Literature Review

Eunhae Hannah Bae & Norman Goodacre

August 25, 2017

Below is the email I received from Norman about the project.

Hi guys,

I played around a bit with the virusmentha data that already has grouped virus-host ppis accroding to virus families. Here are the stats (number are the ppis, then family):

**2987 Herpesviridae**

**1593 Retroviridae**

**1305 Orthomyxoviridae**

**820 Paramyxomyxoviridae**

**569 Flaviviridae**

246 Adenoviridae

200 Poxviridae

162 Polyomaviridae

159 Bunyaviridae

114 Filoviridae

59 Reoviridae

56 Arterieviridae

56 Togaviridae

36 Hepadnaviridae

14 Rhabdoviridae\_

11 Arenaviridae

5 Coronaviridae

3 Hepeviridae

1 Baculoviridae

1 Siphoviridae

With those numbers  I guess we should do the analysis with the first 5:

Alternatively, we can skip Paramoxyviridae, as the remaining 4 groups provide the best investigated viruses in terms of their interactions and have a large(r) set of PPIs.

//

**Diversity of human viruses**

Statistics of viruses known to infect humans, genomes sequenced, genetic diversity?

Statistics of human disease caused by these viruses, case numbers, mortality, economic damage etc.

* Herpesviridae (mostly Herpes simplex, HSV)
  + Genomic architecture
    - Davison, Andrew J. "Comparative analysis of the genomes." (2007).

Herpes simplex viruses are placed in the family of *Herpesviridae* and consist of linear, double-stranded DNA molecules of about 125 to 240 kbp depending on virus types.

* + viruses known to infect humans
    - Burrel, Sonia, et al. "Ancient recombination events between human herpes simplex viruses." Molecular Biology and Evolution 34.7 (2017): 1713-1721.

Human herpes simplex virus 1 (HSV-1) is oral-facial herpes that is transmitted by oral-oral contact, known as ‘cold sores’. The number of HSV-1 infections was highest for Africa, South-East Asia and Western Pacific. HSV-1 can also cause genital herpes through oral to genital contact. Looker et al. explained that HSV-1 is an increasing cause of genital infection, and the prevalence of genital HSV-1 infection varies by region. Highest prevalence is in the Americas, followed by Europe and Western Pacific.

Human herpes simplex virus 2 (HSV-2) is primarily/entirely sexually transmitted genital herpes. Looker et al. explained that the highest prevalence of HSV-2 is in Africa, followed by Americas..

Looker et al. mentioned that there is a gender discrepancy in HSV infection that both HSV-1 and HSV-2 infections are more common among women than men.

* + - Looker, Katharine J., et al. "Global and regional estimates of prevalent and incident herpes simplex virus type 1 infections in 2012." PloS one 10.10 (2015): e0140765.
    - Brown, Zane A., et al. "Effect of serologic status and cesarean delivery on transmission rates of herpes simplex virus from mother to infant." Jama 289.2 (2003): 203-209.
    - Kropp, Rhonda Y., et al. "Neonatal herpes simplex virus infections in Canada: results of a 3-year national prospective study." Pediatrics 117.6 (2006): 1955-1962.

(although it has been known that HSV-2 is more common type of genital herpes) In Looker et al paper:: HSV-1 is more likely the cause of Neonatal herpes than HSV-2. A study found that in a 58,000 live births, mother’s genital HSV at delivery had transmitted HSV-1 more than HSV-2 (Brown et al., 2003). Also, 63% of neonatal herpes cases were due to HSV-1 in Canada from 2000 to 2003 (Kropp et al., 2006).

* + genomes sequenced

??

* + genetic diversity
    - Burrel, Sonia, et al. "Ancient recombination events between human herpes simplex viruses." Molecular Biology and Evolution 34.7 (2017): 1713-1721.

HSV-1 and HSV-2 show low overall genomic variability. HSV-1 has accumulated more genomic variation.

* + - Park, Donglim, et al. "Functional comparison of herpes simplex virus 1 (HSV-1) and HSV-2 ICP27 homologs reveals a role for ICP27 in virion release." Journal of virology 89.5 (2015): 2892-2905.

