Journal of Medical Virology

The phylogenetic approach for infectious disease evolution and epidemiology: an updating review

Journal:	*Journal of Medical Virology*
Manuscript ID	JMV-19-7387.R1
Wiley - Manuscript type:	Review
Date Submitted by the Author:	n/a
Complete List of Authors:	Ciccozzi, Massimo; Universita Campus Bio-Medico di Roma Lai, A.; University of Milan Zehender, Gianguglielmo; University of Milan Borsetti, Alessandra; Istituto Superiore di Sanita Cella, Eleonora; Universita Campus Bio-Medico di Roma Ciotti, Marco; University of Rome Tor Vergata Sagnelli, Evangelista; University of Campania Luigi Vanvitelli Sagnelli, Caterina; University of Campania Luigi Vanvitelli Angeletti, Silvia; Universita Campus Bio-Medico di Roma
Keywords:	Evolution, Epidemiology, Research and Analysis Methods

The phylogenetic approach for viral infectious disease evolution and epidemiology: an updating review

Massimo Ciccozzi,[1*] Alessia Lai, [2] Gianguglielmo Zehender, [2] Alessandra Borsetti, [3] Eleonora Cella,[1] Marco Ciotti, [4] Evangelista Sagnelli, [5] Caterina Sagnelli, [5] Silvia Angeletti [6]

[1] Unit of Medical Statistics and Molecular Epidemiology, University Campus Bio-Medico of Rome, Italy.

[2] Department of Biomedical and Clinical Sciences 'L. Sacco', University of Milan, Milan, Italy.

[3] National HIV/AIDS Research Center, Istituto Superiore di Sanità, Roma, Italy.

[4] Laboratory of Molecular Virology, Polyclinic Tor Vergata Foundation, Rome, Italy.

[5] Department of Mental Health and Public Medicine, Section of Infectious Diseases, University of Campania Luigi Vanvitelli, Naples, Italy.

[6] Unit of Clinical Laboratory Science, University Campus Bio-Medico of Rome, Italy.

Running Head: Phylogenesis in viral infectious disease

*Corresponding Author

Prof Massimo Ciccozzi

Unit of Medical Statistics and Molecular Epidemiology,

University Campus Bio-Medico of Rome

Via Alvaro del Portillo, 21

00128 Rome, Italy

Abstract

In the last decade, phylogenetic approach is recurrent in molecular evolutionary analysis. On May 12th, 2019, about 2.296.213 papers are found, but typing "phylogeny" or "epidemiology AND phylogeny" only 199.804 and 20.133 are retrieved, respectively. Molecular epidemiology in infectious diseases is widely used to define the source of infection as so as the ancestral relationships of individuals sampled from a population. Coalescent theory and phylogeographic analysis have had scientific application in several, recent pandemic events, and nosocomial outbreaks. Hepatitis viruses and Immunodeficiency Virus (HIV) have been largely studied. Phylogenetic analysis has been recently applied on Polyomaviruses so as in the more recent outbreaks due to different Arboviruses type as Zika and Chikungunya viruses discovering the source of infection and the geographic spread.

Data on sequences isolated by the microorganism are essential to apply the phylogenetic tools and research in the field of infectious disease phylodinamics is growing up. There is the need to apply molecular phylogenetic and evolutionary methods in areas out of infectious diseases, as translational genomics and personalized medicine. Lastly the application of these tools in vaccine strategy so as in antibiotic and antiviral researchers are encouraged.

Key word: Evolution; Epidemiology; Research and analysis methods

Introduction

Typing in PubMed "Epidemiology" at the date of May 12th, 2019, about 2,296,213 papers can be found, but typing "Phylogeny" only 199,804 are retrieved. Curiously typing "epidemiology AND phylogeny" 20,133 articles appear (Figure 1). This is conceivable thinking to the origin of epidemiology and phylogeny. Epidemiological concept was born at the time of Hippocrates considered to be the first epidemiologist (Figure 2A). Hippocrates was a Greek physician, he wrote a book called "On AIRS WATERS, PLACES" were he described epidemics. On the other hand, phylogeny probably has as father Charles Darwin (Figure 2B) and his principles in a theory of evolution due to natural selection, principles and theory divide with the thinking of Russel Wallace. The first phylogenetic tree is due to Walter Fitch and Emanuel Margoliash, in 1967 "Construction of phylogenetic trees. A method based on mutation distances as estimated from cytochrome c sequences is of general applicability "(1). Phylogenetics is a branch of molecular epidemiology that infers knowledge about taxonomy so as the evolution of microorganisms (2). It is a powerful tool, widely used in the study of rapidly evolving RNA viruses as well as recently in bacteria applying phylogenetic analysis in nosocomial infection (3-5). In the last years, a number of new methods that infer phylogenetic trees have been introduced. These methods are based on Bayesian theory introducing the statistical phylogenetic with two important concepts: the molecular clock hypothesis and the coalescent theory (6-8). Genetic distances and phylogenetic trees (coupled with a correct epidemiological design i.e., cross sectional studies), inferred via different sequence evolutionary models, are normally used to assign the genotype (9). In addition, phylogeny has been widely used to define circulating recombinant forms (CRFs), to discover mosaics and complex form of the virus (10-13). Coalescent theory and the molecular clock hypothesis are instead used to study the ancestral relationships of individuals sampled from a population (i.e. longitudinal studies) which can be inferred from a gene genealogy (phylogenetic tree) (10, 13-16). The link between epidemiological design and phylogeny inference can be seen as a hypothetical cycle starting with the specific evolutionary hypothesis and going on through data generation, data analysis and test hypothesis (Figure 3). Depending on the scientific question raised and on the availability of molecular epidemiological data, normally the first step analysis is based on maximum likelihood criteria (12), followed by phylodinamic criteria, molecular clock hypothesis and Bayesian techniques that are normally applied (12). Molecular epidemiological investigation has a historical and scientific application in Global infection diseases. In several relevant and recent pandemic events, as severe acute respiratory syndrome SARS (17-20), avian influenza A subtype H5N1 virus (21-23), H1N1 influenza A virus (24-28), Zaire Ebola virus (EBOV) (29-32) and lastly Zika virus (ZIKV), phylodinamic analysis

has been fruitful applied (33-36). All these epidemic events have been resolved, first applying the principles of classical epidemiology, quarantine and isolation, and then using phylodinamic principles representing drivers to the solution of the origin and the diffusion of the epidemics.

Hepatitis

The population dynamics and the circulation of hepatitis viruses has been investigated to better understand the origin and the migrations of these viruses by phylogenetic and phylodinamic approaches (Figure 4).

Phylogenesis and phylodinamic of Hepatitis A Virus (HAV)

Molecular epidemiology and phylogenetic analysis of HAV isolates have been proved crucial for understanding the demographic, social and environmental factors responsible for the current worldwide epidemiological patterns.

HAV is non-enveloped, single-stranded RNA virus belonging to the Picornaviridae family, genus Hepatovirus.

Molecular epidemiological investigation of HAV outbreaks is usually done using a short genomic region, like the VP1-P2A junction for the greatest simplicity in obtaining reliable results during an epidemic.

Several foodborne HAV outbreaks have been described and HAV infected food handlers have been frequently identified as a source of HAV outbreaks (37-40). The food-borne outbreak of hepatitis A occurred in the U.S.A. in 2005 was associated with the consumption of contaminated row oysters and phylogenetic analysis identified the same VP1-P2B sequence in the oysters and in oyster consumers (41, 42); a more detailed analysis of whole-genome sequences showed that HAV strains involved in the outbreak were genetically heterogeneous but closely related (43).

Two important outbreaks related to the consumption of imported raw or undercooked mussels have occurred in 1996-1997 and in 2002 in Campania and Apulia regions of Southern Italy, respectively (44), with most of the HAV isolates clustering as genotype Ia by phylogenetic analysis (45). Between September and November 2003, 1,023 cases of hepatitis A cases had been registered in in Tennessee, North Carolina, Georgia and Pennsylvania (U.S.A.) (46, 47), associated to the consumption of green onions grown in Mexico. Sequence analysis of the VP1-P2B region identified 3 HAV variants tightly clustered in phylogenetic tree (46, 47). However, the evaluation of the HAV whole-genome sequences showed that a much greater number of strains were involved (43).

An outbreak of acute hepatitis A due to the consumption of semi-dried tomatoes occurred in England in 2012 (48). The phylogenetic analysis identified multiple HAV strains of subtype IB and attributed the

genetic variability to viral evolution at the local source of the production rather than to substitutions in each infected individual (43).

Genetic analyses of HAV outbreaks has demonstrated the occurrence of person-to-person transmission in large families, day care centers or schools, commonly due to contacts with infected asymptomatic children (49) acting as reservoirs of HAV infection. In accordance with the above, Pelletier et al. reported a limited outbreak of hepatitis A in a large family in the U.S.A., in which the cases were primary or secondary contacts of an HAV infected foreign adoptee (50).

Molecular analysis is crucial for the identification of the source of HAV infection, especially when infected cases do not declare risk factors (51, 52), especially in countries with low HAV endemicity, where outbreaks are usually associated with a single HAV strain (53). Conversely, when outbreaks occur in highly endemic geographic areas an exposure to several HAV strains should be expected (54-60).

In recent years, numerous outbreaks of acute hepatitis A have occurred in men who have sex with men (MSM) in Europe (61, 62), in the U.S.A. (63) and in Japan (64), with specific HAV IIIA strains circulating among MSM communities identified by phylogenetic trees (65-68).

A large outbreak of hepatitis A has been identified in central Italy, with 523 cases identified between January 2016 and March 2017. This outbreak had an extremely high male to female ratio, suggesting that sexual practices between males were implicated in the transmission of this infection. In support of this hypothesis the Authors found that several cases were due to the HAV variant circulating in Europe in the same period among MSM (VRD_521_2016) (69).

Monophyletic clusters of HAV IA variants have been also identified in MSM in Spain, Italy, U.S.A. (66, 67, 70) and in previous HAV outbreaks occurred in MSM in several European countries.

Outbreaks of acute hepatitis A in IDUs has been reported in North American and North European countries (71-73). In Norway, IIIA strains have been identified in IDUs with parenterally acquired HAV infection (74). Two samples clustered in sub-genotype IIIA, closely related to variants identified in outbreaks occurred in Norwegian and Swedish IDUs communities in the same period (75).

The analysis of sequences based on sub-genomic regions is of practical use for epidemic investigation, but it does not allow a specific identification of the HAV strain. Instead, a more precise genetic resolution of the HAV strains can be achieved by the analysis of the whole genome sequences and intra-host heterogeneity, which represent a great opportunity for phylogenetic analysis and deeper knowledge of HAV molecular evolution.

Phylogenesis and phylodinamic of Hepatitis B Virus (HBV)

Phylogenetic studies may help to understand the origin, evolution and geographic dispersion of HBV due to its high genetic heterogeneity.

The main difficulty in reconstructing the HBV phylodinamic is the lack of a consensus in the estimation of the rate of evolution of the virus that affects also the tMRCA estimates and the timescales of HBV evolution (76). Studies based on external ("fossil") calibration points (77, 78, 79) describe HBV as a slowly evolving virus, whereas those based on internal calibration approaches (such as the use of heterochronous sequences) indicate that it is a highly variable virus evolving at a rate comparable with that of retroviruses (80). A recent study, re-estimating the viral evolutionary rate in known transmission chains, observed different rates at intra and inter host levels suggesting a reduced rate in transmission event with in general lower evolutionary rates on transmitted lineages (80).

1. Phylogenesis of genotype D

Several studies have reconstructed very different histories for the different HBV genotypes (81).

From the first studies, the original introduction of genotype D has been dated in the early 20th century (76, 82).

The phylogeographical reconstruction of the HBV genotype D history (83) indicates that it originated in India being the subgenotype D5 the first to diverge. A common ancestor of the remaining HBV-D subgenotypes left India in the first decade of the twentieth century and reached central Asia, where D1 and D3 diverged. Subsequently, they spread to Europe and Mediterranean area by means of a Southwestern pathway (mainly followed by subgenotype D1), reaching North Africa and the Southeastern Mediterranean through the Middle East. A second north-western pathway (due to the spread of D2) that overflow in the eastern Europe and through the former Soviet Union and iron curtain countries reached the Mediterranean basin (Albania) (83).

The World Wars played a crucial role in the global spread of HBV-D from India to the rest of the world, but the further spread of the infection was probably sustained by unsafe use of injections in medical practice (84).

2. The phylogeny of genotype A

Genotype A originated in Africa and penetrated in Europe in the XVI and XVII centuries following the Portuguese trades and ships and in Asia because of trade and travel between eastern Africa and southern Asia (85).

