EDITORIAL

Viral evolution and transmission effectiveness

Patsarin Rodpothong, Prasert Auewarakul

Patsarin Rodpothong, Prasert Auewarakul, Department of Microbiology, Faculty of Medicine Siriraj Hospital, Bangkok 10700, Thailand
Patsarin Rodpothong, Prasert Auewarakul, Center for Emerging and Neglected Infectious Diseases, Mahidol University, Bangkok 10700, Thailand
Author contributions: Rodpothong P and Auewarakul P contributed to this paper.
Correspondence to: Prasert Auewarakul, MD, Professor, Department of Microbiology, Faculty of Medicine Siriraj Hospital, 2 Prannok Road, BangkokNoi, Bangkok 10700, Thailand. prasert.aue@mahidol.ac.th
Telephone: +66-2-4197059 Fax: +66-2-4184148

Abstract

Different viruses transmit among hosts with different degrees of efficiency. A basic reproductive number (R_0) indicates an average number of cases getting infected from a single infected case. R_0 can vary widely from a little over 1 to more than 10. Low R_0 is usually found among rapidly evolving viruses that are often under a strong positive selection pressure, while high R_0 is often found among viruses that are highly stable. The reason for the difference between antigenically diverse viruses with low R_0, such as influenza A virus, and antigenically stable viruses with high R_0, such as measles virus, is not clear and has been a subject of great interest. Optimization of transmissibility fitness considering intra-host dynamics and inter-host transmissibility was shown to result in strategies for tradeoff between transmissibility and diversity. The nature of transmission, targeting either a naïve children population or an adult population with partial immunity, has been proposed as a contributing factor for the difference in the strategies used by the two groups of viruses. The R_0 determines the levels of threshold heard immunity. Lower R_0 requires lower herd immunity to terminate an outbreak. Therefore, it can be assumed that the outbreak saturation can be reached more readily when the R_0 is low. In addition, one may assume that when the outbreak saturation is reached, herd immunity may provide a strong positive selection pressure that could possibly result in an occurrence of escape mutants. Studies of these hypotheses will give us an important insight into viral evolution. This review discusses the above hypotheses as well as some possible mechanistic explanation for the difference in transmission efficiency of viruses

REPRODUCTIVE RATE AND ANTIGENIC DIVERSITY

Viruses transmit between hosts with various degrees of efficiency. A parameter called basic reproductive number (R_0) provides a quantitative estimate of the transmission efficiency. R_0 is the mean number of secondary infections that one case would produce in a fully susceptible population[1]. In order for an outbreak to occur, R_0 needs to be higher than 1. For some pathogens, R_0 can be relatively low. For example, the 2009 pandemic H1N1 influenza virus has been estimated to have R_0 of 1.3-1.7[2,3]. This

Table 1 Basic reproductive number and antigenic diversity of certain viruses

Virus	Reproductive No.	Antigenic diversity
Measles	12-18[4]	Stable, single serotype, long-lasting immunity
Mumps	5-7[4]	Stable, single serotype, long-lasting immunity
Rubella	5-7[4]	Stable, single serotype, long-lasting immunity
Poliovirus	5-7[4]	Stable, 3 serotypes, long-lasting immunity
Respiratory syncytial virus	1-2[16]	Two major antigenic groups, multiple antigenic variants within the groups, re-infection can occur[17,18]
Influenza	1-2[2,3]	Antigenic drift, numerous antigenic variants, re-infection is common

range is comparable to the R_0 estimates of other pandemic and seasonal influenza viruses. It may seem surprising that a virus that can cause devastating outbreaks at a global scale has such a low transmission efficiency. Despite the low R_0, influenza virus can spread rapidly owing to its short generation time. Some other viruses have much higher R_0. For example, R_0 of measles virus was estimated to be 12-18[4]. Table 1 shows R_0 of some common human viral pathogens.

It is remarkable that most viruses with high R_0 are antigenically stable and contain a single or a limited number of serotypes. Vaccine is usually highly effective, and escaping mutants have not been a problem. This group of virus often causes childhood diseases, such as mumps, measles, rubella, and poliomyelitis. In contrast, the antigenicity of viruses with low R_0, such as influenza virus, is highly diverse. New antigenic variants are constantly emerging to replace the preexisting strains. This so called "antigenic drift" is the most important problem in influenza vaccine production and implementation. Vaccine strains have to be predicted and selected for each outbreak season based on extensive surveillance data. Vaccination needs to be repeated annually and there is a risk of mismatch between vaccine and circulating strains[5].