HSV-2 is biologically distinct from HSV-1. Genetically, it is highly related to HSV-1 with 83% protein sequence identify with nearly identical open reading frame arrangement. HSV-1 and HSV-2 can recombine in tissue cultures (*in vitro*).

In vivo, Park et al. suspects that recombinants of HSV-1 and HSV-2 are at a replicative disadvantage *in vivo*. And naturally occurring recombinants have not been described.

Burrel et al explained that HSV-2 infection seems to protect from further infection of the genitals with HSV-1. The transmission of recombinant HSV-1 comprising HSV-2 genome fragments can be expected to be very rare.

* + - Szpara, Moriah L., et al. "Evolution and diversity in human herpes simplex virus genomes." Journal of virology 88.2 (2014): 1209-1227.

HSV-1 can vary between infected individuals, over sequential isolates from the same individual, and by geographic region. High-throughput sequencing can be useful to see sequence diversity among HSV-1 strains, as PCR or polypeptide analysis has limitation for studying the whole genome. (earlier studies focused on recombination used PCR-based approaches ad only covered 2-4% of genome Burrel et al., 2017)

“During the coevolution of humans and HSV-1, spatial segregation of ancestral host populations is likely to have generated some geographic isolation of viral lineages” (Szpara et al).

* + human disease caused by these viruses
    - case numbers
      * Looker, Katharine J., et al. "Global and regional estimates of prevalent and incident herpes simplex virus type 1 infections in 2012." PloS one 10.10 (2015): e0140765.

Looker explained that the WHO has made estimates of the global burden of HSV-2 twice for 2003 and 2012. An estimated 417 million people were living with the infection in 2012.

Looker et al. found that 67% of the population (3.7 billion people under age of 50) had HSV-1 infection for 2012 globally.

140 million people aged 15-49-years were estimated to have genital HSV-1 infection worldwide in 2012

In the study of Burrel et al. infection with HSV-2 is less frequent but still reaches a global prevalence of 11%

* + - Mortality
      * Both infections are lifelong.
      * Most are asymptomatic (mild symptoms or blisters)
    - economic damage
      * Szucs, Thomas D., et al. "The estimated economic burden of genital herpes in the United States. An analysis using two costing approaches." BMC Infectious Diseases 1.1 (2001): 5.

Estimated 3.1 million symptomatic episodes per year in the USA, and the annual direct medical costs were at a minimum of $283 million (corresponding to 0.1% of the US health care expenditure ($1,007,300 million) to maximum of $984 million in 1996.

Indirect costs accounted for further $217 million.

* + - * Wang, Xiao, et al. "Meta-analysis of infectious agents and depression." Scientific reports 4 (2014).

A study (Wang et al. 2014) found that there were statistically significant associations between depression and infection with HSV-1.

HSV-2 infection and depression, which failed to show a significant association between the presence of HSV-2 DNA in brain tissue and depression. This was determined from finding out that there was no differences in HSV-2 IgG antibody titers between depressed patients and healthy controls. However, the author suggests that HSV-2 infection at childbirth might play a role in later development of depression.