Several studies suggested a recent origin of HBV-A2 in Europe, between 1960s and 1980s through the sexual transmission, particularly men-having-sex-with-men (MSM) (76, 86) among who a single clonal strain has been isolated among high-risk subjects all over the world (87, 88).

HBV-A1 most probably originated in Africa and the slave trade and colonization played a major role in its global dispersion (89).

A recent phylogeographic study suggested that the origin of the currently circulating HBV-A subgenotypes originated in West Africa more than 1,000 years ago and was exported to Asia in the 17th century through the Portuguese or Arab trades and to Latin America in the following century through the transatlantic slave trade. On the contrary, HBV-A2 subgenotype originated more recently, in the 20th century and spread in Western countries between 1970s and 1990s, mainly among subjects at high risk of infection (90).

3. The phylogeny of genotype E

The observation that all isolates of genotype E are included in a single monophyletic group together and are absent among Afro-Americans (91). On the contrary, the phylodinamic studies, reconstructing the timescale of the HBV-E evolution suggested its spread in west Africa between 200 and 60 years ago (92). These observations support the view that the explosive spread of HBV-E in Africa must have been due to a new and highly efficient route of transmission, probably the unsafe use of needles during numerous mass-vaccination campaigns, in the 1920s and 1960s (92, 93). A recent study confirmed the origin of HBV-E to half XIX century, and showed that it was introduced in Madagascar before 1970s but remained confined until recently, when an exponential growth was observed as far as the first decade of 2000s, when the skyline plot showed a plateau (94).

4. The phylogeny of genotypes F and H

The observation that genotypes F and H share a common ancestor and are evolutionarily distant from the Old World genotypes but closely related to an isolate obtained from a woolly monkey of the New World suggested a long history of evolution of these genotypes, possibly representing the result of a cross-species transfer (95).

The phylogeography of HBV-F indicates the pre-Columbian origin of HBV-F and its expansion during the rapid increase in the Latin American population since the XVIII century (78).

A study of HBV-F in Colombia, demonstrated that HBV-F3 was probably the oldest F sub-genotype being the most related to genotype H (96).

Other studies (97) have suggested a possible origin of the F and H genotypes going back to the initial human New World settlements around 13,000 years ago and a divergence time at least 10,000 years ago.

5. The phylogeny of genotypes B and C

A recent study, using a "fossil" calibration approach have proposed that HBV genotype C as the oldest human genotype originating about 30.0 kya (98). This result seems to be confirmed by the detection of HBV nucleic acids in a Korean mummy of the 16th century which phylogenetically grouped with subgenotype C2 (99) suggesting a relatively low rate of HBV-C evolution.

To date no phylogenetic studies were conducted on genotype G. Two new genotypes (I and J) have been recently proposed (100, 101); but they have been shown to be recombinant forms and their identification remained as far as mainly sporadic (102).

Phylogenesis and phylodinamic of Hepatitis C Virus (HCV)

About 3% of the world's population are chronically infected with HCV, accounting for 150–180 million people worldwide, with about 3.5 million new chronic infections per year (103, 104). It has ben also estimated that about 350,000 deaths per year and a quarter of cases of hepatocellular carcinoma (HCC) are attributable to HCV infection (105).

Although a progressive decline in the incidence of HCV acute infection has been registered in Western countries, HCV chronic HCV infection remains a serious health problem (106, 107).

The phylogenetic analysis verifies how the viral strains are genetically correlated, if mutation is the only cause of viral differences. Regarding HCV, phylogenetic clustering has been used as a tool to identify the characteristics associated with a higher probability of HCV transmission. (108, 109).

In Western countries, HCV-genotypes 1A, genotype 1B and genotype 3A have been widely spread through the transfusion of blood or its derivate before a reliable test to identify HCV infection had become available and through needle sharing between drug users (IDUs) (110). At present, these HCV genotypes affect most HCV patients in these geographic areas (111-113).

Both HCV genotypes and subtypes (114) are differently distributed around the world (115-117). HCV genotyping is achievable by phylogenetic methods. HCV- genotypes 1 and 3 are worldwide distributed, (118). Genotypes 1 and 2 are endemic in Western African countries (Burkina Faso, Ghana, Guinea Bissau, Republic of Benin and Nigeria) but sporadically found also in Italy (119-124), genotypes 1 and 4 predominates in the Democratic Republic of Congo and Gabon and genotype 4 in Central and Middle East Africa so as in Turkey (120, 125-131). HCV genotype 3 is endemic in Asia (39% of all infections), where also genotype 6 is widely distributed in the South-East regions (132). Genotype 5 is the most common HCV genotype in South Africa and in Belgium and genotype 7 has been detected only in Central African immigrants tested in Canada (133).

Phylogenetic studies on HCV subtypes showed that genotype 3A strongly predominates in North America, Europe and Oceania, whereas it accounts for only half chronic infection in India, where other subtypes have been detected like genotype 3B in 20,3% of cases, genotype 3G in 6,8% and 3I in 3,6% (134, 135).

Among factor impairing the response to Direct Acting Antivirals (DAAs) is the presence of baseline Resistance-Associated Substitutions (RASs) as evidenced in case of DAA-naive patients with genotype 3 infection showing RASs in the NS5A region (136). The relevance and the clinical impact of these strains strongly suggest further phylogenetic and epidemiological investigations.

Phylogenesis and phylodinamic of Hepatitis E virus (HEV)

Only limited works focused their attention in the phylogenetic studies of HEV genotypes, in particular those causing zoonotic infections (HEV-3 and -4). An early study performed from Purdy et al. (137) estimated a divergence time of the genotypes 1–4 ancestor between anthropotropic and enzootic genotypes about 536 to 1344 years ago and suggested a possible enzootic ancestor concomitantly with the split of human and swine variants.

By Skyline plot analysis, genotype 1a number of infection increased until 1970-1980 years followed by a plateau in the last twenty years, whereas genotypes 3 and 4 increased around 1940 to 1945 years and decreased around 1990. The hypothesis was that the observed increases could be related to World War II and the population movement from urban to more rural settings or, additionally, to more lax sanitation procedures.

HEV-3 is the best described genotype. It is classified in 10 subtypes (3a-j) with a strong geographic structure, which group into two main clades: one including subtypes 3abchij and the second including 3efg (138). Based on phylodinamic and phylogeographic studies, the second clade (called B in the study of Zehender et al.) (139) consists of mainly European strains, while the first clade (A) includes a subclade of European isolates (3c) and a second subclade of mainly Asian and all the North American (grouping into subtype 3a) isolates. Dated trees suggested the origin of the HEV-3 in the XIX or XVIII century (139, 140), while the penetration of this virus in Asia dated the first decades of 1900s. The oldest clade resulted 3b, with an origin in 1920s including Japanese isolates (141) suggesting an introduction of this strain in Japan after the fall of the ban on meat consumption and the increase of pork importation mainly from Europe (139, 141). The phylodinamic analysis showed an exponential growth of the infection between 1950s until recently (2000s). A recent study published two complete genomes of genotype 3 isolated, one was classified as 3i, a subtype already detected in boars in Europe, but never

detected before in Italy either in animals or in humans, the second strain did not resulted classified in the subtypes defined to date (142).

Genotype 4 has been described mainly in Asia and it is classified into nine subtypes (4a-i) but was recently isolated in a small epidemic in Rome (143) in five patients living in the same area without travel in endemic areas. Given the scarcity of available sequences of HEV-4, only a single study described the phylodinamic of this HEV genotype, which suggests that the most probable origin of this genotype is Japan in the first decade of 1900s. The virus then was exported to China and from there to India, Indonesia, Korea, Taiwan and, sporadically, also to Europe (144).

Phylogenetic and phylodinamic analysis contributed to knowledge deepening in the evolutionary reconstruction of HEV virus history and transmission dynamic.

Phylogenesis and phylodinamic of Human Immuno deficiency Virus (HIV)

In the field of HIV-1 the phylogeny was primarily used to identify new genetic forms and trace the spread of HIV-1 subtypes. The combination with recombination analysis permitted the definition of an increasing number of CRFs (circulating recombinant forms), 98 published to date.

Tracing of genetic forms is important for epidemiological purposes but can also be of relevance in clinical settings, as it is well known that some biological properties differ among subtypes. The subtype variants can be affect antiviral drug resistance development (145-150), the genetic barrier to drug resistance development and/or disease progression (151-154). A recent paper from Abecasis et al. (155) studied the HIV epidemics in Europe and analyzed the circulation of HIV-1 non B subtypes. Authors concluded that it is clear that some non-B subtypes imported into Europe remained largely limited to migrant populations and risk groups analysis, where subtypes are clearly compartmentalized. The prevalence of subtype B was higher in the MSM risk group compared to heterosexuals that are more frequently infected with non-B subtypes. This evidence indicated highly stratified epidemics in different countries and risk groups suggesting the need of target preventive measures to specific populations. The work by Magiorkinis et al. (156) describes the global pattern of HIV-1 migration across the Western Hemisphere. On the contrary, of Gilbert et al that evidenced the role of the American continent and the Caribbean as “outwards” for the Western epidemic at the initial random migration event (157). Magiorkinis et al. evidenced a constant subsequent spread to the rest of the world. In the global context, this study defined the limited role of Europe in the spread of HIV Western epidemic, highlighting that the incoming infections spread mainly among regional populations.

Many recent studies have focused their attention on the study of migrants in different countries as they frequently are at increased risk of many diseases compared with native populations due to many factors

such as structural barriers and social inequalities. Determining when HIV-1 acquisition has occurred, pre migration or post migration, is crucial to design adequate HIV-1prevention programs, HIV-1 testing strategies and treatment services. Recently, some evidence supports that migrants acquire HIV-1 infection post migration. In reference to the European situation, data from Alvarez-del Arco et al. (158) reported that 63% of HIV-positive migrants acquired infection after migrating into Europe. Also in this work, the proportion of post migration infection varies according to patient origin and risk categories with a highest percentage of HIV-1 diagnosis in subjects from Latin America and Caribbean (71%), in particular in MSM (79%). Similarly, Paraskevis et al. (159) found that the HIV-1 infections in migrants were more frequently acquired after the arrival in Greece and through contacts between migrants. Comparable results were reported from Pantazis et al. (160) that reported the post migration infection in half of studied subjects (55,3%).

A number of published works evaluated at local level the circulation of HIV variants in many countries and in specific population such as MSM (161-163) and IDU (164-165). These papers highlighted the presence of multiple epidemic clusters, tried to reconstruct the date of HIV introduction in different areas and analyzed their dynamics at population level. Theses knowledge represents a substantial increase in our depth of knowledge on which interventions can be based.

Phylogenetic analysis is an important way to trace epidemiological relationship in cases with dubious or unknown links. In this context, phylogeny has been recurrently used as a forensic tool in HIV transmission investigations. Phylogeny permit to identify groups of genetic isolates that form genetically related clades with strains closely related. The application of phylogenetic methods was so extended in the context of HIV-specific criminal laws that precise rolls were defined (166, 167). Lemey et al. (168) applied phylogeny to a possible HIV- 1 transmission case, in which 6 females presumably became HIV infected after a sexual assault from an African suspect. The conclusions of this work highlighted that viral strains from the victims were more closely related to the virus carried by the suspect with respect to used controls, but did not exclude the possibility of the presence of an additional subject that could have infected both suspect and victims. In another case (169), the results of phylogenic analyses resulted to be more definitive permitting the liberation the foreign medical staff who was accused of transmitting the HIV strain to children attending Al-Fateh Hospital in Benghazi in Libya. The authors demonstrated that the HIV-1 and HCV strains related to young victims were already circulating and prevalent in this hospital and its environs before the arrival in March 1998 of suspects.

Considering the very rapid HIV evolution and the fact that molecular sequences taken at different points in time showed a statistically significant number of genetic differences; HIV-1 evolution was also

studied in the microcosm represented of viral sequences sampled longitudinally i.e. heterochronous sequences obtained from an infected host. The main results from these studies showed a correlation between the intra-host evolutionary rate disease progression, linear or inverse (170, 171). The founding of supported monophyletic clades including sequences isolated from a specific tissue supported the existence of diverse HIV-1 subpopulations infecting different body tissues (172, 173).

The comparison of HIV-1 intra-host genealogies from patients sampled under different conditions could help in understanding the viral reservoir dynamics, the most important knowledge for the HIV-1 vaccine design.

Phylogenetic and phylodinamic analysis contributed to consistent improvement in the evolutionary reconstruction of HIV virus history.

Phylogenesis and phylodinamic of Arboviruses

Phylogenetic and population genetic inference (phylodinamics) based on viral whole genome data can resolve putative outbreaks, investigate their aetiology, and provide spatiotemporal context during investigations (174).

