-Directed Drug Therapies for Neglected Tropical Diseases Caused by Protozoan Parasites. *Front. Microbiol.* **2018**, *9*, 2655. [CrossRef]
- 42. Burmeister, A.R.; Hansen, E.; Cunningham, J.J.; Rego, E.H.; Turner, P.E.; Weitz, J.S.; Hochberg, M.E. Fighting Microbial Pathogens by Integrating Host Ecosystem Interactions and Evolution. *Bioessays* **2021**, *43*, e2000272. [CrossRef]
- 43. Wang, Z.-Y.; Fu, L.-Y.; Zhang, H.-Y. Can Medical Genetics and Evolutionary Biology Inspire Drug Target Identification? *Trends Mol. Med.* **2012**, *18*, 69–71. [CrossRef]
- 44. Kamb, A.; Harper, S.; Stefansson, K. Human Genetics as a Foundation for Innovative Drug Development. *Nat. Biotechnol.* 2013, 31, 975–978. [CrossRef]
- 45. Quan, Y.; Wang, Z.Y.; Chu, X.Y.; Zhang, H.Y. Evolutionary and Genetic Features of Drug Targets. *Med. Res. Rev.* 2018, 38, 1536–1549. [CrossRef] [PubMed]
- 46. Tripathi, S.; García-Sastre, A. Antiviral Innate Immunity through the Lens of Systems Biology. *Virus Res.* 2016, 218, 10–17. [CrossRef] [PubMed]
- 47. Kenney, A.D.; Dowdle, J.A.; Bozzacco, L.; McMichael, T.M.; St. Gelais, C.; Panfil, A.R.; Sun, Y.; Schlesinger, L.S.; Anderson, M.Z.; Green, P.L.; et al. Human Genetic Determinants of Viral Diseases. *Annu. Rev. Genet.* **2017**, *51*, 241–263. [CrossRef] [PubMed]
- 48. Klebanov, N. Genetic Predisposition to Infectious Disease. Cureus 2018, 10, e3210. [CrossRef]
- 49. Yudin, N.S.; Barkhash, A.V.; Maksimov, V.N.; Ignatieva, E.V.; Romaschenko, A.G. Human Genetic Predisposition to Diseases Caused by Viruses from Flaviviridae Family. *Mol. Biol.* **2018**, *52*, 190–209. [CrossRef]
- 50. Akcay, I.M.; Katrinli, S.; Ozdil, K.; Doganay, G.D.; Doganay, L. Host Genetic Factors Affecting Hepatitis B Infection Outcomes: Insights from Genome-Wide Association Studies. *World J. Gastroenterol.* **2018**, *24*, 3347–3360. [CrossRef]
- 51. Verma, R.; Sharma, P.C. Next Generation Sequencing-Based Emerging Trends in Molecular Biology of Gastric Cancer. *Am. J. Cancer Res.* **2018**, *8*, 207–225.