* + - vaccine
* Retroviridae (mostly HIV)
  + Genomic architecture
  + viruses known to infect humans
  + genomes sequenced
  + genetic diversity / origin
    - AIDS is caused by two lentiviruses, HIV-1 and HIV-2, originated at around the same time from multiple cross-species transmission of SIVs in African primates, specifically from SIVcpz in chimpanzees to HIV-1 group M in human. These genetic changes are needed to accommodate cross-species. SIV newly introduced and accumulated in the new host, and there need to be enough accumulation to adapt to host proteins and restriction factors, and for further human-to-human spread (Sharp et al., 2011).
  + human disease caused by these viruses
    - case numbers
      * CDC reports (<https://www.cdc.gov/hiv/basics/statistics.html> )
        + Since the beginning of epidemic (which was officially began on June 5, 1981
        + 37,600 new HIV infections in 2014
        + 39,513 were diagnosed in 2015
        + 2.1 million new cases in HIV in 2015
        + Annual number of new diagnoses has been declining
        + 1.1 million people living with HIV at the end of 2014
        + 36.7 million living with HIV around the world as of June 2016, and approximately 60 million have been infected by the pandemic (M) form since 1981
    - Mortality
      * 1.1 million died from AIDS-related illnesses in 2015
      * 25 million deaths total due to M form since discovery in 1981 (Merson et al., 2008)
    - economic damage
      * A growing economy in a LMIC could shrink to a subsistence-level economy as a result of AIDS mortality
      * Application to S. Africa predicts that this could happen to S. Africa, with a population of over 55 million, within 4 generations (Bell et al., 2006)
    - vaccine
      * The neutralizing antibody (Nab) target is the gp120 (env) fusion protein
      * However, HIV-1 env is often glycosylated with non-immunogenic glycans, preventing binding of Nabs (Burton et al., 2004)
      * In addition, env forms a trimer that shields the binding site compared to monomer (Burton et al., 2004)
      * Kinetic and spatial constraints of Nab binding – binding of env to CD4 + CCR5/CXCR4 (Burton et al., 2004)
      * Nabs target the variable region of HIV-1 env, which is highly strain-specific (Burton et al., 2012)
      * A huge amount of diversity exists (Burton et al., 2012)
      * Attenuated vaccines expressing non-functional env tend to be unstable an induce non-neutralizing antibodies (Burton et al., 2012)
* Orthomyxoviridae (mostly Influenza)
  + Genomic architecture
    - Influenza viruses have segmented, negative sense, single strand RNA genomes and are placed in the family Orthomyxoviridae. Genome is organized into distinct segments \_\_\_
  + viruses known to infect humans
    - Primarily influenza viruses – three genera, A, B, and C. These three are also then divided into groups according to their fusion and exit proteins hemagglutinin (H) and neuraminidase (N). H is responsible for binding to the host cell surface, while N is responsible for cleaving sialic acid in glycoproteins of the host cell surface to allow the release of the particle.
    - The most recent epidemics of seasonal influenza A include H1N1 (swine flu), H3N2 (swine flu), H5N1 (avian flu), and H7N9 (avian flu). (WHO, 2009). Historical pandemics include 1918 H1N1, 1957 H2N2, and 1968 H3N2.
  + genomes sequenced
    - 443 complete genomes for family orthomyxoviridae have been sequenced as of May, 2017 – although ~450,000 entries exist in GenBank, many of them influenza A segments
  + genetic diversity
    - Swine flu: A(H1N1) viruses circulated in humans from 1918 until the A(H2N2) influenza pandemic of 1957. During this period there was substantial antigenic drift of A(H1N1) viruses in humans away from the 1918 virus ([2](http://science.sciencemag.org/content/325/5937/197.full#ref-2), [13](http://science.sciencemag.org/content/325/5937/197.full#ref-13)). A(H1N1) influenza viruses from the early 1950s reemerged in humans in 1977 ([14](http://science.sciencemag.org/content/325/5937/197.full#ref-14)). From 1977 to 2009, there was substantial further antigenic evolution of the human A(H1N1) viruses that was sufficient to warrant eight updates of the H1 component of the influenza virus vaccine ([15](http://science.sciencemag.org/content/325/5937/197.full#ref-15)) (Garten et al., 2009).