Theoretically, mutation rate and selective pressure determine the rate of evolution and diversity of any organisms. The mutation rate can explain differences between DNA and RNA viruses. Because of the lack of proof reading mechanism in the genome replication, RNA viruses have mutation rates at several orders of magnitude higher than those of DNA viruses[6]. This explains why RNA viruses are much more diverse than DNA viruses, both antigenically and phenotypically. All the above mentioned viruses with different antigenic diversity are however RNA viruses, and they were shown to have comparable high mutation rates[6]. Therefore, the determining factor for the difference in antigenic diversity must be the selective pressures. However, selective pressure is a complex phenomenon that involves several aspects of host-pathogen interaction, such as the immune response and transmission environment. While host immune responses provide positive selective pressure favoring escape mutations, constraints for optimizing transmission and replication fitness exerts negative selective pressure to keep the optimal wild type unless there are changes in transmission conditions. Viruses use different strategies to tradeoff between these two types of selective pressure resulting in different levels of diversity.

TRANSMISSION FITNESS AS A FUNCTION OF WITHIN- AND BETWEEN-HOST DYNAMICS

Viruses transmit *via* different routes and under different conditions. In order to transmit efficiently, viruses need to shed high levels of infectious virions. However, producing high levels of progeny requires high replication rate, which makes them vulnerable to immune response or consumes too much resources (infected cells) leading to rapid progressive fatal diseases. Either way, high levels of viral replication cannot be sustained in long term. Viruses therefore need to choose either to replicate at a maximum rate in a short period, or to extend the period of replication and shedding with reduced replication rate. In order to persist in an extended period, viruses need to be able to replicate in the presence of specific immune responses and this requires an immune escape mechanism. Escape mutations leading to antigenic diversity is a common mechanism for the immune escape. A mathematical modeling showed that optimization of transmission fitness results in 3 groups of pathogens[7]: (1) childhood diseases, which are highly efficient in transmission over a short period; (2) sexually transmitted diseases, which have low transmission efficiency but can establish persistent infection; and (3) viruses with high antigenic diversity and low transmission rate, such as influenza. Contact rate may be the major determinant for the optimization options. A high contact rate, such as in childhood diseases, favors high replication and transmission rates, whereas a low contact rate favors an extended period of transmission and hence a persistent infection[8].

MAXIMIZING TRANSMISSION FITNESS *VS* TOLERATING MUTATIONS

Because viruses with high R_0 are often antigenically conserved, it is plausible to assume that they are under a strong negative selective pressure. It is likely that this selective pressure stems from the necessity to maintain the optimal structure for maximum transmission and replication fitness. Because protective epitopes usually overlap with receptor-binding domains, escape mutations often affect the binding affinity to the viral receptors and impair the viral fitness[9]. The ability to tolerate mutations would likely mean that the virus has given up its optimal fitness in exchange for structural flexibility. Along this

line of evolution, a virus would optimize its fitness with a constraint to keep the flexibility. In contrast, a virus could optimize its fitness to a higher level without this constraint, resulting in a highly efficient but un-flexible structure.

IMMUNE EFFECTIVENESS AS A BARRIER TO IMMUNE ESCAPE

A highly effective immune response provides a strong positive selective pressure that drives the emergence of an escape mutant. However, it may at the same time be a barrier for these escape mutations to occur. For example in the case of high levels of high-affinity antibodies that can recognize several minor structural changes on the virions, only drastic changes will be able to escape this antigen-antibody interaction[10]. These drastic changes will require several simultaneous mutations that may not be tolerated by the viral fitness. Similar to the case where an effective immune response targeting multiple targets, the virus will require multiple mutations to escape this response. Therefore, a highly effective immune response can be viewed as a barrier against the escape mutants.

It is well known that conserved viruses of childhood diseases elicit highly effective immune response with a lifelong protection, whereas immune response to influenza virus is short-lived and not always protective. This suggests that the effectiveness of specific immune response against those childhood diseases plays a role as a barrier to prevent the occurrence of escape mutants and antigenic diversity. On the other hand, an ineffective immune response would cause little antigenic changes on the virus because of the lack of selective pressure, and partially effective immune response may be the most effective force that drives most of the viral antigenic variation. Partially effective immune response enough to exert selective pressure, but not effective enough to suppress escape viral mutants is the most effective driving force of antigenic variation.