- 52. Ma, B.B.Y.; Hui, E.P.; Chan, A.T.C. Investigational Drugs for Nasopharyngeal Carcinoma. *Expert Opin. Investig. Drugs* 2017, 26, 677–685. [CrossRef]
- 53. Hashemi, S.M.A.; Thijssen, M.; Hosseini, S.Y.; Tabarraei, A.; Pourkarim, M.R.; Sarvari, J. Human Gene Polymorphisms and Their Possible Impact on the Clinical Outcome of SARS-CoV-2 Infection. *Arch. Virol.* **2021**, *166*, 1–20. [CrossRef]
- 54. Guedj, J.; Yu, J.; Levi, M.; Li, B.; Kern, S.; Naoumov, N.V.; Perelson, A.S. Modeling Viral Kinetics and Treatment Outcome during Alisporivir Interferon-Free Treatment in Hepatitis C Virus Genotype 2 and 3 Patients. *Hepatology* **2014**, *59*, 1706–1714. [CrossRef]
- 55. Nissen, S.K.; Christiansen, M.; Helleberg, M.; Kjær, K.; Jørgensen, S.E.; Gerstoft, J.; Katzenstein, T.L.; Benfield, T.; Kronborg, G.; Larsen, C.S.; et al. Whole Exome Sequencing of HIV-1 Long-Term Non-Progressors Identifies Rare Variants in Genes Encoding Innate Immune Sensors and Signaling Molecules. *Sci. Rep.* 2018, *8*, 15253. [CrossRef] [PubMed]
- 56. Daniloski, Z.; Jordan, T.X.; Wessels, H.-H.; Hoagland, D.A.; Kasela, S.; Legut, M.; Maniatis, S.; Mimitou, E.P.; Lu, L.; Geller, E.; et al. Identification of Required Host Factors for SARS-CoV-2 Infection in Human Cells. *Cell* **2021**, *184*, 92–105. [CrossRef] [PubMed]
- 57. Brest, P.; Mograbi, B.; Hofman, P.; Milano, G. Using Genetics To Dissect SARS-CoV-2 Infection. *Trends Genet*. **2021**, *37*, 203–204. [CrossRef] [PubMed]
- De Wilde, A.H.; Snijder, E.J.; Kikkert, M.; van Hemert, M.J. Host Factors in Coronavirus Replication. *Curr. Top. Microbiol. Immunol.* 2018, 419, 1–42. [CrossRef] [PubMed]
- 59. Liu, D.; Yang, J.; Feng, B.; Lu, W.; Zhao, C.; Li, L. Mendelian Randomization Analysis Identified Genes Pleiotropically Associated with the Risk and Prognosis of COVID-19. *J. Infect.* **2021**, *82*, 126–132. [CrossRef] [PubMed]
- 60. Landrum, M.J.; Lee, J.M.; Riley, G.R.; Jang, W.; Rubinstein, W.S.; Church, D.M.; Maglott, D.R. ClinVar: Public Archive of Relationships among Sequence Variation and Human Phenotype. *Nucleic Acids Res.* **2014**, *42*, D980–D985. [CrossRef]
- 61. Amberger, J.S.; Bocchini, C.A.; Schiettecatte, F.; Scott, A.F.; Hamosh, A. OMIM.Org: Online Mendelian Inheritance in Man (OMIM®), an Online Catalog of Human Genes and Genetic Disorders. *Nucleic Acids Res.* **2015**, *43*, D789–D798. [CrossRef]
- Stenson, P.D.; Mort, M.; Ball, E.V.; Evans, K.; Hayden, M.; Heywood, S.; Hussain, M.; Phillips, A.D.; Cooper, D.N. The Human Gene Mutation Database: Towards a Comprehensive Repository of Inherited Mutation Data for Medical Research, Genetic Diagnosis and next-Generation Sequencing Studies. *Hum. Genet.* 2017, 136, 665–677. [CrossRef] [PubMed]