    - Avian flu: All 16 hemagglutinin and 10 neuraminidase variants have been detected in avian flu (Alexander et al., 2007; Chen et al., 2013). Modern influenza in birds as well as humans is thought to have been strongly influenced by a global “sweep” in the 19th and 20th centuries, which selected the most infectious internal genes of influenza A (Worobey et al., 2014). Because birds and swine serve as reservoirs for the virus in which recombination can occur, the exchange of lifestock across the global over the past two centuries has given influenza a unique opportunity to adapt to humans.
    - Evolution of the virus: reassortment and genetic drift are known to be common mechanisms that expand the diversity of both influenza A (Nelson and Holmes, 2007; Ghedin et al., 2009). Recombination, long suspected to play a role due to the observation the different segments of recent influenza strains’ genomes are from different original hosts (swine, avian, equine), was demonstrated to play a role in the evolution of influenza A ( He et al., 2008).
  + human disease caused by these viruses
    - case numbers
      * Swine flu: In April 2009, a previously undescribed A(H1N1) influenza virus was isolated from humans in Mexico and the United States ([19](http://science.sciencemag.org/content/325/5937/197.full#ref-19)). As of 18 May 2009, there have been 8829 laboratory-confirmed cases in 40 countries, resulting in 74 deaths ([20](http://science.sciencemag.org/content/325/5937/197.full#ref-20)–[23](http://science.sciencemag.org/content/325/5937/197.full#ref-23)). Of the 2009 A(H1N1) viruses, we have sequenced full or partial genomes of 17 isolated in Mexico, and 59 from 12 states in the United States (table S1) (Garten et al., 2009).
      * Avian flu: 200 people as well as millions of poultry were killed in a recent avian flu outbreak (H5N1) (Chang et al., 2007).
    - Mortality
      * Swine flu: In April 2009, a previously undescribed A(H1N1) influenza virus was isolated from humans in Mexico and the United States ([19](http://science.sciencemag.org/content/325/5937/197.full#ref-19)). As of 18 May 2009, there have been 8829 laboratory-confirmed cases in 40 countries, resulting in 74 deaths ([20](http://science.sciencemag.org/content/325/5937/197.full#ref-20)–[23](http://science.sciencemag.org/content/325/5937/197.full#ref-23)). Of the 2009 A(H1N1) viruses, we have sequenced full or partial genomes of 17 isolated in Mexico, and 59 from 12 states in the United States (table S1). Announced as the first pandemic of the 21st century by the WHO (Chang et al., 2009)
      * Avian flu: 2013 H7N9 associated with deaths, but no apparent outbreaks, in China. Individuals become infected as a result of contact with poultry (Chen et al., 2013).
    - economic damage
      * Poultry – HPAI (highly pathogenic avian influenza) usually due to H5 or H7 variants, can cause the death of hundreds of millions of agricultural birds, resulting in drastic economic setbacks as well as posing significant zoonotic threats (Alexander et al., 2007). In recent years, the number of birds infected as well as the methods and costs of disease control have increased dramatically, for reasons not yet fully understood.
      * Humans - Based on 2003 US population, we estimated that annual influenza epidemics resulted in an average of 610,660 life-years lost (undiscounted), 3.1 million hospitalized days, and 31.4 million outpatient visits. Direct medical costs averaged $10.4 billion (95% confidence interval [C.I.], $4.1, $22.2) annually. Projected lost earnings due to illness and loss of life amounted to $16.3 billion (C.I., $8.7, $31.0) annually. The total economic burden of annual influenza epidemics using projected statistical life values amounted to $87.1 billion (C.I., $47.2, $149.5) (Molinari et al., 2007).
    - Vaccine
      * In a study that followed seasonal influenza across several seasons, using hundreds of patient samples, it was found that patients often had multiple difference influenza A subtypes, and even combination of influenza A and influenza B (Ghedin et al., 2009)
* Paramyxoviridae (mostly measles)
  + Genomic architecture
  + viruses known to infect humans
  + genomes sequenced
  + genetic diversity
  + human disease caused by these viruses
    - case numbers
    - mortality
    - economic damage
    - vaccine
* Flaviviridae (Hepatitis C)
  + viruses known to infect humans
  + genomes sequenced
  + genetic diversity
  + human disease caused by these viruses
    - case numbers
    - mortality
    - economic damage
    - vaccine