Many of the important zoonotic arboviruses belong to the families Togaviridae (Chikungunya virus, CHIKV) and Flaviviridae (Zika Virus, ZIKV; Yellow Fever virus, YFV; Dengue virus, DENV) (175). The main vector able to spread zoonotic viruses are the mosquitoes, especially (Culex spp., Aedes spp., etc.). These may act as vectors for the same virus in different vertebrate hosts, depending on different geographical and ecological locations.

Phylogenetic analyses can reveal a better picture of arboviral outbreak, as have been occurred in Brazil in CHIKV outbreak. Phylogenetic reconstruction suggests several separate introductions of the Asian genotype (ECSA) strain, in contrast to a unique introduction of the ECSA genotype followed by virus dissemination inside the country (176, 177) also in Dengue virus outbreaks (178). In CHIKV Italian outbreak in September 2017, phylodinamic analysis showed that the virus was in loco at least 3 months before the outbreak and in particular it probably originated in Pakistan and not India as the classical epidemiology supposed (179).

As well as in ZIKV outbreaks, phylogenetic inference showed how the outbreak in Brazil was closely related to the French Polynesia isolates that circulated in November 2013 (35, 180) and that all viruses sampled in the Americas, including those from Brazil, form a robust monophyletic cluster the Asian genotype. Meanwhile in United States, ZIKV outbreaks appeared to had multiples introductions linked to the Caribbean (181)

The application of Bayesian statistical inference framework, can allow the reconstruction of the geographic history of the still ongoing and never reported before epidemic on the basis of the first isolates sampled at known times (182). It has been showed that CHIKV was imported to the Americas from Southeast Asia (183-187).

During CHIKV outbreak in 2005-2006, non-synonymous mutation (A226V) occurred both on La Reunion and in some Indian area (188). The phylogenetic analysis highlighted how the sporadic finding of A226V, that allowed a vector jump from Aedes aegypti to ae. albopictus, was possible probably due to the importation of mutated strains from Indian Ocean islands to the Indian subcontinent. Moreover, the analysis showed the likely eastward path from Africa to Indian Ocean Island to India, and from there to other South East Asian countries (188).

The phylogenetic analyses allowed to understand the epidemiology history and how the viruses spread in the infected countries and inside the country with the possibility to plan prevention strategies (181, 189, 190)

Phylogeny of human polyomaviruses

Polyomaviruses belong the *Polyomaviridae* family that according to the International Committee on Taxonomy of Viruses (ICTV) includes four genera named *alpha-, beta-, gamma- and delta-polyomavirus* that together include 83 species (https://talk.ictvonline.org/taxonomy). This taxonomic classification is based on the phylogenetic analysis of the Large T Antigen (LT Ag) coding sequences.

Up to 2007, two human polyomaviruses were known: BKPyV and JCPyV both uncovered in 1971 and named after the initials of the patients where they were first isolated (191, 192). According to Jin et al., four serotypes of BKPyV exist based on the differences between amino acids 61-83 in the VP1 region. These four serotypes corresponded to the four subtypes characterized by sequencing or restriction fragment length polymorphisms (193, 194).

Subtype I has a worldwide distribution, subtype IV is present in Asia and part of Europe, and subtypes II and III are rarely detected. Phylogenetic analysis of full- length genomes revealed the presence of subgroups within subtypes I (Ia, Ib1, Ib2 and Ic) and IV (IV a1, IV a2, IVb1, IVb2, IVc1, IVc2) (195, 196).

JCPyV, the etiological agent of Progressive Multifocal Leukoencephalopathy (PML) usually observed in immunocompromised patients, especially HIV positive, is classified into a major VP1 serotype and at least seven major genotypes by phylogenetic analysis. These genotypes present a characteristic geographic distribution (197, 198) that allowed us to trace the human migrations from the African continent where human beings first appeared to the rest of the world. JCPyV type 1 predominates

between Europeans and European-Americans, types 2 and 7 are found in Asia, and types 3 and 6 in Africa. Genotype 4 found between European-Americans in USA is closely related to the genotype 1 from which it differs of about 1% at DNA level (199).

In 2007 the discovery of two new human polyomaviruses named KIPyV and WUPyV isolated from the respiratory secretions of patients affected by acute respiratory tract infection was reported (200, 201). At phylogenetic analysis analysis, these two viruses formed a new subclass of polyomaviruses. KIPyV was phylogenetically related to other primate polyomaviruses in the early region of the genome, but had a little homology, <30% amino acid identity, in the late region of the genome when compared to known polyomaviruses (200). Phylogenetic analysis clearly revealed that WUPyV was a novel virus closely related to KIPyV. In the early VP1region, WU/KI was most related to the primate polyomaviruses BK, SV40, JC, and baboon polyomavirus. On the other hand, analysis of the VP2 open reading frame showed that WUPyV was very much divergent from other polyomaviruses aside from KIPyV (201). In the following years, phylogenetic analysis of whole genome sequences evidenced three WUPyV genotypes and five subtypes: Ia, Ib, Ic, II, IIIa and IIIb. No association was noted between genotype and distinct clinical features (202).

In 2008, a fifth novel human polyomavirus was discovered. This novel human polyomavirus was named Merkel cell polyomavirus (MCPyV) (203) belonging to the genus *Alphapolyomavirus* was described as causing agent of Merkel cell carcinoma a rare and aggressive skin cancer usually affecting the elderly and immunocompromised individuals (204). Neighbor-joining trees for putative MCPyV LT, sT, VP1, and VP2 proteins revealed that the four known human polyomaviruses BKPyV, JCPyV, KIPyV, and WUPyV clustered together in the SV40 subgroup, whereas MCPyV was most closely related to MuPyV subgroup viruses and African green monkey (AGM) lymphotropic polyomavirus (LPyV) (203).

Phylogenetic analysis, performed by the NJ method and using MCPyV complete genome sequences indicated the presence of two clades highly supported by high bootstrap values. These clades were provisionaly named by the authors as A for MCPyV strains isolated from individuals of Asian background and C for those strained isolated from patients of Caucasian background (205-207). Other authors have more recently confirmed the exsistence of an Asian genotype (208). These data suggest a geographic distribution of MCPyV genotypes similarly to BKPyV and JCPyV viruses (209).

Following MCPyV discovery, two other human polyomaviruses were identified in the skin of healthy individuals named HPyV6 and HPyV7 (207). HPVy7 has been associated with a pruritic rash in two lung transplant recipients (210). HPyV6 has been linked to Kimura disease, a chronic inflammatory disorder with subcutaneous nodules that is endemic in East Asia (211). Phylogenetic analysis performed

on the LT gene sequences revealed the presence of two distinct clades. All the HPyV strains originated from the Japanese patients belonged to the Asian/Japanese clade indicating an Asian/Japanese genotype (252).

The trichodysplasia spinulosa-associated polyomavirus (TSPyV) genome was amplified from in plucked facial spines of a heart transplant patient with trichodysplasia spinulosa (212). Phylogenetic analysis revealed a close relationship of TSPyV with the Bornean orangutan polyomavirus and, more distantly, the Merkel cell polyomavirus; the causative agent of Merkel Cell Carcinoma (MCC).

. Maximum likelihood phylogenetic tree analysis of the VP1, VP2 and LTAg proteins revealed that MWPyV was highly divergent from all known polyomaviruses. Based on VP2 and LTAg sequences, MWPyV clustered with the clade containing HPyV9, LPyV, HaPyV, MPyV, TSPyV, MCV, ChPyV and the orangutan polyomaviruses. These discordant phylogenetic relationships suggest that MWPyV is probably derived from an ancestral recombination event (213). Currently, MWPyV has not been associated to any specific human disease.

STLPyV was uncovered screening the fecal microbiota of a child living in Malawi (214). This novel polyomavirus is closely related to MWPyV but phylogenetic analysis revealed that MWPyV and STLPyV form distinct clades with high confidence.

HPyV12 is novel human polyomavirus originally identified in a resected liver tissue by a generic PCR targeting the VP1 gene. Phylogenetic analysis showed that HPyV12 was not closely related to any known human or animal polyomavirus (215).

NJPyV-2013, is the last novel human polyomavirus identified in a young pancreatic transplant recipient with retinal blindness and vascular myopathy (216). Phylogenetic analysis performed on whole genome sequences, showed that NJPyV-2013 is most closely related to chimpanzee polyomavirus. Phylogenetic and phylodinamic analysis contributed to a further improvement in the evolutionary reconstruction of Polyoma virus history causing human infections.

Role of phylogenetic analysis in vaccine strategies

One of the most effective ways to prevent infectious diseases is to use vaccines that stimulate a protective immune response upon subsequent contact with the pathogen.

An effective vaccination strategy is strictly connected to the knowledge of the microbial origin, spread and evolution, especially whenever a novel pathogen emerges, and predicting the course of evolution is one of the most challenging and important areas in infectious diseases (217). Availability of microbe species genome sequences and application of system biology to the field of vaccinology, have provided new ways for identification and evaluation of potential vaccine candidates. Phylogenetic analysis is one

of the best tools for studying the adaptive evolution of pathogens, and it is used to investigate outbreaks or endemic diseases and the history of the pandemics-including the pattern (when, where, how) of disease spreading throughout the world (218, 219).

Despite the wide use of vaccines, vaccination failure is possible due to the emergence of variants that escape vaccination. Thus, it is pivotal to determine the genetic diversity between circulating field strains and the currently used vaccine strains. Indeed, the change of genetic material is one of the major mechanisms that pathogens use to escape the host immune response, and that leads to outbreaks of infectious disease (220). Vaccine development deals with the interplay of fitness, diversity, dynamics and virulence, which are the most important factors of viral evolution, and phylogenetic analysis is a key tool to evaluate all together these factors.

Here, some examples in which phylogenetic analysis has improved design and development of vaccines have been provided.

We consider first RNA viruses, like influenza and HIV, which pose grave risk to public health worldwide. These fast evolving viruses continuously adapt to the highly variable environments they live, by evading host immune responses and altering the severity of disease (221). Influenza viruses, that infects about one fifth of the worldwide human population each year, are characterized by a very rapid evolution, resulting from mutations or genetic re-assortment with other of influenza viruses, such as pig horses and birds (222). This is why there is a new vaccine each year (223). Since 90's it was clearly shown by phylogeny that, strains with amino acid changes at hemagglutinin sites tended to become the dominant strain next year (224). This permanent evolution necessitates of vaccine composition updates, based on the selection of strains that antigenically match currently circulating influenza viruses (225). Thus, use of phylogenetic methods to influenza sequence data, has become an integral part of the yearly vaccine design cycle. To date, computational analysis-assisted vaccine design established for influenza, has not been equally successful for HIV. While every years, influenza strains diverge by about 1 to 2 percent in the population (226), in comparison HIV strains, diverge about one percent per year within a single individual. Phylogenetic analysis showed that the genetic structure of the HIV epidemic tends to spread out from an ancestor in a radial fashion to generate high variation (227). The evolutionary dynamics of HIV strains involve multiple processes as high replication rates and recombination, selective pressure exerted by the host immune system and by drugs, and genetic drift (228). Thus far, designing a vaccine to cover all, or fraction of HIV diversity has appeared to be impossible (229). The high evolutionary rate of HIV makes the needed progress in virus vaccine strategies, which are a critical component for controlling the HIV epidemic. In this regard, new vaccine approaches are coming by use

of computational vaccine design. For example, to conceive computer-generated artificial proteins, that potentially may stimulate a broader immune response than natural HIV proteins, phylogenetics and computational techniques are used. Researchers make a “consensus” sequences, including the most common amino acid at each position from a broad number of sequences HIV’s variation, to create proteins that will serve as immunogens in vaccines. This methodology, that uses bioinformatics tools for vaccine development starting from genomic information of the pathogen, is called reverse vaccinology. However, until today, this approach has failed to deliver an effective, preventive HIV-1 vaccine (230).

Another rapidly evolving virus, which is a major public health problem, is Hepatitis B virus (HBV) (231). HBV infection is caused by a partially double-stranded reverse-transcribed DNA virus, with genetic variability related to high levels of virus production, and absence of viral polymerase proofreading activity during the reverse transcription step of the replication cycle. Because of genomic structure and unusual replication cycle, it is difficult to estimate exactly rate of virus evolution. HBV is more comparable to an RNA virus that use reverse transcriptase than a DNA virus, with a mutation rate (0,0005 substitutions per site per year) as compared with HIV (0,003 substitutions per site per year). Phylogenetic analysis has revealed ten human genotypes of HBV, from A to J, with a sequence divergence larger than 8% (232). Pyogenesis of novel variants, generated by recombination events between different HBV genotypes, have documented their presence worldwide, confirming their potential for spreading in a wide range of human populations and developing their own epidemiology (233). For example, in low endemicity countries with high immigration rates, phylogenetic analyses provided essential information on the molecular epidemiology and population dynamics of HBV. Current HBV vaccines contain correctly folded HBsAg and neutralizing epitopes of the S antigen, induce rapid protection and overcome nonresponse to second-generation vaccines. However, recombination, evasion of host immunity, vaccine escape and resistance to drugs have been associated with evolutionary aspects of HBV genetic variability. Mutations of the determinant "a" in S antigen, generating variants that "escape" from vaccination, called escape mutants, and that induce infection also in subjects who have protective titers of anti-HBs, have been described (234). With an ever better vaccine formulation and vaccination coverage, eradication of HBV would be possible.