HERD IMMUNITY AND SELECTIVE PRESSURE

At a population level, non-immune individuals could be protected from an infectious agent if a sufficient fraction of the population is immuned. The benefit of indirect protection is caused by the interruption of transmission chain by immuned individuals. An outbreak can be aborted if the fraction of immuned individuals reaches a threshold level. This immune protection of a partially immuned population is called herd immunity. The level of herd immunity required to terminate an outbreak depends on R_0[4]. Pathogens with higher R_0 require a higher herd immunity level. The threshold herd immunity level can be calculated by $1 - 1/R_0$. For example, a disease with R_0 of 2 will need a herd immunity level at $1 - 1/2$ or 50% of the population, whereas a disease with R_0 of

10 will need a herd immunity level of $1 - 1/10$ or 90%. Theoretically, in a homogenous population with sufficient and evenly distributed contact rate, herd immunity level will be reached regardless of R_0. Although a virus with higher R_0 requires a higher herd immunity level, it can spread more rapidly resulting in a higher number of infected cases, and the required herd immunity level will be eventually reached. In reality where populations are heterogeneous, the threshold herd immunity level may never be reached. Contact rate does not distribute evenly in a population and can vary demographically and geographically. A realistic population consists of multiples compartments interacting with one another at various contact rates[11]. Although a pathogen with a high R_0 is highly effective in transmitting in a small population with a high contact rate, it may not be as effective in a larger population with multiple compartments interacting with a lower contact rate. Therefore, a high herd immunity level required to eliminate an outbreak with a high R_0 may not be reached and pockets of naïve hosts will always be available to perpetuate transmission chain. In contrast, a disease with a low R_0, which requires only a low level of herd immunity, will be more likely to acquire enough herd immunity. In addition, a disease with a lower R_0 depends less on high contact rate, thus it will be more effective in transmitting between compartments in a population. If an effective herd immunity level is reached, it will provide a strong selective pressure to select for immune escape. This would explain lineage extinction usually observed in influenza phylogeny and frequent emergence of a new escape variant as an antigenic drift. Without a capacity to change its antigenic epitopes, influenza would be extinct after sufficient herd immunity has been reached. Measles, on the other hand, can stay antigenically conserved because the required high herd immunity level has never been reached.

Influenza viruses are circulating in many animal species and interspecies transmission occasionally happens. Interspecies transmission can bring a new virus into human population, which would cause a pandemic due to the lack of herd immunity in the human population. However, after entering human population the course of viral evolution is mainly dictated by the interaction between the virus and human hosts since seasonal influenza viruses do not transmit back and forth between human and animal species. In contrast to the constant antigenic drift of seasonal influenza viruses, avian and swine influenza viruses do not exhibit rapid antigenic changes. Constant influx of new piglets and hatchlings as naïve hosts continues the transmission chain without the need for antigenic changes. It is interesting that the R_0 in this situation does not need to be as low as those observed in human outbreaks. Although a wide range of R_0 was estimated for influenza outbreaks in avian species, the number can be above 5 in some settings[12,13]. The higher R_0 of avian influenza outbreaks in birds and poultry, and the antigenic stability of these viruses contrast to the lower R_0 of seasonal influenza viruses with rapid antigenic

changes. This evidence supports the role of transmission effectiveness in the viral evolution strategy.

POTENTIAL MECHANISTIC DIFFERENCES IN TRANSMISSION FITNESS

Although the difference in R_0 among viruses clearly indicates that transmission fitness can vary among different viruses, mechanistic explanation for this variability is still lacking. Viral infectivity can be quantified *in vitro*, and physical quantity of viruses can be measured in genome copy number. These *in vitro* studies, however, may not accurately represent *in vivo* infectivity and transmission fitness. Theoretically, in order to obtain maximum transmission fitness, viruses should be shed in high quantity, spread effectively and stably maintained in the environment. They should also be able to efficiently penetrate protective barriers to the target cells after entering a new host, bind to target cells with high affinity, withstand host extra- and intra-cellular innate defense. Among these contributing factors, receptor-binding affinity is probably the most crucial one since most neutralizing epitopes overlap with receptor-binding domains on viral surface proteins. Optimization of the receptor binding may exert structural constraints on mutations of the antigenic epitopes. Evidence for any of these possible differences is still lacking. For example, viral load in respiratory secretion as quantified in viral RNA copy number did not seem to be different between influenza and measles[14,15], although the data were from different studies and the comparison may not be fully reliable. Similarly, direct comparison of the other viral characteristics between viruses with high and low R_0 should be studied. This will improve our understanding of these viruses.