- 63. Pavan, S.; Rommel, K.; Mateo Marquina, M.E.; Höhn, S.; Lanneau, V.; Rath, A. Clinical Practice Guidelines for Rare Diseases: The Orphanet Database. *PLoS ONE* 2017, *12*, e0170365. [CrossRef]
- Li, M.J.; Liu, Z.; Wang, P.; Wong, M.P.; Nelson, M.R.; Kocher, J.-P.A.; Yeager, M.; Sham, P.C.; Chanock, S.J.; Xia, Z.; et al. GWASdb v2: An Update Database for Human Genetic Variants Identified by Genome-Wide Association Studies. *Nucleic Acids Res.* 2016, 44, D869–D876. [CrossRef]
- 65. Chen, J.; Tian, W. Explaining the Disease Phenotype of Intergenic SNP through Predicted Long Range Regulation. *Nucleic Acids Res.* **2016**, *44*, 8641–8654. [CrossRef] [PubMed]
- 66. Becker, K.G.; Barnes, K.C.; Bright, T.J.; Wang, S.A. The Genetic Association Database. Nat. Genet. 2004, 36, 431–432. [CrossRef]

- Piñero, J.; Queralt-Rosinach, N.; Bravo, À.; Deu-Pons, J.; Bauer-Mehren, A.; Baron, M.; Sanz, F.; Furlong, L.I. DisGeNET: A Discovery Platform for the Dynamical Exploration of Human Diseases and Their Genes. *Database* 2015, 2015, bav028. [CrossRef] [PubMed]
- 68. Wishart, D.S.; Feunang, Y.D.; Guo, A.C.; Lo, E.J.; Marcu, A.; Grant, J.R.; Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z.; et al. DrugBank 5.0: A Major Update to the DrugBank Database for 2018. *Nucleic Acids Res.* **2018**, *46*, D1074–D1082. [CrossRef] [PubMed]
- 69. Weitzman, M.D.; Fradet-Turcotte, A. Virus DNA Replication and the Host DNA Damage Response. *Annu. Rev. Virol.* 2018, *5*, 141–164. [CrossRef]
- 70. Boulant, S.; Stanifer, M.; Lozach, P.-Y. Dynamics of Virus-Receptor Interactions in Virus Binding, Signaling, and Endocytosis. *Viruses* 2015, 7, 2794–2815. [CrossRef] [PubMed]
- 71. Maginnis, M.S. Virus-Receptor Interactions: The Key to Cellular Invasion. J. Mol. Biol. 2018, 430, 2590-2611. [CrossRef]
- 72. Rajendran, L.; Knölker, H.-J.; Simons, K. Subcellular Targeting Strategies for Drug Design and Delivery. *Nat. Rev. Drug. Discov.* 2010, *9*, 29–42. [CrossRef]
- 73. Clapham, P.R.; McKnight, Á. HIV-1 Receptors and Cell Tropism. Br. Med. Bull. 2001, 58, 43–59. [CrossRef] [PubMed]
- 74. Sundaravaradan, V.; Das, S.R.; Ramakrishnan, R.; Sehgal, S.; Gopalan, S.; Ahmad, N.; Jameel, S. Role of HIV-1 Subtype C Envelope V3 to V5 Regions in Viral Entry, Coreceptor Utilization and Replication Efficiency in Primary T-Lymphocytes and Monocyte-Derived Macrophages. *Virol. J.* 2007, *4*, 126. [CrossRef] [PubMed]
- 75. Kim, M.B.; Giesler, K.E.; Tahirovic, Y.A.; Truax, V.M.; Liotta, D.C.; Wilson, L.J. CCR5 Receptor Antagonists in Preclinical to Phase II Clinical Development for Treatment of HIV. *Expert Opin. Investig. Drugs* **2016**, *25*, 1377–1392. [CrossRef]
- Lenz, J.C.C.; Rockstroh, J.K. Vicriviroc, a New CC-Chemokine Receptor 5 Inhibitor for Treatment of HIV: Properties, Promises and Challenges. *Expert Opin. Drug Metab. Toxicol.* 2010, 6, 1139–1150. [CrossRef] [PubMed]