Varicella–zoster virus (VZV) is the only herpesvirus for which a vaccine has been developed. Human herpes viruses present low natural mutation rate, and stability of VZV genome is likely related in part to the small number of replication cycles during primary infection (235). To date, five major clades confirmed by full-genome sequencing have been established for VZV. It is thought that recombination to have played a crucial role in the evolution of VZV. Indeed, a point of phylogenetic interest for VZV,

is that vaccine in use for many years, may have significantly altered the natural epidemiology of wild-type virus in the human populations (236). Although VZV wild-type recombinants have been rarely reported (237) recombination might occur presumably also between wild-type and vaccine strains. Thus, surveillance of VZV molecular epidemiology is indicated to detect changes in strain prevalence where VZV vaccine is in use.

Integration of mathematical evolutionary modeling with experimental data will produce important information on protein sequence and structure analysis to make a prevision on the effect that selective pressure could have on aminoacidic residue essential for vaccine development or drugs activity (Figure 5). play an increasingly important role in the development of new vaccines and the control of infectious disease.

Phylogenetic analysis of whole genome sequencing for viral nosocomial outbreak investigation.

During last years, the introduction of advanced molecular techniques and of bioinformatics analysis in the routine clinical practice has largely modified the diagnostic approach to infectious disease. By Next Generation Sequencing (NGS) technology, the whole genome sequencing (WGS) of pathogen microorganisms has been introduced improving the ability to identify and control epidemic outbreaks, thus preventing the spread of the pathogen and decreasing morbidity and mortality (240). Phylogenetic analysis applied to whole genome sequences has been used for the epidemiological surveillance of nosocomial infections as recently reported (3, 4, 241)**.**

The application of WGS and phylogenetic analysis for nosocomial investigation has been used prevalently in case of hospital acquired infection caused by bacteria commonly isolated by simple culture of different biological samples (3,4, 241, 242) The use of this new technology in nosocomial viral infection has been more limited but some reports have been published.

A first report on the use of metagenomics NGS for the investigation of a nosocomial outbreak by Parainfluenza Virus 3 was published by Greninger et al in 2016. By Bayesian phylogenetic analysis, these authors demonstrated that two strains causing HAI have the same sequence confirming the hypothesis of a possible transmission within the medical ward where the two patients were admitted. Authors concluded that NGS and WGS represent important tools for pathogen transmission investigation and a significant advance in clinical care providing pathogen detection with high resolution and reliability (243).

The same authors used this technology to rule out a nosocomial outbreak caused by respiratory virus, in children at high risk for pulmonary disease. Applying WGS and phylogenetic analysis, they excluded the possibility of a single source-transmission with a turnaround time of 24 hour (244).

Phylogenetic analysis applied to WGS revealed to be essential to discriminate nosocomial outbreak strains from other strains isolated in the same setting but not related to the outbreak event, as recently demonstrated by Zhu et al. during an epidemiological investigation performed on respiratory syncytial virus-B strains circulating within a Unit of hematology and stem cell transplant. Authors compared the routine sequencing of glycoprotein (G) gene to the WGS and concluded that phylogenetic analysis applied to single gene sequence was not able to identify the cluster outbreak and that at this aim WGS was essential (245).

Recently phylogenetic analysis was used to investigate the measles outbreak of nosocomial origin occurred in Milan, Italy between March-August 2017. This study by phylogenetic inference demonstrated the occurrence of three nosocomial clusters within an emergency department initially notified as unrelated cases, suggesting the importance of measles vaccination especially for the healthcare workers (246).

The importance of pyogenesis and phylodinamic applied to WGS has been further highlighted in case of viral nosocomial infection occurring in immunocompromised patients, such as Norovirus infection causing gastroenteritis worldwide affecting pediatric or immunosuppressed populations. Brown et al. in 2018 by phylogenetic analysis of norovirus WGS showed that the nosocomial setting is characterized by multiple strains introduction and evidenced how this type of analysis is superior to the classic epidemiological investigation for infection control and prevention and should be routinely introduced in the hospital setting (247).

Conclusion

Epidemiology has been and still today is the first way to study viral infection and to face an epidemic event to reach a huge number of information to use in clinic and therapy. In these last years, the question that spring is “May phylogenetic analysis support and or complement an epidemiological investigation?” To answer this question is more complex as appear (Figure 6). Phylogeny and phylodinamic can give an immediate and important support studying the epidemiologic events not by the human point of view but by the microorganism point of view. Data on sequences isolated by the microorganism implied in the epidemic or pandemic events can give information on the origin and the spread of the event so as can track the geographical path of diffusion. In terms of nosocomial infection, can give important information on the source of the infection pinpointing the exact way of diffusion. On the other hand, classical epidemiology has focused on research of the risk factors related to behavioral patterns. When classical epidemiology is integrated by the phylogenetic and phylodinamics approaches

goes into a single analytical framework, called evolutionary epidemiology. This review and the examples discussed within showed the necessity to enrich the classical surveillance by the molecular phylogenetic analysis to better understand the emergence and spread of global infectious diseases in terms of regional epidemic so as in nosocomial outbreak. Despite research in the field of infectious disease phylodinamics is growing up, there is a need to apply molecular phylogenetic and evolutionary methods in areas out of infectious diseases, as well as translational genomics, and personalized medicine. The huge availability of genetic and clinical data give to the researchers a great opportunity to apply this molecular approach to studies of tumors and chronic infections, try to better understand the mechanism to the basis of complex transmission dynamics among tissues and cells. Phylogenetic studies can also give a great hand in vaccine studies and in antibiotic and antiviral resistant the last great battle in these times.

The major improvements from phylogenetic and phylodinamic analysis application in the field of viral infectious disease are represented by the possibility to enrich the classical epidemiology with genetic information thus widening the knowledge on viral epidemic and pathogen transmission.

REFERENCES

1. Fitch WM, Margoliash E. Construction of phylogenetic trees. Science 1967;155:279-84.
2. Lemey P, Rambau A, Drummond AJ, Suchard MA. Bayesian phylogeography finds its roots. PLOS Comput. Biol. 2009; 5:e1000520.
3. Angeletti S, Cella E, Prosperi M, Spoto S, Fogolari M, De Florio L, Antonelli F, Dedej E, De Flora C, Ferraro E, Incalzi RA, Coppola R, Dicuonzo G, Francescato F, Pascarella S, Ciccozzi M. Multi-drug resistant Pseudomonas aeruginosa nosocomial strains: Molecular epidemiology and evolution. Microb Pathog. 2018;123:233-241.
4. Cella E, Ciccozzi M, Lo Presti A, Fogolari M, Azarian T, Prosperi M, Salemi M, Equestre M, Antonelli F, Conti A, Cesaris M, Spoto S, Incalzi RA, Coppola R, Dicuonzo G, Angeletti S.

Multi-drug resistant Klebsiella pneumoniae strains circulating in hospital setting: whole-genome sequencing and Bayesian phylogenetic analysis for outbreak investigations. Sci Rep. 2017;7:3534.

5. Angeletti S, Presti AL, Cella E, Fogolari M, De Florio L, Dedej E, Blasi A,Milano T, Pascarella S, Incalzi RA, Coppola R, Dicuonzo G, Ciccozzi M. Klebsiella pneumoniae blaKPC-3 nosocomial epidemic: Bayesian and evolutionary analysis. Infect Genet Evol. 2016;46:85-93
6. Pybus OG, Rambaut A, Harvey PH. An integrated framework for the inference of viral population history from reconstructed genealogies. Genetics. 2000;155:1429-37
7. Zuckerkandl E, Pauling L. Molecular disease, evolution, and genetic heterogeneity. In Horizons in Biochemistry, di M Kasha e B. Pullman. New York: Academic Press; 1962 189–225 p.
8. Lemey, P, M Salemi, e AM. Vandamme. The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing. New York: Cambridge University Press; 2009.
9. Pol D. Empirical problems of the hierarchical likelihood ratio test for model selection. Syst Biol. 2004;53:949-62.
10. Zehender G, Ebranati E, Lai A, Santoro MM, Alteri C, Giuliani M, Palamara G, Perno CF, Galli M, Lo Presti A, Ciccozzi M. Population dynamics of HIV-1 subtype B in a cohort of men-having-sex-with-men in Rome, Italy. J Acquir Immune Defic Syndr. 2010;55:156-60
11. Callegaro A, Svicher V, Alteri C, Lo Presti A, Valenti D, Goglio A, Salemi M, Cella E, Perno CF, Ciccozzi M, Maggiolo F. Epidemiological network analysis in HIV-1 B infected patients diagnosed in Italy between 2000 and 2008. Infect Genet Evol. 2011;11:624-32.
12. Ciccozzi M, Zehender G, Cento V, Lo Presti A, Teoharov P, Pavlov I, Bogdanova V, Perno CF, Ciotti M. Molecular analysis of hepatitis C virus infection in Bulgarian injecting drug users. J Med Virol. 2011;83:1565-70
13. Ciccozzi M, Equestre M, Costantino A, Marascio N, Quirino A, Lo Presti A, Cella E, Bruni R, Liberto MC, Focà A, Pisani G, Zehender G, Ciccaglione AR. Hepatitis C virus genotype 4d in Southern Italy: reconstruction of its origin and spread by a phylodinamic analysis. J Med Virol. 2012;84:1613-9
14. de Oliveira T, Pybus OG, Rambaut A, Salemi M, Cassol S, Ciccozzi M, Rezza G, Gattinara GC, D'Arrigo R, Amicosante M, Perrin L, Colizzi V, Perno CF; Benghazi Study Group. Molecular epidemiology: HIV-1 and HCV sequences from Libyan outbreak. Nature. 2006;444:836-7.

15. Salemi M, de Oliveira T, Ciccozzi M, Rezza G, Goodenow MM. High-resolution molecular epidemiology and evolutionary history of HIV-1 subtypes in Albania.PLoS One. 2008;3:e1390.
16. Bon I, Ciccozzi M, Zehender G, Biagetti C, Verrucchi G, Lai A, Lo Presti A,Gibellini D, Re MC. HIV-1 subtype C transmission network: the phylogenetic reconstruction strongly supports the epidemiological data. J Clin Virol. 2010;48(3):212-4.
17. Salemi M, Fitch WM, Ciccozzi M, Ruiz-Alvarez MJ, Rezza G, Lewis MJ. Severe acute respiratory syndrome coronavirus sequence characteristics and evolutionary rate estimate from maximum likelihood analysis. J Virol. 2004;78:1602-3
18. Min J, Cella E, Ciccozzi M, Pelosi A, Salemi M, Prosperi M. The global spread of Middle East respiratory syndrome: an analysis fusing traditional epidemiological tracing and molecular phylodinamics. Glob Health Res Policy. 2016;1:14.
19. Beau De Rochars VM, Lednicky J, White S, Loeb J, Elbadry MA, Telisma T, Chavannes S, Anilis MG, Cella E, Ciccozzi M, Okech BA, Salemi M, Morris JG Jr. Isolation of Coronavirus NL63 from Blood from Children in Rural Haiti: Phylogenetic Similarities with Recent Isolates from Malaysia. Am J Trop Med Hyg. 2017;96(1):144-147.
20. Moreno A, Lelli D, de Sabato L, Zaccaria G, Boni A, Sozzi E, Prosperi A, Lavazza A, Cella E, Castrucci MR, Ciccozzi M, Vaccari G. Detection and full genome characterization of two beta CoV viruses related to Middle East respiratory syndrome from bats in Italy. Virol J. 2018 Jan 12;15(1):10.
21. Babakir-Mina M, Ciccozzi M, Ciotti M, Marcuccilli F, Balestra E, Dimonte S, Perno CF, Aquaro S. Phylogenetic analysis of the surface proteins of influenza A (H5N1) viruses isolated in Asian and African populations. New Microbiol. 2009;32:397-403.
22. Ciccozzi M, Montieri S, Facchini M, Rezza G, Donatelli I, Campitelli L. Evolutionary analysis of HA and NS1 genes of H5N1 influenza viruses in 2004-2005 epidemics. Avian Dis. 2007;51:455-60.
23. Yegani S, Shoushtari AH, Eshratabadi F, Molouki A. Full sequence analysis of hemagglutinin and neuraminidase genes and proteins of highly pathogenic avian influenza H5N1 virus detected in Iran, 2015. Trop Anim Health Prod. 2019;51:605-612.
24. Gao J, Couzens L, Burke DF, Wan H, Wilson P, Memoli MJ, Xu X, Harvey R, Wrammert J, Ahmed R, Taubenberger JK, Smith DJ, Fouchier RAM, Eichelberger MC. Antigenic Drift of the Influenza A(H1N1)pdm09 Virus Neuraminidase Results in Reduced Effectiveness of A/California/7/2009 (H1N1pdm09)-Specific Antibodies. MBio. 2019;10: e00307-19.