Chen, Chien-Jen, et al. "Epidemiology of virus infection and human cancer." *Viruses and Human Cancer*. Springer Berlin Heidelberg, 2014. 11-32.

The International Agency for Research (IARC) is part of the World Health Organization and its main research focus is on human cancer and causes based on their epidemiological and laboratory data. There are viruses that cause cancer in human and identified as Group 1 carcinogen (Group 1 is classified as “known to be human carcinogens”) by IARC. The 7 viruses that fall into Group 1 category are: Epstein-Barr virus (EBV), human papillomavirus (HPV), human T cell lymphotrophic virus, type-1 (HTLV-1), Kaposi’s sarcoma herpes virus (KSHV) (4 direct chronic carcinogens), hepatitis B virus (HBV), hepatitis C virus (HCV) (2 indirect carcinogens through chronic inflammation), and human immunodeficiency virus, type-1 (HIV-1) (indirect carcinogen through immune suppression).

References:

Bell, C., Devarajan, S. and Gersbach, H., 2006. The long-run economic costs of AIDS: A model with an application to South Africa. *The World Bank Economic Review*, *20*(1), pp.55-89.

Burton, D.R., Desrosiers, R.C., Doms, R.W., Koff, W.C., Kwong, P.D., Moore, J.P., Nabel, G.J., Sodroski, J., Wilson, I.A. and Wyatt, R.T., 2004. HIV vaccine design and the neutralizing antibody problem. *Nature immunology*, *5*(3), pp.233-236.

Burton, D.R., Ahmed, R., Barouch, D.H., Butera, S.T., Crotty, S., Godzik, A., Kaufmann, D.E., McElrath, M.J., Nussenzweig, M.C., Pulendran, B. and Scanlan, C.N., 2012. A blueprint for HIV vaccine discovery. *Cell host & microbe*, *12*(4), pp.396-407.

Chang, L.Y., Shih, S.R., Shao, P.L., Huang, D.T.N. and Huang, L.M., 2009. Novel swine-origin influenza virus A (H1N1): the first pandemic of the 21st century. *Journal of the Formosan Medical Association*, *108*(7), pp.526-532.

Chen, Y., Liang, W., Yang, S., Wu, N., Gao, H., Sheng, J., Yao, H., Wo, J., Fang, Q., Cui, D. and Li, Y., 2013. Human infections with the emerging avian influenza A H7N9 virus from wet market poultry: clinical analysis and characterisation of viral genome. *The Lancet*, *381*(9881), pp.1916-1925.

Chen, Chien-Jen, et al. "Epidemiology of virus infection and human cancer." *Viruses and Human Cancer*. Springer Berlin Heidelberg, 2014. 11-32.

Garten, R.J., Davis, C.T., Russell, C.A., Shu, B., Lindstrom, S., Balish, A., Sessions, W.M., Xu, X., Skepner, E., Deyde, V. and Okomo-Adhiambo, M., 2009. Antigenic and genetic characteristics of swine-origin 2009 A (H1N1) influenza viruses circulating in humans. *science*, *325*(5937), pp.197-201.

Ghedin, E., Fitch, A., Boyne, A., Griesemer, S., DePasse, J., Bera, J., Zhang, X., Halpin, R.A., Smit, M., Jennings, L. and George, K.S., 2009. Mixed infection and the genesis of influenza virus diversity. *Journal of virology*, *83*(17), pp.8832-8841.

He, C.Q., Xie, Z.X., Han, G.Z., Dong, J.B., Wang, D., Liu, J.B., Ma, L.Y., Tang, X.F., Liu, X.P., Pang, Y.S. and Li, G.R., 2008. Homologous recombination as an evolutionary force in the avian influenza A virus. *Molecular biology and evolution*, *26*(1), pp.177-187.

Merson, M.H., O'Malley, J., Serwadda, D. and Apisuk, C., 2008. The history and challenge of HIV prevention. *The lancet*, *372*(9637), pp.475-488.

Nelson, M.I. and Holmes, E.C., 2007. The evolution of epidemic influenza. *Nature reviews. Genetics*, *8*(3), p.196.

Sharp, Paul M., and Beatrice H. Hahn. "Origins of HIV and the AIDS pandemic." Cold Spring Harbor perspectives in medicine 1.1 (2011): a006841. (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3234451/>)

World Health Organization. Global Influenza Programme, 2009. *Pandemic influenza preparedness and response: a WHO guidance document*. World Health Organization.

Worobey, M., Han, G.Z. and Rambaut, A., 2014. A synchronized global sweep of the internal genes of modern avian influenza virus. *Nature*, *508*(7495), p.254.