- 77. UniProt: A Worldwide Hub of Protein Knowledge. Nucleic Acids Res. 2019, 47, D506–D515. [CrossRef]
- Gorter, F.A.; Hall, A.R.; Buckling, A.; Scanlan, P.D. Parasite Host Range and the Evolution of Host Resistance. J. Evol. Biol. 2015, 28, 1119–1130. [CrossRef]
- Dallas, T.; Huang, S.; Nunn, C.; Park, A.W.; Drake, J.M. Estimating Parasite Host Range. Proc. Biol. Sci. 2017, 284, 20171250. [CrossRef]
- 80. Paterson, S.; Vogwill, T.; Buckling, A.; Benmayor, R.; Spiers, A.J.; Thomson, N.R.; Quail, M.; Smith, F.; Walker, D.; Libberton, B.; et al. Antagonistic Coevolution Accelerates Molecular Evolution. *Nature* **2010**, *464*, 275–278. [CrossRef]
- 81. Koonin, E.V. Viruses and Mobile Elements as Drivers of Evolutionary Transitions. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 2016, 371, 20150442. [CrossRef]
- 82. Voskarides, K.; Christaki, E.; Nikolopoulos, G.K. Influenza Virus—Host Co-Evolution. A Predator-Prey Relationship? *Front. Immunol.* **2018**, *9*, 2017. [CrossRef] [PubMed]
- Smith, S.E.; Busse, D.C.; Binter, S.; Weston, S.; Diaz Soria, C.; Laksono, B.M.; Clare, S.; van Nieuwkoop, S.; van den Hoogen, B.G.; Clement, M.; et al. Interferon-Induced Transmembrane Protein 1 Restricts Replication of Viruses That Enter Cells via the Plasma Membrane. J. Virol. 2019, 93, e02003-18. [CrossRef]
- Zhai, Y.; Franco, L.M.; Atmar, R.L.; Quarles, J.M.; Arden, N.; Bucasas, K.L.; Wells, J.M.; Niño, D.; Wang, X.; Zapata, G.E.; et al. Host Transcriptional Response to Influenza and Other Acute Respiratory Viral Infections—A Prospective Cohort Study. *PLoS Pathog.* 2015, *11*, e1004869. [CrossRef] [PubMed]
- 85. Bull, J.J.; Levin, B.R.; Molineux, I.J. Promises and Pitfalls of In Vivo Evolution to Improve Phage Therapy. *Viruses* **2019**, *11*, 1083. [CrossRef] [PubMed]
- Gurney, J.; Brown, S.P.; Kaltz, O.; Hochberg, M.E. Steering Phages to Combat Bacterial Pathogens. *Trends Microbiol.* 2020, 28, 85–94. [CrossRef]
- 87. Phillips, A.M.; Gonzalez, L.O.; Nekongo, E.E.; Ponomarenko, A.I.; McHugh, S.M.; Butty, V.L.; Levine, S.S.; Lin, Y.-S.; Mirny, L.A.; Shoulders, M.D. Host Proteostasis Modulates Influenza Evolution. *eLife* **2017**, *6*, e28652. [CrossRef]
- Domazet-Lošo, T.; Tautz, D. Phylostratigraphic Tracking of Cancer Genes Suggests a Link to the Emergence of Multicellularity in Metazoa. *BMC Biol.* 2010, *8*, 1–10. [CrossRef]
- 89. Diamond, M.S.; Farzan, M. The Broad-Spectrum Antiviral Functions of IFIT and IFITM Proteins. *Nat. Rev. Immunol.* **2013**, *13*, 46–57. [CrossRef]
- 90. Liebeskind, B.J.; McWhite, C.D.; Marcotte, E.M. Towards Consensus Gene Ages. Genome Biol. Evol. 2016, 8, 1812–1823. [CrossRef]
- Domazet-Lošo, T.; Carvunis, A.-R.; Albà, M.M.; Šestak, M.S.; Bakaric, R.; Neme, R.; Tautz, D. No Evidence for Phylostratigraphic Bias Impacting Inferences on Patterns of Gene Emergence and Evolution. *Mol. Biol. Evol.* 2017, 34, 843–856. [CrossRef] [PubMed]