25. Soli R, Kaabi B, Barhoumi M, Maktouf C, Ahmed SB. Bayesian phylogenetic analysis of the influenza-A virus genomes isolated in Tunisia, and determination of potential recombination events. Mol Phylogenet Evol. 2019;134:253-268.

26. Mugosa B, Vujosevic D, Ciccozzi M, Valli MB, Capobianchi MR, Lo Presti A,Cella E, Giovanetti M, Lai A, Angeletti S, Scarpa F, Terzić D, Vratnica Z.Genetic diversity of the haemagglutinin (HA) of human influenza a (H1N1) virus in montenegro: Focus on its origin and evolution. J Med Virol. 2016;88:1905-13

27. Zehender G, Pariani E, Piralla A, Lai A, Gabanelli E, Ranghiero A, Ebranati E, Amendola A, Campanini G, Rovida F, Ciccozzi M, Galli M, Baldanti F, Zanetti AR. Reconstruction of the evolutionary dynamics of the A(H1N1)pdm09 influenza virus in Italy during the pandemic and post-pandemic phases. PLoS One. 2012;7:e47517.

28. Ciccozzi M, Babakir-Mina M, Lo Presti A, Marcuccilli F, Perno CF, Ciotti M. Phylogenesis and Clinical Aspects of Pandemic 2009 Influenza A (H1N1) Virus Infection. Open Virol J. 2011;5:22-6.

29. Jansen van Vuren P, Ladner JT, Grobbelaar AA, Wiley MR, Lovett S, Allam M, Ismail A, le Roux C, Weyer J, Moolla N, Storm N, Kgaladi J, Sanchez-Lockhart M,Conteh O, Palacios G, Paweska JT. Phylodinamic Analysis of Ebola Virus Disease Transmission in Sierra Leone. Viruses. 2019;11: E71

30. Arias A, Watson SJ, Asogun D, Tobin EA, Lu J, Phan MVT, Jah U, Wadoum REG,Meredith L, Thorne L, Caddy S, et al. Rapid outbreak sequencing of Ebola virus in Sierra Leone identifies transmission chains linked to sporadic cases. Virus Evol. 2016;2:vew016.

31. Giovanetti M, Grifoni A, Lo Presti A, Cella E, Montesano C, Zehender G, Colizzi V, Amicosante M, Ciccozzi M. Amino acid mutations in Ebola virus glycoprotein of the 2014 epidemic. J Med Virol. 2015 Jun;87(6):893-8.

32. Azarian T, Lo Presti A, Giovanetti M, Cella E, Rife B, Lai A, Zehender G,Ciccozzi M, Salemi M. Impact of spatial dispersion, evolution, and selection on Ebola Zaire Virus epidemic waves. Sci Rep. 2015 May 14;5:10170.

33. Ebranati E, Veo C, Carta V, Percivalle E, Rovida F, Frati ER, Amendola A,Ciccozzi M, Tanzi E, Galli M, Baldanti F, Zehender G. Time-scaled phylogeography of complete Zika virus genomes using discrete and continuous space diffusion models. Infect Genet Evol. 2019;73:33-43

34. Iovine NM, Lednicky J, Cherabuddi K, Crooke H, White SK, Loeb JC, Cella E,Ciccozzi M, Salemi M, Morris JG Jr. Coinfection With Zika and Dengue-2 Viruses in a Traveler Returning From Haiti, 2016: Clinical Presentation and Genetic Analysis. Clin Infect Dis. 2017;64:72-75.
35. Angeletti S, Lo Presti A, Giovanetti M, Grifoni A, Amicosante M, Ciotti M,Alcantara LJ, Cella E, Ciccozzi M. Phylogenesys and homology modeling in Zika virus epidemic: food for thought. Pathog Glob Health. 2016;110:269-274.
36. Lednicky J, Beau De Rochars VM, El Badry M, Loeb J, Telisma T, Chavannes S,Anilis G, Cella E, Ciccozzi M, Rashid M, Okech B, Salemi M, Morris JG Jr. Zika Virus Outbreak in Haiti in 2014: Molecular and Clinical Data. PLoS Negl Trop Dis.2016;10:e0004687.
37. Schmid D, Fretz R, Buchner G, Konig C, Perner H, Sollak R, Tratter A, Hell M, Maass M, Strasser M, Allerberger F. Foodborne outbreak of hepatitis A, November 2007-January 2008, Austria. European Journal of Clinical Microbiology & Infectious Diseases 2009; 28:385–391.
38. Harries M, Monazahian M, Wenzel J, Jilg W, Weber M, Ehlers J, Dreesman J, Mertens E. Foodborne hepatitis A outbreak associated with bakery products in northern Germany, 2012. Eurosurveillance 2014; 19:20992.
39. Schenkel K, Bremer V, Grabe C, Van Treeck U, Schreier E, Hohne M, Ammon A, Alpers K. Outbreak of hepatitis A in two federal states of Germany: Bakery products as vehicle of infection. Epidemiology & Infection 2006; 134:1292–1298.
40. Wenzel JJ, Allerberger F. Hepatitis A as a foodborne infection. Lancet Infectious Diseases 2014; 14:907–908.
41. Bialek SR, George PA, Xia GL, Glatzer MB, Motes ML, Veazey JE, Hammond RM, Jones T, Shieh YC, Wamnes J, Vaughan G, Khudyakov Y, Fiore AE. Use of molecular epidemiology to confirm a multistate outbreak of hepatitis A caused by consumption of oysters. Clinical Infectious Diseases 2007; 44: 838–840.
42. Shieh YC, Khudyakov YE, Xia G, Ganova-Raeva LM, Khambaty FM, Woods JW, Veazey JE, Motes ML, Glatzer MB, Bialek SR, Fiore AE. Molecular confirmation of oysters as the vector for hepatitis A in a 2005 multistate outbreak. Journal of Food Protection 2007; 70:145–150.
43. Vaughan G, Xia G, Forbi JC, Purdy MA, Rossi LM, Spradling PR, Khudyakov YE. Genetic relatedness among hepatitis A virus strains associated with food-borne outbreaks. PLoS One. 2013;8:e74546

44. Lopalco PL, Malfait P, Menniti-Ippolito F, Prato R, Germinario C, Chironna M, Quarto M, Salmaso S. Determinants of acquiring hepatitis A virus disease in a large Italian region in endemic and epidemic periods. J Viral Hepat. 2005;12:315-21

45. D'Amelio R, Mele A, Mariano A, Romanò L, Biselli R, Lista F, Zanetti A, Stroffolini T. Hepatitis A, Italy. Emerg Infect Dis. 2005;11:1155-6.

46. Amon JJ, Devasia R, Xia G, Nainan OV, Hall S, Lawson B, Wolthuis JS, Macdonald PD, Shepard CW, Williams IT, Armstrong GL, Gabel JA, Erwin P, Sheeler L, Kuhnert W, Patel P, Vaughan G, Weltman A, Craig AS, Bell BP, Fiore A. Molecular epidemiology of foodborne hepatitis a outbreaks in the United States, 2003. J Infect Dis. 2005;192:1323-30.

47. Wheeler C, Vogt TM, Armstrong GL, Vaughan G, Weltman A, Nainan OV, Dato V, Xia G, Waller K, Amon J, Lee TM, Highbaugh-Battle A, Hembree C, Evenson S, Ruta MA, Williams IT, Fiore AE, Bell BP. An outbreak of hepatitis A associated with green onions. N Engl J Med. 2005;353:890-7.

48. Carvalho C, Thomas H, Balogun K, Tedder R, Pebody R, Ramsay M, Ngui S. A possible outbreak of hepatitis A associated with semi-dried tomatoes, England, July-November 2011. Euro Surveill. 2012;17:20083

49. Poovorawan Y, Theamboonlers A, Chongsrisawat V, Jantaradsamee P, Chutsirimongkol S, Tangkijvanich P. Clinical features and molecular characterization of hepatitis A virus outbreak in a child care center in Thailand. J Clin Virol. 2005;32:24-8

50. Pelletier AR, Mehta PJ, Burgess DR, Bondeson LM, Carson PJ, Rea VE, Sharapov UM, Hu DJ. An outbreak of hepatitis A among primary and secondary contacts of an international adoptee. Public Health Rep. 2010;125:642-6.

51. Perevoscikovs J, Lucenko I, Magone S, Brila A, Curikova J, Vennema H. Community-wide outbreak of hepatitis A in Latvia in 2008 - an update. Euro Surveill. 2009;14: 19092.

52. Rodrigues L, Pista A, Oliveira A, Agua-Doce I, Manita C, Paixão MT. Molecular epidemiology of hepatitis A virus in a group of Portuguese citizens living in Lisbon area. J Med Virol. 2007;79:483-7.

53. McFarland N, Dryden M, Ramsay M, Tedder RS, Ngui SL; 2008 Winchester HAV Outbreak Team. An outbreak of hepatitis A affecting a nursery school and a primary school. Epidemiol Infect. 2011;139:336-43

54. de Paula VS, Baptista ML, Lampe E, Niel C, Gaspar AM. Characterization of hepatitis A virus isolates from subgenotypes IA and IB in Rio de Janeiro, Brazil. J Med Virol. 2002;66:22-7

55. Amado LA, Villar LM, de Paula VS, Pinto MA, Gaspar AM. Exposure to multiple subgenotypes of hepatitis A virus during an outbreak using matched serum and saliva specimens. J Med Virol. 2011 May;83(5):768-75.

56. de Paula VS, Saback FL, Gaspar AM, Niel C. Mixed infection of a child care provider with hepatitis A virus isolates from subgenotypes IA and IB revealed by heteroduplex mobility assay. J Virol Methods. 2003;107:223-8

57. Morais LM, de Paula VS, Arantes MR, Oliveira ML, Gaspar AM. Early infection and asymptomatic spread of hepatitis A virus in a public child care center in Rio de Janeiro, Brazil: should attending children under two years of age be vaccinated? Mem Inst Oswaldo Cruz. 2006;101:401-5

58. Villar LM, Morais LM, Aloise R, Melo MM, Calado IA, Lampe E, Gaspar AM.Co-circulation of genotypes IA and IB of hepatitis A virus in Northeast Brazil. Braz J Med Biol Res. 2006 Jul;39(7):873-81.

59. Villar LM, Morais LM. GL, Han XH, Williams I, Bell BP, Margolis HS. Hepatitis A molecular epidemiology in the United States, 1996–1997: sources of infection and implications of vaccination policy. J Infect Dis 2005;191:957–63.

60. Hofmeister MG, Foster MA, Teshale EH. Epidemiology and transmission of hepatitis A virus and hepatitis E virus infections in the United States. Cold Spring Harb Perspect Med 2018;a033431.

61. Franco E, Giambi C, Ialacci R, Coppola RC, Zanetti AR. Risk groups for hepatitis A virus infection. Vaccine 2003;21:2224–33

62. Urbanus AT, van Houdt R, van de Laar TJ, Coutinho RA. Viral hepatitis among men who have sex with men, epidemiology and public health consequences. EuroSurveill 2009;14:19421

63. Cotter SM, Sansom S, Long T, Koch E, Kellerman S, Smith F, Averhoff F, Bell BP. Outbreak of hepatitis A among men who have sex with men: implications for hepatitis A vaccination strategies. J Infect Dis. 2003;187:1235-40

64. Takahashi H, Yotsuyanagi H, Yasuda K, Koibuchi T, Suzuki M, Kato T, Nakamura T, Iwamoto A, Nishioka K, Iino S, Koike K, Itoh F. Molecular epidemiology of hepatitis A virus in metropolitan areas in Japan. J Gastroenterol. 2006;41:981-6.

65. Bondarenko TIu, Ternovoĭ VA, Netesov SV. The hepatitis A virus: structural and functional organization of the genome, its molecular diagnostic value and cultivation. Mol Gen Mikrobiol Virusol. 2013;(3):12-21.