REFERENCES

1   **Heffernan JM**, Smith RJ, Wahl LM. Perspectives on the basic reproductive ratio. *J R Soc Interface* 2005; **2**: 281-293
2   **Fraser C**, Donnelly CA, Cauchemez S, Hanage WP, Van Kerkhove MD, Hollingsworth TD, Griffin J, Baggaley RF, Jenkins HE, Lyons EJ, Jombart T, Hinsley WR, Grassly NC, Balloux F, Ghani AC, Ferguson NM, Rambaut A, Pybus OG, Lopez-Gatell H, Alpuche-Aranda CM, Chapela IB, Zavala EP, Guevara DM, Checchi F, Garcia E, Hugonnet S, Roth C. Pandemic potential of a strain of influenza A (H1N1): early findings. *Science* 2009; **324**: 1557-1561
3   **Yang Y**, Sugimoto JD, Halloran ME, Basta NE, Chao DL, Matrajt L, Potter G, Kenah E, Longini IM. The transmissibility and control of pandemic influenza A (H1N1) virus. *Science* 2009; **326**: 729-733
4   **Fine PE**. Herd immunity: history, theory, practice. *Epidemiol Rev* 1993; **15**: 265-302
5   **Boni MF**. Vaccination and antigenic drift in influenza. *Vaccine* 2008; **26** Suppl 3: C8-14
6   **Drake JW**. Rates of spontaneous mutation among RNA viruses. *Proc Natl Acad Sci USA* 1993; **90**: 4171-4175
7   **Lange A**, Ferguson NM. Antigenic diversity, transmission mechanisms, and the evolution of pathogens. *PLoS Comput Biol* 2009; **5**: e1000536
8   **Frank SA**, Bush RM. Barriers to antigenic escape by pathogens: trade-off between reproductive rate and antigenic mutability. *BMC Evol Biol* 2007; **7**: 229
9   **Watabe T**, Kishino H. Structural considerations in the fitness landscape of a virus. *Mol Biol Evol* 2010; **27**: 1782-1791
10 **Watabe T**, Kishino H, de Oliveira Martins L, Kitazoe Y. A likelihood-based index of protein protein binding affinities with application to influenza HA escape from antibodies. *Mol Biol Evol* 2007; **24**: 1627-1638
11 **Koopman JS**, Jacquez G, Chick SE. New data and tools for integrating discrete and continuous population modeling strategies. *Ann N Y Acad Sci* 2001; **954**: 268-294
12 **Stegeman A**, Bouma A, Elbers AR, de Jong MC, Nodelijk G, de Klerk F, Koch G, van Boven M. Avian influenza A virus (H7N7) epidemic in The Netherlands in 2003: course of the epidemic and effectiveness of control measures. *J Infect Dis* 2004; **190**: 2088-2095
13 **Comin A**, Klinkenberg D, Marangon S, Toffan A, Stegeman A. Transmission dynamics of low pathogenicity avian influenza infections in Turkey flocks. *PLoS One* 2011; **6**: e26935
14 **Li CC**, Wang L, Eng HL, You HL, Chang LS, Tang KS, Lin YJ, Kuo HC, Lee IK, Liu JW, Huang EY, Yang KD. Correlation of pandemic (H1N1) 2009 viral load with disease severity and prolonged viral shedding in children. *Emerg Infect Dis* 2010; **16**: 1265-1272
15 **Thomas B**, Beard S, Jin L, Brown KE, Brown DW. Development and evaluation of a real-time PCR assay for rapid identification and semi-quantitation of measles virus. *J Med Virol* 2007; **79**: 1587-1592
16 **Weber A**, Weber M, Milligan P. Modeling epidemics caused by respiratory syncytial virus (RSV). *Math Biosci* 2001; **172**: 95-113
17 **Galiano MC**, Luchsinger V, Videla CM, De Souza L, Puch SS, Palomo C, Ricarte C, Ebekian B, Avendaño L, Carballal G. Intragroup antigenic diversity of human respiratory syncytial virus (group A) isolated in Argentina and Chile. *J Med Virol* 2005; **77**: 311-316
18 **Martínez I**, Dopazo J, Melero JA. Antigenic structure of the human respiratory syncytial virus G glycoprotein and relevance of hypermutation events for the generation of antigenic variants. *J Gen Virol* 1997; **78** (Pt 10): 2419-2429