- Sonnhammer, E.L.L.; Gabaldón, T.; Sousa da Silva, A.W.; Martin, M.; Robinson-Rechavi, M.; Boeckmann, B.; Thomas, P.D.; Dessimoz, C. Quest for Orthologs consortium Big Data and Other Challenges in the Quest for Orthologs. *Bioinformatics* 2014, 30, 2993–2998. [CrossRef] [PubMed]
- Boeckmann, B.; Marcet-Houben, M.; Rees, J.A.; Forslund, K.; Huerta-Cepas, J.; Muffato, M.; Yilmaz, P.; Xenarios, I.; Bork, P.; Lewis, S.E.; et al. Quest for Orthologs Entails Quest for Tree of Life: In Search of the Gene Stream. *Genome Biol. Evol.* 2015, 7, 1988–1999. [CrossRef] [PubMed]
- 94. Plenge, R.M.; Scolnick, E.M.; Altshuler, D. Validating Therapeutic Targets through Human Genetics. *Nat. Rev. Drug. Discov.* 2013, 12, 581–594. [CrossRef] [PubMed]

- 95. Cong, F.; Cheung, A.K.; Huang, S.-M.A. Chemical Genetics-Based Target Identification in Drug Discovery. *Annu. Rev. Pharmacol. Toxicol.* **2012**, 52, 57–78. [CrossRef]
- 96. Chen, X.; Ji, Z.L.; Chen, Y.Z. TTD: Therapeutic Target Database. Nucleic Acids Res. 2002, 30, 412–415. [CrossRef] [PubMed]
- Cotto, K.C.; Wagner, A.H.; Feng, Y.-Y.; Kiwala, S.; Coffman, A.C.; Spies, G.; Wollam, A.; Spies, N.C.; Griffith, O.L.; Griffith, M. DGIdb 3.0: A Redesign and Expansion of the Drug-Gene Interaction Database. *Nucleic Acids Res.* 2018, 46, D1068–D1073. [CrossRef]
- 98. Betz, U.A.K. How Many Genomics Targets Can a Portfolio Afford? Drug Discov. Today 2005, 10, 1057–1063. [CrossRef]
- 99. Hopkins, A.L.; Groom, C.R. The Druggable Genome. Nat. Rev. Drug. Discov. 2002, 1, 727–730. [CrossRef]
- 100. Betz, U.A.; Farquhar, R.; Ziegelbauer, K. Genomics: Success or Failure to Deliver Drug Targets? *Curr. Opin. Chem. Biol.* 2005, 9, 387–391. [CrossRef]
- Borrel, A.; Regad, L.; Xhaard, H.; Petitjean, M.; Camproux, A.-C. PockDrug: A Model for Predicting Pocket Druggability That Overcomes Pocket Estimation Uncertainties. J. Chem. Inf. Model. 2015, 55, 882–895. [CrossRef]
- 102. Hajduk, P.J.; Huth, J.R.; Fesik, S.W. Druggability Indices for Protein Targets Derived from NMR-Based Screening Data. *J. Med. Chem.* **2005**, *48*, 2518–2525. [CrossRef] [PubMed]
- 103. Kana, O.; Brylinski, M. Elucidating the Druggability of the Human Proteome with EFindSite. *J. Comput. Aided Mol. Des.* **2019**, *33*, 509–519. [CrossRef]
- 104. Fernández, A. Deep Learning to Therapeutically Target Unreported Complexes. *Trends Pharmacol. Sci.* 2019, 40, 551–554. [CrossRef] [PubMed]
- 105. Burley, S.K.; Berman, H.M.; Bhikadiya, C.; Bi, C.; Chen, L.; Di Costanzo, L.; Christie, C.; Dalenberg, K.; Duarte, J.M.; Dutta, S.; et al. RCSB Protein Data Bank: Biological Macromolecular Structures Enabling Research and Education in Fundamental Biology, Biomedicine, Biotechnology and Energy. Nucleic Acids Res. 2019, 47, D464–D474. [CrossRef]