66. Bordi L, Rozera G, Scognamiglio P, Minosse C, Loffredo M, Antinori A, Narciso P, Ippolito G, Girardi E, Capobianchi MR; GEAS Group. Monophyletic outbreak of Hepatitis A involving HIV-infected men who have sex with men, Rome, Italy 2008-2009. J Clin Virol. 2012;54:26-9

67. Pérez-Sautu U, Costafreda MI, Lite J, Sala R, Barrabeig I, Bosch A, Pintó RM. Molecular epidemiology of hepatitis A virus infections in Catalonia, Spain, 2005-2009: circulation of newly emerging strains. J Clin Virol. 2011;52:98-102

68. van Steenbergen JE, Tjon G, van den Hoek A, Koek A, Coutinho RA, Bruisten SM. Two years' prospective collection of molecular and epidemiological data shows limited spread of hepatitis A virus outside risk groups in Amsterdam, 2000-2002. J Infect Dis. 2004;189:471-82.

69. Lanini S, Minosse C, Vairo F, Garbuglia A, Di Bari V, Agresta A, Rezza G, Puro V, Pendenza A, Loffredo MR, Scognamiglio P, Zumla A, Panella V, Ippolito G,Capobianchi MR; Gruppo Laziale Sorveglianza Epatiti Virali (GLaSEV). A large ongoing outbreak of hepatitis A predominantly affecting young males in Lazio, Italy; August 2016 - March 2017. PLoS One. 2017 Nov 10;12(11):e0185428

70. Nainan OV, Armstrong GL, Han XH, Williams I, Bell BP, Margolis HS. Hepatitis a molecular epidemiology in the United States, 1996-1997: sources of infection and implications of vaccination policy. J Infect Dis. 2005 Mar 15;191(6):957-63

71. Harkess J, Gildon B, Istre G. Outbreaks of hepatitis A among illicit drug users, Oklahoma, 1984–87. Am J Publ Hlth 1989; 79: 463–6.

72. Grinde B, Stene-Johansen K, Sharma B, Hoel T, Jensenius M, Skaug K. Characterisation of an epidemic of hepatitis A virus involving intravenous drug abusers – infection by needle sharing? J Med Virol 1997; 5369–75.

73. Jensenius M, Espinoza R, Hoel T, Oktedalen O, Heger B, Skar A, et al. An outbreak of hepatitis A among intravenous drug addicts in Oslo 1995–96. Tidsskrift Den Norske Laegforening 1997; 117: 935–40. 14.

74. Stene-Johansen K, Jonassen TØ, Skaug K. Characterization and genetic variability of Hepatitis A virus genotype IIIA. J Gen Virol. 2005;86:2739-45

75. O'Donovan D, Cooke RP, Joce R, Eastbury A, Waite J, Stene-Johansen K. An outbreak of hepatitis A amongst injecting drug users. Epidemiol Infect. 2001;127:469-73

76. Zehender G, De Maddalena C, Giambelli C, Milazzo L, Schiavini M, Bruno R, Tanzi E, Galli M. Different evolutionary rates and epidemic growth of hepatitis B virus genotypes A and D. Virology 2008;380:84-90.

77. Holmes EC. Evolutionary history and phyleogeography of human virus. Annu Rev Microbiol 2008;62:307-328.

78. Kramvis A, Weitzmann L, Owiredu WK, Kew MC. Analysis of the complete genome of subgroup A' hepatitis B virus isolates from South Africa. J Gen Virol 2002; 83:835-839

79. Torres C, Pineiro y Leone FG, Pezzano SC, Mbayed VA, Campos RH. New perspectives on the evolutionary history of hepatitis B virus genotype F. Mol Phylogenet Evol 2011;59:114-122.

80. Vrancken B, Suchard MA, Lemey P. Accurate quantification of within- and between-host HBV evolutionary rates requires explicit transmission chain modelling. Virus Evol. 2017;3:vex028

81. Michitaka K, Tanaka Y, Horiike N, Duong TN, Chen Y, Matsuura K, Hiasa Y, Mizokami M, Onji M. Tracing the history of hepatitis B virus genotype D in western Japan. Journal of medical virology 2006; 78 :44-52.

82. Zehender G, Ebranati E, Gabanelli E, Shkjezi R, Lai A, Sorrentino C, Lo Presti A, Basho M, Bruno R, Tanzi E, Bino S, Ciccozzi M, Galli M. Spatial and temporal dynamics of hepatitis B virus D genotype in Europe and the Mediterranean Basin. PLoS One 2012;7 (5):e37198.

83. Zehender G, Shkjezi R, Ebranati E, Gabanelli E, Abazaj Z, Tanzi E, Kraja D, Bino S, Ciccozzi M, Galli M. Reconstruction of the epidemic history of hepatitis B virus genotype D in Albania. Infection, genetics and evolution 2012; 12 (2):291-298.

84. Mortimer PP. Arsphenamine jaundice and the recognition of instrument-borne virus infection. Genitourin Med 1995;71 (2):109-119

85. Hannoun C, Soderstrom A, Norkrans G, Lindh M. Phylogeny of African complete genomes reveals a West African genotype A subtype of hepatitis B virus and relatedness between Somali and Asian A1 sequences. J Gen Virol 2005;86 :2163-2167.

86. van Houdt R, Bruisten SM, Geskus RB, Bakker M, Wolthers KC, Prins M, Coutinho RA. Ongoing transmission of a single hepatitis B virus strain among men having sex with men in Amsterdam. Journal of viral hepatitis 2010;17 (2):108-114.

87. Hahne S, van Houdt R, Koedijk F, van Ballegooijen M, Cremer J, Bruisten S, Coutinho R, Boot H. Selective hepatitis B virus vaccination has reduced hepatitis B virus transmission in the Netherlands. PLoS One 2013;8 (7):e67866.

88. Rantala M, van de Laar MJ. Surveillance and epidemiology of hepatitis B and C in Europe - a review. Euro Surveill 2008;13: 18880.

89. Kramvis A, Paraskevis D. Subgenotype A1 of HBV--tracing human migrations in and out of Africa. Antiviral therapy 2013;18:513-521

90. Zehender G, Svicher V, Gabanelli E, Ebranati E, Veo C, Lo Presti A, Cella E, Giovanetti M, Bussini L, Salpini R, Alteri C, Lai A, Tanzi E, Perno CF, Galli M, Ciccozzi M. Reliable timescale inference of HBV genotype A origin and phylodinamics. Infect Genet Evol. 2015;32:361-9.

91. Mulders MN, Venard V, Njayou M, Edorh AP, Bola Oyefolu AO, Kehinde MO, Muyembe Tamfum JJ, Nebie YK, Maiga I, Ammerlaan W, Fack F, Omilabu SA, Le Faou A, Muller CP. Low genetic diversity despite hyperendemicity of hepatitis B virus genotype E throughout West Africa. The Journal of infectious diseases 2004;190 (2):400-408.

92. Andernach IE, Hubschen JM, Muller CP. Hepatitis B virus: the genotype E puzzle. Rev Med Virol 2009; 19 (4):231-240.

93. Forbi JC, Vaughan G, Purdy MA, Campo DS, Xia GL, Ganova-Raeva LM, Ramachandran S, Thai H, Khudyakov YE. Epidemic history and evolutionary dynamics of hepatitis B virus infection in two remote communities in rural Nigeria. PLoS One 2010; 5 (7):e11615.

94. Sagnelli C, Ciccozzi M, Coppola N, Minichini C, Lo Presti A, Starace M, Alessio L, Macera M, Cella E, Gualdieri L, Caprio N, Pasquale G, Sagnelli E. Molecular diversity in irregular or refugee immigrant patients with HBV-genotype-E infection living in the metropolitan area of Naples. J Med Virol. 2017;89(6):1015-1024.

95. Arauz-Ruiz P, Norder H, Robertson BH, Magnius LO. Genotype H: a new Amerindian genotype of hepatitis B virus revealed in Central America. J Gen Virol 2002; 83 (Pt 8):2059-2073

96. Alvarado-Mora MV, Pinho JR. Distribution of HBV genotypes in Latin America. Antiviral therapy 2013; 18:459-465.

97. Godoy BA, Alvarado-Mora MV, Gomes-Gouvea MS, Pinho JR, Fagundes N, Jr. Origin of HBV and its arrival in the Americas--the importance of natural selection on time estimates. Antiviral therapy 2013; 18 :505-512.

98. Paraskevis D, Magiorkinis G, Magiorkinis E, Ho SY, Belshaw R, Allain JP, Hatzakis A. Dating the origin and dispersal of hepatitis B virus infection in humans and primates. Hepatology (Baltimore, Md) 2013;57:908-916.

99. Kahila Bar-Gal G, Kim MJ, Klein A, Shin DH, Oh CS, Kim JW, Kim TH, Kim SB, Grant PR, Pappo O, Spigelman M, Shouval D. Tracing hepatitis B virus to the 16th century in a Korean mummy. Hepatology (Baltimore, Md) 2012;56 (5):1671-1680.

100. Tran TT, Trinh TN, Abe K. New complex recombinant genotype of hepatitis B virus identified in Vietnam. J Virol. 2008;82(11):5657-63.

101. Tatematsu K, Tanaka Y, Kurbanov F, Sugauchi F, Mano S, Maeshiro T, Nakayoshi T, Wakuta M, Miyakawa Y, Mizokami M. A genetic variant of hepatitis B virus divergent from known human and ape genotypes isolated from a Japanese patient and provisionally assigned to new genotype J. J Virol. 2009;83(20):10538-47.

102. Arankalle VA, Gandhe SS, Borkakoty BJ, Walimbe AM, Biswas D, Mahanta J. A novel HBV recombinant (genotype I) similar to Vietnam/Laos in a primitive tribe in eastern India. J Viral Hepat. 2010;17(7):501-10.

103. Ashfaq UA, Javed T, Rehman S, Nawaz Z, Riazuddin S. An overview of HCV molecular biology, replication and immune responses. Virology Journal 2011, 8; 161.

104. Kapoor A, Simmonds P, Gerold G, Qaisar N, Jain K, Henriquez JA et al. Characterization of a canine homolog of hepatitis C virus. Proc Natl Acad Sci USA. 2011, 12; 108 (28): 11608–13.

105. Stroffolini T, Sagnelli E, Sagnelli C, Smedile A, Furlan C, Morisco F, Coppola N, Andriulli A, Almasio PL. The burden of HBV infection in HCV patients in Italy and the risk of reactivation under DAA therapy. Dig Liver Dis. 2019;51(3):434-437.

106. Orland JR, Wright TL, Cooper S. Acute hepatitis C. Hepatology. 2001;33:321–7.

107. Blackard JT, Shata MT, Shire NJ, Sherman KE. Acute hepatitis C virus infection: a chronic problem. Hepatology. 2008;47:321–31.

108. Jacka B, Applegate T, Krajden M, Olmstead A, Harrigan PR, Marshall BDL, et al. Phylogenetic clustering of hepatitis C virus among people who inject drugs in Vancouver, Canada. Hepatology (Baltimore, Md). 2014; 60(5):1571–80.

109. van de Laar T, Pybus O, Bruisten S, Brown D, Nelson M, Bhagani S, et al. Evidence of a large, international network of international hepatitis C virus transmission in HIV-positive men who have sex with men. Gastroenterology. 2009; 136(5):1609–17.

110. Zhou S, Cella E, Zhou W, Kong WH, Liu MQ, Liu PL, Ciccozzi M, Salemi M, Chen X. Population dynamics of hepatitis C virus subtypes in injecting drug users on methadone maintenance treatment in China associated with economic and health reform. J Viral Hepat. 2017;24(7):551-560

111. Hagan H, Pouget ER, Williams IT, Garfein RL, Strathdee SA, Hudson SM, Latka MH, Ouellet LJ. Attribution of Hepatitis C Virus Seroconversion Risk in Young Injection Drug Users in 5 US Cities. Journal of Infectious Diseases. 2010; 201(3):378–85

112. Nelson PK, Mathers BM, Cowie B, Hagan H, Des Jarlais D, Horyniak D, et al. Global epidemiology of hepatitis B and hepatitis C in people who inject drugs: results of systematic reviews. The Lancet. 2011; 378(9791):571–83.

113. Zibbell JE, Iqbal K, Patel RC, Suryaprasad A, Sanders KJ, Moore-Moravian L, et al. Increases in Hepatitis C Virus Infection Related to Injection Drug Use Among Persons Aged ≤30 Years — Kentucky, Tennessee, Virginia, and West Virginia, 2006–2012. MMWR Morb Mortal Wkly Rep. 2015; 64(17):453–8.