- 106. Wang, Z.-Y.; Zhang, H.-Y. Rational Drug Repositioning by Medical Genetics. Nat. Biotechnol. 2013, 31, 1080–1082. [CrossRef] [PubMed]
- 107. Li, Z.; Yao, Y.; Cheng, X.; Chen, Q.; Zhao, W.; Ma, S.; Li, Z.; Zhou, H.; Li, W.; Fei, T. A Computational Framework of Host-Based Drug Repositioning for Broad-Spectrum Antivirals against RNA Viruses. *iScience* **2021**, *24*, 102148. [CrossRef]
- Boraschi, D.; Italiani, P.; Weil, S.; Martin, M.U. The Family of the Interleukin-1 Receptors. *Immunol. Rev.* 2018, 281, 197–232. [CrossRef] [PubMed]
- 109. Li, X.; Wu, N.; Ji, H.; Huang, Y.; Hu, H.; Li, J.; Mi, S.; Duan, S.; Chen, X. A Male-Specific Association between AGTR1 Hypermethylation and Coronary Heart Disease. *Bosn. J. Basic Med. Sci.* **2020**, *20*, 31–36. [CrossRef] [PubMed]
- 110. Al-Salama, Z.T.; Frampton, J.E. Glycopyrronium/Formoterol: A Review in COPD. Drugs 2019, 79, 1455–1466. [CrossRef]
- 111. Engering, A.; Geijtenbeek, T.B.H.; van Vliet, S.J.; Wijers, M.; van Liempt, E.; Demaurex, N.; Lanzavecchia, A.; Fransen, J.; Figdor, C.G.; Piguet, V.; et al. The Dendritic Cell-Specific Adhesion Receptor DC-SIGN Internalizes Antigen for Presentation to T Cells. J. Immunol. 2002, 168, 2118–2126. [CrossRef] [PubMed]
- 112. Lin, G.; Simmons, G.; Pöhlmann, S.; Baribaud, F.; Ni, H.; Leslie, G.J.; Haggarty, B.S.; Bates, P.; Weissman, D.; Hoxie, J.A.; et al. Differential N-Linked Glycosylation of Human Immunodeficiency Virus and Ebola Virus Envelope Glycoproteins Modulates Interactions with DC-SIGN and DC-SIGNR. *J. Virol.* **2003**, *77*, 1337–1346. [CrossRef] [PubMed]
- 113. Tassaneetrithep, B.; Burgess, T.H.; Granelli-Piperno, A.; Trumpfheller, C.; Finke, J.; Sun, W.; Eller, M.A.; Pattanapanyasat, K.; Sarasombath, S.; Birx, D.L.; et al. DC-SIGN (CD209) Mediates Dengue Virus Infection of Human Dendritic Cells. *J. Exp. Med.* 2003, 197, 823–829. [CrossRef]
- 114. Glinka, A.; Dolde, C.; Kirsch, N.; Huang, Y.-L.; Kazanskaya, O.; Ingelfinger, D.; Boutros, M.; Cruciat, C.-M.; Niehrs, C. LGR4 and LGR5 Are R-Spondin Receptors Mediating Wnt/β-Catenin and Wnt/PCP Signalling. *EMBO Rep.* 2011, 12, 1055–1061. [CrossRef]
- 115. Becker, L.; Huang, Q.; Mashimo, H. Immunostaining of Lgr5, an Intestinal Stem Cell Marker, in Normal and Premalignant Human Gastrointestinal Tissue. *Sci. World J.* 2008, *8*, 1168–1176. [CrossRef] [PubMed]
- Gaudet, P.; Livstone, M.S.; Lewis, S.E.; Thomas, P.D. Phylogenetic-Based Propagation of Functional Annotations within the Gene Ontology Consortium. *Brief. Bioinform.* 2011, 12, 449–462. [CrossRef]
- 117. Cascino, I.; Fiucci, G.; Papoff, G.; Ruberti, G. Three Functional Soluble Forms of the Human Apoptosis-Inducing Fas Molecule Are Produced by Alternative Splicing. *J. Immunol.* **1995**, *154*, 2706–2713.
- 118. Cao, P.; Zhang, W.; Gui, W.; Dong, Y.; Jiang, T.; Gong, Y. Structural Insights into the Mechanism of Calmodulin Binding to Death Receptors. *Acta Crystallogr. D Biol. Crystallogr.* 2014, 70, 1604–1613. [CrossRef]
- Oehm, A.; Behrmann, I.; Falk, W.; Pawlita, M.; Maier, G.; Klas, C.; Li-Weber, M.; Richards, S.; Dhein, J.; Trauth, B.C. Purification and Molecular Cloning of the APO-1 Cell Surface Antigen, a Member of the Tumor Necrosis Factor/Nerve Growth Factor Receptor Superfamily. Sequence Identity with the Fas Antigen. J. Biol. Chem. 1992, 267, 10709–10715. [CrossRef]
- 120. Fang, F.; Lue, L.-F.; Yan, S.; Xu, H.; Luddy, J.S.; Chen, D.; Walker, D.G.; Stern, D.M.; Yan, S.; Schmidt, A.M.; et al. RAGE-Dependent Signaling in Microglia Contributes to Neuroinflammation, Aβ Accumulation, and Impaired Learning/Memory in a Mouse Model of Alzheimer's Disease. *FASEB J.* 2010, 24, 1043–1055. [CrossRef]
- 121. Xue, J.; Rai, V.; Singer, D.; Chabierski, S.; Xie, J.; Reverdatto, S.; Burz, D.S.; Schmidt, A.M.; Hoffmann, R.; Shekhtman, A. Advanced Glycation End Product Recognition by the Receptor for AGEs. *Structure* **2011**, *19*, 722–732. [CrossRef]
- 122. Wu, T.; Shi, Z.; Baumgart, T. Mutations in BIN1 Associated with Centronuclear Myopathy Disrupt Membrane Remodeling by Affecting Protein Density and Oligomerization. *PLoS ONE* **2014**, *9*, e93060. [CrossRef] [PubMed]

- 123. Miyagawa, T.; Ebinuma, I.; Morohashi, Y.; Hori, Y.; Young Chang, M.; Hattori, H.; Maehara, T.; Yokoshima, S.; Fukuyama, T.; Tsuji, S.; et al. BIN1 Regulates BACE1 Intracellular Trafficking and Amyloid-β Production. *Hum. Mol. Genet.* 2016, 25, 2948–2958. [CrossRef] [PubMed]
- 124. Lingel, A.; Weiss, T.M.; Niebuhr, M.; Pan, B.; Appleton, B.A.; Wiesmann, C.; Bazan, J.F.; Fairbrother, W.J. Structure of IL-33 and Its Interaction with the ST2 and IL-1RAcP Receptors-Insight into Heterotrimeric IL-1 Signaling Complexes. *Structure* 2009, 17, 1398–1410. [CrossRef] [PubMed]
- 125. Liu, X.; Hammel, M.; He, Y.; Tainer, J.A.; Jeng, U.-S.; Zhang, L.; Wang, S.; Wang, X. Structural Insights into the Interaction of IL-33 with Its Receptors. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 14918–14923. [CrossRef] [PubMed]
- 126. Qiu, R.G.; Abo, A.; Steven Martin, G. A Human Homolog of the C. Elegans Polarity Determinant Par-6 Links Rac and Cdc42 to PKCzeta Signaling and Cell Transformation. *Curr. Biol.* **2000**, *10*, 697–707. [CrossRef]
- 127. Drenan, R.M.; Doupnik, C.A.; Boyle, M.P.; Muglia, L.J.; Huettner, J.E.; Linder, M.E.; Blumer, K.J. Palmitoylation Regulates Plasma Membrane-Nuclear Shuttling of R7BP, a Novel Membrane Anchor for the RGS7 Family. J. Cell Biol. 2005, 169, 623–633. [CrossRef]