114. Gededzha MP, Selabe SG, Kyaw T, Rakgole JN, Blackard JT, Mphahlele MJ. Introduction of new subtypes and variants of hepatitis C virus genotype 4 in South Africa. J Med Virol. 2012; 84 (4):601–7.

115. Hajarizadeh B, Grebely J, Dore GJ. Epidemiology and natural history of HCV infection. Nature Reviews Gastroenterology and Hepatology. 2013;10(9):553-62.

116. Magiorkinis G, Magiorkinis E, Paraskevis D, Ho SY, Shapiro B, Pybus OG, Allain JP, Hatzakis A. The global spread of hepatitis C virus 1a and 1b: a phylodinamic and phylogeographic analysis. PLoS Med. 2009;6(12): e1000198.

117. Forbi JC, Purdi MA, Campo DS, Vaughan G, Dimitrova GE, Ganova-Raeva LM et al. Epidemic history of hepatitis C virus infection in two remote communities in Nigeria, West Africa. J Gen Virol. 2012; 93 (Pt 7):1410–21.

118. Gower E, Estes C, Blach S, Razavi-Shearer K, Razavi H. Global epidemiology and genotype distribution of the hepatitis C virus infection. J. Hepatol. 2014, 61, S45–S57.

119. Marascio N, Ciccozzi M, Equestre M, Lo Presti A, Costantino A, Cella E, Bruni R, Liberto MC, Pisani G, Zicca E, Barreca GS, Torti C, Focà A, Ciccaglione AR. Back to the origin of HCV 2c subtype and spreading to the Calabria region (Southern Italy) over the last two centuries: a phylogenetic study. Infect Genet Evol. 2014;26:352-8.

120. Mellor J, Holmes EC, Jarvis LM, Yap PL, Simmonds P. Investigation of the pattern of hepatitis C virus sequence diversity in different geographical regions: implications for virus classification. The International HCV Collaborative Study Group. J Gen Virol. 1995;76:2493-507.

121. Ruggieri A, Argentini C, Kouruma F, Chionne P, D'Ugo E, Spada E, Dettori S,Sabbatani S, Rapicetta M. Heterogeneity of hepatitis C virus genotype 2 variants in West Central Africa (Guinea Conakry). J Gen Virol. 1996;77:2073-6

122. Jeannel D, Fretz C, Traore Y, Kohdjo N, Bigot A, Pê Gamy E, Jourdan G, Kourouma K, Maertens G, Fumoux F, Fournel JJ, Stuyver L. Evidence for high genetic diversity and long-term endemicity of hepatitis C virus genotypes 1 and 2 in West Africa. J Med Virol. 1998;55(2):92-7

123. Wansbrough-Jones MH, Frimpong E, Cant B, Harris K, Evans MR, Teo CG. Prevalence and genotype of hepatitis C virus infection in pregnant women and blood donors in Ghana. Trans R Soc Trop Med Hyg. 1998;92(5):496-9

124. Candotti D, Temple J, Sarkodie F, Allain JP. Frequent recovery and broad genotype 2 diversity characterize hepatitis C virus infection in Ghana, West Africa. J Virol. 2003;77(14):7914-23.

125. Bukh J, Purcell RH, Miller RH. At least 12 genotypes of hepatitis C virus predicted by sequence analysis of the putative E1 gene of isolates collected worldwide. Proc Natl Acad Sci U S A. 1993 Sep 1;90(17):8234-8.

126. Stuyver L, Rossau R, Wyseur A, Duhamel M, Vanderborght B, Van Heuverswyn H,Maertens G. Typing of hepatitis C virus isolates and characterization of new subtypes using a line probe assay. J Gen Virol. 1993;74 (Pt 6):1093-102.

127. Xu LZ, Larzul D, Delaporte E, Bréchot C, Kremsdorf D. Hepatitis C virus genotype 4 is highly prevalent in central Africa (Gabon). J Gen Virol. 1994;75 :2393-8.

128. Fretz C, Jeannel D, Stuyver L, Hervé V, Lunel F, Boudifa A, Mathiot C, de Thé G, Fournel JJ. HCV infection in a rural population of the Central African Republic (CAR): evidence for three additional subtypes of genotype 4. J Med Virol. 1995;47(4):435-7

129. Menendez C, Sanchez-Tapias JM, Kahigwa E, Mshinda H, Costa J, Vidal J, Acosta C, Lopez-Labrador X, Olmedo E, Navia M, Tanner M, Rodes J, Alonso PL. Prevalence and mother-to-infant transmission of hepatitis viruses B, C, and E in Southern Tanzania. J Med Virol. 1999;58(3):215-20

130. Ndjomou J, Pybus OG, Matz B. Phylogenetic analysis of hepatitis C virus isolates indicates a unique pattern of endemic infection in Cameroon. J Gen Virol. 2003;84(Pt 9):2333-41

131. Ciccozzi M, Zehender G, Polat C, Lai A, Kayman T, Sorrentino C, Ebranati E,Cella E, Lo Presti A, Abacioglu H. Phylogenetic analysis of HCV-4d in Turkey: the curious case of Kayseri province. J Med Virol. 2014 Mar;86(3):454-60

132. Wasitthankasem R, Vongpunsawad S, Siripon N, Suya C, Chulothok P, Chaiear K, Rujirojindakul P, Kanjana S, Theamboonlers A, Tangkijvanich P, Poovorawan Y. Genotypic distribution of hepatitis C virus in Thailand and Southeast Asia. PloS One. 2015;10(5):e0126764

133. Murphy DG, Sablon E, Chamberland J, Fournier E, Dandavino R, Tremblay CL. Hepatitis C virus genotype 7, a new genotype originating from Central Africa. J Clin Microbiol 2015; 53 (3): 967–972.

134. Welzel TM, Bhardwaj N, Hedskog C, Chodavarapu K, Camus G, McNally J, Brainard D, Miller MD, Mo H, Svarovskaia E, et al. Global epidemiology of HCV subtypes and resistance-associated substitutionsevaluated by sequencing-based subtype analyses. J. Hepatol. 2017, 67, 224–236.

135. Mugosa B, Cella E, Lai A, Lo Presti A, Blasi A, Vratnica Z, Vujoševic D, Ebranati E, Lauševic D, Guarino M, Zehender G, Milano T, Pascarella S, Spoto S,Angeletti S, Ciccozzi M. Hepatitis C virus genotype 3A in a population of injecting drug users in Montenegro: Bayesian and evolutionary analysis. Arch Virol. 2017;162(6):1549-1561.

136. AASLD/IDSA/IAS–USA. Recommendations for Testing, Managing, and Treating Hepatitis C. Available online: http://www.hcvguidelines.org

137. Purdy MA, Khudyakov YE. Evolutionary history and population dynamics of hepatitis E virus. PLoS One. 2010;5(12):e14376.

138. Doceul V, Bagdassarian E, Demange A, Pavio N. Zoonotic Hepatitis E Virus: Classification, Animal Reservoirs and Transmission Routes. Viruses. 2016;8(10). pii: E270.

139. Zehender G, Ebranati E, Lai A, Luzzago C, Paladini S, Tagliacarne C, Galli C, Galli M, Ciccozzi M, Zanetti AR, Romanò L. Phylogeography and phylodinamics of European genotype 3 hepatitis E virus. Infect Genet Evol. 2014;25:138-43.

140. Mirazo S, Mir D, Bello G, Ramos N, Musto H, Arbiza J. New insights into the hepatitis E virus genotype 3 phylodinamics and evolutionary history. Infect Genet Evol. 2016;43:267-73.

141. Nakano T, Takahashi K, Pybus OG, Hashimoto N, Kato H, Okano H, Kobayashi M, Fujita N, Shiraki K, Takei Y, Ayada M, Arai M, Okamoto H, Mishiro S. New findings regarding the epidemic history and population dynamics of Japan-indigenous genotype 3 hepatitis E virus inferred by molecular evolution. Liver Int. 2012;32(4):675-88.

142. De Sabato L, Vaccari G, Lemey P, Amoroso MG, Fusco G, Ianiro G, Di Bartolo I. Phylogenetic analysis of two genotype 3 Hepatitis E viruses from wild boar, Italy. Virus Genes. 2018;54(6):812-817

143. Garbuglia AR, Scognamiglio P, Petrosillo N, Mastroianni CM, Sordillo P, Gentile D, La Scala P, Girardi E, Capobianchi MR. Hepatitis E virus genotype 4 outbreak, Italy, 2011. Emerg Infect Dis. 2013;19(1):110-4.

144. Nakano T, Takahashi K, Takahashi M, Nishigaki Y, Watanabe N, Ishida S, Fujimoto S, Kato H, Okano H, Takei Y, Ayada M, Tomita E, Arai M, Okamoto H, Mishiro S. Investigating the origin and global dispersal history of hepatitis E virus genotype 4 using phylogeographical analysis. Liver Int. 2016;36(1):31-41.

145. Abecasis AB, Deforche K, Snoeck J, Bacheler LT, McKenna P, Carvalho AP, Gomes P, Camacho RJ, Vandamme AM: Protease mutation M89I/V is linked to therapy failure in patients infected with the HIV-1 non-B subtypes C. F or G. Aids 2005, 19:1799.

146. Abecasis AB, Deforche K, Bacheler LT, McKenna P, Carvalho AP, Gomes P, Vandamme AM, Camacho RJ: Investigation of baseline susceptibility to protease inhibitors in HIV-1 subtypes C, F, G and CRF02_AG. Antivir Ther 2006, 11:581.

147. Abecasis AB, Vandamme AM, Lemey P: Quantifying differences in the tempo of human immunodeficiency virus type 1 subtype evolution. J Virol 2009, 83:12917.

148. Camacho RJ, Vandamme AM: Antiretroviral resistance in different HIV-1 subtypes: impact on therapy outcomes and resistance testing interpretation. Curr Opin HIV AIDS 2007, 2:123.

149. Brenner B, Turner D, Oliveira M, Moisi D, Detorio M, Carobene M, Marlink RG, Schapiro J, Roger M, Wainberg MA: A V106M mutation in HIV-1 clade C viruses exposed to efavirenz confers cross-resistance to nonnucleoside reverse transcriptase inhibitors. AIDS 2003, 17:F1.

150. Grossman Z, Paxinos EE, Averbuch D, Maayan S, Parkin NT, Engelhard D, Lorber M, Istomin V, Shaked Y, Mendelson E, et al: Mutation D30N is not preferentially selected by human immunodeficiency virus type 1 subtype C in the development of resistance to nelfinavir. Antimicrob Agents Chemother 2004, 48:2159.

151. van de Vijver DA, Wensing AMJ, Angarano G, Asjö B, Balotta C, Boeri E, Camacho R, Chaix M-L, Costagliola D, De Luca A, Derdelinckx I, Grossman Z, Hamouda O, Hatzakis A, Hemmer R, Hoepelman A, Horban A, Korn K, Kücherer C, Leitner T, Loveday C, MacRae E, Maljkovic I, de Mendoza C, Meyer L, Nielsen C, Op de Coul ELM, Ormaasen V, Paraskevis D, Perrin L, Puchhammer-Stöckl E, Ruiz L, Salminen M, Schmit J-C, Schneider F, Schuurman R, Soriano V, Stanczak G, Stanojevic M, Vandamme A-M, Van Laethem K, Violin M, Wilbe K, Yerly S, Zazzi M, Boucher CAB: The calculated genetic barrier for antiretroviral drug resistance substitutions is largely similar for different HIV-1 subtypes. J Acquir Immune Defic Syndr 2006, 41:352–360.

152. Baeten JM, Chohan B, Lavreys L, Chohan V, McClelland RS, Certain L, Mandaliya K, Jaoko W, Julie O. HIV-1 subtype D infection is associated with faster disease progression than subtype A in spite of similar plasma HIV-1 loads. J Infect Dis 2007, 195:1177.

153. Kouri V, Aleman Y, Perez L, Perez J, Fonseca C, Correa C, Aragones C, Campos J, Alvarez D, Schrooten Y, Vinken L, Limia C, Soto Y, Vandamme AM, Van Laethem K. High frequency of antiviral drug resistance and non-b subtypes in HIV-1 patients failing antiviral therapy in Cuba. J Int AIDS Soc 2014; 17: 19754.

154. Kouri V, Khouri R, Aleman Y, Abrahantes Y, Vercauteren J, Pineda-Pena AC, Theys K, Megens S, Moutschen M, Pfeifer N, Van Weyenbergh J, Perez AB, Perez J, Perez L, Van Laethem K, Vandamme AM. CRF19_cpx is an Evolutionary fit HIV-1 Variant Strongly Associated With Rapid Progression to AIDS in Cuba. EBioMedicine 2015; 2: 244-254.