- 128. Hunt, R.A.; Edris, W.; Chanda, P.K.; Nieuwenhuijsen, B.; Young, K.H. Snapin Interacts with the N-Terminus of Regulator of G Protein Signaling 7. *Biochem. Biophys. Res. Commun.* **2003**, 303, 594–599. [CrossRef]
- 129. Tao, W.; Pennica, D.; Xu, L.; Kalejta, R.F.; Levine, A.J. Wrch-1, a Novel Member of the Rho Gene Family That Is Regulated by Wnt-1. *Genes Dev.* 2001, *15*, 1796–1807. [CrossRef]
- Pitti, R.M.; Marsters, S.A.; Ruppert, S.; Donahue, C.J.; Moore, A.; Ashkenazi, A. Induction of Apoptosis by Apo-2 Ligand, a New Member of the Tumor Necrosis Factor Cytokine Family. J. Biol. Chem. 1996, 271, 12687–12690. [CrossRef]
- Ramamurthy, V.; Yamniuk, A.P.; Lawrence, E.J.; Yong, W.; Schneeweis, L.A.; Cheng, L.; Murdock, M.; Corbett, M.J.; Doyle, M.L.; Sheriff, S. The Structure of the Death Receptor 4-TNF-Related Apoptosis-Inducing Ligand (DR4-TRAIL) Complex. *Acta Crystallogr. F Struct. Biol. Commun.* 2015, *71*, 1273–1281. [CrossRef]
- 132. Doud, M.B.; Lee, J.M.; Bloom, J.D. How Single Mutations Affect Viral Escape from Broad and Narrow Antibodies to H1 Influenza Hemagglutinin. *Nat. Commun.* **2018**, *9*, 1–12. [CrossRef] [PubMed]
- Patel, M.C.; Shirey, K.A.; Pletneva, L.M.; Boukhvalova, M.S.; Garzino-Demo, A.; Vogel, S.N.; Blanco, J.C. Novel Drugs Targeting Toll-like Receptors for Antiviral Therapy. *Future Virol.* 2014, *9*, 811–829. [CrossRef]
- 134. Brelot, A.; Chakrabarti, L.A. CCR5 Revisited: How Mechanisms of HIV Entry Govern AIDS Pathogenesis. Available online: https://pubmed.ncbi.nlm.nih.gov/29932942/?from_term=CCR5+&from_pos=2 (accessed on 18 May 2020).
- 135. Grande, F.; Occhiuzzi, M.A.; Rizzuti, B.; Ioele, G.; De Luca, M.; Tucci, P.; Svicher, V.; Aquaro, S.; Garofalo, A. CCR5/CXCR4 Dual Antagonism for the Improvement of HIV Infection Therapy. *Molecules* **2019**, *24*, 550. [CrossRef]
- 136. Kumar, N.; Sharma, S.; Kumar, R.; Tripathi, B.N.; Barua, S.; Ly, H.; Rouse, B.T. Host-Directed Antiviral Therapy. *Clin. Microbiol. Rev.* **2020**, 33, e00168-19. [CrossRef]
- 137. McHutchison, J.G.; Manns, M.P.; Muir, A.J.; Terrault, N.A.; Jacobson, I.M.; Afdhal, N.H.; Heathcote, E.J.; Zeuzem, S.; Reesink, H.W.; Garg, J.; et al. Telaprevir for Previously Treated Chronic HCV Infection. *N. Engl. J. Med.* 2010, 362, 1292–1303. [CrossRef] [PubMed]
- 138. Imran, M.; Manzoor, S.; Khattak, N.M.; Khalid, M.; Ahmed, Q.L.; Parvaiz, F.; Tariq, M.; Ashraf, J.; Ashraf, W.; Azam, S.; et al. Current and Future Therapies for Hepatitis C Virus Infection: From Viral Proteins to Host Targets. Arch. Virol. 2014, 159, 831–846. [CrossRef] [PubMed]
- 139. Kaufmann, S.H.E.; Dorhoi, A.; Hotchkiss, R.S.; Bartenschlager, R. Host-Directed Therapies for Bacterial and Viral Infections. *Nat. Rev. Drug Discov.* **2018**, *17*, 35–56. [CrossRef]