155. Abecasis AB, Wensing AM, Paraskevis D, Vercauteren J, Theys K, Van de Vijver DA, Albert J, Asjö B, Balotta C, Beshkov D, Camacho RJ, Clotet B, De Gascun C, Griskevicius A, Grossman Z, Hamouda O, Horban A, Kolupajeva T, Korn K, Kostrikis LG, Kücherer C, Liitsola K, Linka M, Nielsen C, Otelea D, Paredes R, Poljak M, Puchhammer-Stöckl E, Schmit JC, Sönnerborg A, Stanekova D, Stanojevic M, Struck D, Boucher CA, Vandamme AM. HIV-1 subtype distribution and its demographic determinants in newly diagnosed patients in Europe suggest highly compartmentalized epidemics. Retrovirology. 2013;10:7.

156. Magiorkinis G, Angelis K, Mamais I, Katzourakis A, Hatzakis A, Albert J, Lawyer G, Hamouda O, Struck D, Vercauteren J, Wensing A, Alexiev I, Åsjö B, Balotta C, Gomes P, Camacho RJ, Coughlan S, Griskevicius A, Grossman Z, Horban A, Kostrikis LG, Lepej SJ, Liitsola K, Linka M, Nielsen C, Otelea D, Paredes R, Poljak M, Puchhammer-Stöckl E, Schmit JC, Sönnerborg A, Staneková D, Stanojevic M, Stylianou DC, Boucher CAB; SPREAD program, Nikolopoulos G, Vasylyeva T, Friedman SR, van de Vijver D, Angarano G, Chaix ML, de Luca A, Korn K, Loveday C, Soriano V, Yerly S, Zazzi M, Vandamme AM, Paraskevis D. The global spread of HIV-1 subtype B epidemic. Infect Genet Evol. 2016;46:169-179.

157. Gilbert MT, Rambaut A, Wlasiuk G, Spira TJ, Pitchenik AE, Worobey M. The emergence of HIV/AIDS in the Americas and beyond. Proc. Natl. Acad. Sci. USA 2007; 104: 18566–18570.

158. Alvarez-Del Arco D, Fakoya I, Thomadakis C, et al. High levels of postmigration HIV acquisition within nine European countries. AIDS 2017; 31: 1979–88.

159. Paraskevis D, Kostaki E, Nikolopoulos GK, et al. Molecular Tracing of the Geographical Origin of Human Immunodeficiency Virus Type 1 Infection and Patterns of Epidemic Spread Among Migrants Who Inject Drugs in Athens. Clin Infect Dis 2017; 65: 2078–84.

160. Pantazis N, Thomadakis C, Del Amo J, Alvarez-Del Arco D, Burns FM, Fakoya I, Touloumi G; aMASE and CASCADE study groups. Determining the likely place of HIV acquisition for migrants in Europe combining subject-specific information and biomarkers data. Stat Methods Med Res 2017; : 962280217746437.

161. Hué S, Pillay D, Clewley JP, Pybus OG. Genetic analysis reveals the complex structure of HIV-1 transmission within defined risk groups. Proc Natl Acad Sci U S A. 2005;102(12):4425-9.

162. Lewis F, Hughes GJ, Rambaut A, Pozniak A, Leigh Brown AJ. Episodic sexual transmission of HIV revealed by molecular phylodinamics. PLoS Med. 2008;5(3):e50.

163. Zehender G, Ebranati E, Lai A, Santoro MM, Alteri C, Giuliani M, Palamara G, Perno CF, Galli M, Lo Presti A, Ciccozzi M. Population dynamics of HIV-1 subtype B in a cohort of men-having-sex-with-men in Rome, Italy. J Acquir Immune Defic Syndr. 2010;55:156-60.

164. Bobkov AF, Pokrovskiĭ VV, Selimova LM, Ladnaia NN, Kazennova EV, Bobkova MR, Khanina TA, Kravchenko AV, Iurin OG, Saukhat SR. [Genetic characteristics of variants of human immunodeficiency virus type 1, causing an epidemic among substance abusers in Commonwealth of Independent States countries]. Vopr Virusol. 1998;43:253-6.

165. Thomson MM, Vinogradova A, Delgado E, Rakhmanova A, Yakovlev A, Cuevas MT, Munoz M, Pinilla M, Vega Y, Perez-Alvarez L, Osmanov S, Najera R. Molecular epidemiology of HIV-1 in St Petersburg, Russia: predominance of subtype A, former Soviet Union variant, and identification of intrasubtype subclusters. J. Acquir. Immune Defic. Syndr. 2009; 51: 332–339.

166. Abecasis AB, Pingarilho M, Vandamme AM. Phylogenetic analysis as a forensic tool in HIV transmission investigations. AIDS. 2018;32:543-554.

167. Bernard EJ, Azad Y, Vandamme AM, Weait M, Geretti AM. HIV forensics: pitfalls and acceptable standards in the use of phylogenetic analysis as evidence in criminal investigations of HIV transmission. HIV Med. 2007;8:382-7.

168. Lemey P, Van Dooren S, Van Laethem K, Schrooten Y, Derdelinckx I, Goubau P, Brun-Vézinet F, Vaira D, Vandamme AM. Molecular testing of multiple HIV-1 transmissions in a criminal case. AIDS. 2005;19:1649-58.

169. de Oliveira T, Pybus OG, Rambaut A, Salemi M, Cassol S, Ciccozzi M, Rezza G, Gattinara GC, D'Arrigo R, Amicosante M, Perrin L, Colizzi V, Perno CF; Benghazi Study Group. Molecular epidemiology: HIV-1 and HCV sequences from Libyan outbreak. Nature. 2006;444(7121):836-7.

170. Williamson S, Perry SM, Bustamante CD, et al. A statistical characterization of consistent patterns of human immunodeficiency virus evolution within infected patients. Mol Biol Evol 2005;22:456-68.

171. Markham RB, Wang WC, Weisstein AE, et al. Patterns of HIV-1 evolution in individuals with differing rates of CD4 T cell decline. Proc Natl Acad Sci USA 1998;95:12568-73.

172. Becquart P, Courgnaud V, Willumsen J, Van de Perre P. Diversity of HIV-1 RNA and DNA in breast milk from HIV-1 infected mothers. Virology 2007;363:256-60.

173. Gray RR, Salemi M, Lowe A, Nakamura KJ, Decker WD, Sinkala M, Kankasa C, Mulligan CJ, Thea DM, Kuhn L, Aldrovandi G, Goodenow MM. Multiple independent lineages of HIV-1 persist in breast milk and plasma. AIDS. 2011;25(2):143-52.

174. Lemey P., Rambau A., Drummond AJ., Suchard MA. Bayesian phylogeography finds its roots. PLOS Comput. Biol. 2009; 5, e1000520.

175. CDC, CENTERS FOR DISEASE CONTROL AND PREVENTION. Arbovirus Catalog. wwwn.cdc.gov.

176. Naveca FG, Claro I, Giovanetti M, de Jesus JG, Xavier J, Iani FCdM, et al. (2019) Genomic, epidemiological and digital surveillance of Chikungunya virus in the Brazilian Amazon. PLoS Negl Trop Dis 13(3): e0007065

177. Tanabe ELL, Tanabe ISB, Santos ECD, Marques JPDS, Borges AA, Lima MC, Anderson L, Bassi ÊJ. Report of East-Central South African Chikungunya virus genotype during the 2016 outbreak in the Alagoas State, Brazil. Rev Inst Med Trop Sao Paulo. 2018;60:e19.

178. Faria NR, da Costa AC, Lourenço J, Loureiro P, Lopes ME, Ribeiro R, Alencar CS, Kraemer MUG, Villabona-Arenas CJ, Wu CH, Thézé J, Khan K, Brent SE, Romano C, Delwart E, Custer B, Busch MP, Pybus OG, Sabino EC; NHLBI Recipient Epidemiology and Donor Evaluation Study-III (REDS-III). Genomic and epidemiological characterisation of a dengue virus outbreak among blood donors in Brazil. Sci Rep. 2017;7(1):15216.

179. Cella E, Riva E, Salemi M, Spoto S, Vita S, Fogolari M, Angeletti S, Ciccozzi M. The new Chikungunya virus outbreak in Italy possibly originated from a single introduction from Asia. Pathog Glob Health. 2018;112(2):93-95.

180. Faria NR, Azevedo RDSDS, Kraemer MUG, Souza R, Cunha MS, Hill SC, Thézé J, Bonsall MB, Bowden TA, Rissanen I, Rocco IM, Nogueira JS, Maeda AY, Vasami FGDS, Macedo FLL, Suzuki A, Rodrigues SG, Cruz ACR, Nunes BT, Medeiros DBA, Rodrigues DSG, Queiroz ALN, da Silva EVP, Henriques DF, da Rosa EST, de Oliveira CS, Martins LC, Vasconcelos HB, Casseb LMN, Simith DB, Messina JP, Abade L, Lourenço J, Alcantara LCJ, de Lima MM, Giovanetti M, Hay SI, de Oliveira RS, Lemos PDS, de Oliveira LF, de Lima CPS, da Silva SP, de Vasconcelos JM, Franco L, Cardoso JF, Vianez-Júnior JLDSG, Mir D, Bello G, Delatorre E, Khan K, Creatore M, Coelho GE, de Oliveira WK, Tesh R, Pybus OG, Nunes MRT, Vasconcelos PFC. Zika virus in the Americas: Early epidemiological and genetic findings. Science. 2016;352(6283):345-349

181. Grubaugh ND, Ladner JT, Kraemer MUG, Dudas G, Tan AL, Gangavarapu K, Wiley MR, White S, Thézé J, Magnani DM, Prieto K, Reyes D, Bingham AM, Paul LM, et al. Genomic epidemiology reveals multiple introductions of Zika virus into the United States. Nature. 2017;546(7658):401-405

182. Giovanetti M, Milano T, Alcantara LC, Carcangiu L, Cella E, Lai A, Lo Presti A, Pascarella S, Zehender G, Angeletti S, Ciccozzi M. Zika Virus spreading in South America: Evolutionary analysis of emerging neutralizing resistant Phe279Ser strains. Asian Pac J Trop Med. 2016;9(5):445-52.

183. Sasmono RT, Perkasa A, Yohan B, Haryanto S, Yudhaputri FA, Hayati RF, Ma'roef CN, Ledermann JP, Aye Myint KS, Powers AM. Chikungunya Detection during Dengue Outbreak in Sumatra, Indonesia: Clinical Manifestations and Virological Profile. Am J Trop Med Hyg. 2017;97(5):1393-1398.

184. de Souza TMA, Ribeiro ED, Corrêa VCE, Damasco PV, Santos CC, de Bruycker-Nogueira F, Chouin-Carneiro T, Faria NRDC, Nunes PCG, Heringer M, Lima MDRQ, Badolato-Corrêa J, Cipitelli MDC, Azeredo EL, Nogueira RMR, Dos Santos FB. Following in the Footsteps of the Chikungunya Virus in Brazil: The First Autochthonous Cases in Amapá in 2014 and Its Emergence in Rio de Janeiro during 2016. Viruses. 2018;10(11).

185. Cunha MS, Cruz NVG, Schnellrath LC, Medaglia MLG, Casotto ME, Albano RM, Costa LJ, Damaso CR. Autochthonous Transmission of East/Central/South African Genotype Chikungunya Virus, Brazil. Emerg Infect Dis. 2017;23(10):1737-1739

186. Jeong YE, Cha GW, Cho JE, Lee EJ, Jee Y, Lee WJ. Viral and serological kinetics in Zika virus-infected patients in South Korea. Virol J. 2017;14(1):70.

187. Stapleford KA, Moratorio G, Henningsson R, Chen R, Matheus S, Enfissi A, Weissglas-Volkov D, Isakov O, Blanc H, Mounce BC, Dupont-Rouzeyrol M, Shomron N, Weaver S, Fontes M, Rousset D, Vignuzzi M. Whole-Genome Sequencing Analysis from the Chikungunya Virus Caribbean Outbreak Reveals Novel Evolutionary Genomic Elements. PLoS Negl Trop Dis. 2016;10(1):e0004402

188. Lo Presti A, Ciccozzi M, Cella E, Lai A, Simonetti FR, Galli M, Zehender G, Rezza G. Origin, evolution, and phylogeography of recent epidemic CHIKV strains. Infect Genet Evol. 2012;12(2):392-8.

189. Faria NR, Kraemer MUG, Hill SC, Goes de Jesus J, Aguiar RS, Iani FCM, et al. Genomic and epidemiological monitoring of yellow fever virus transmission potential. Science. 2018;361(6405):894-899.

190. Thézé J, Li T, du Plessis L, Bouquet J, Kraemer MUG, Somasekar S, Yu G, de Cesare M, Balmaseda A, Kuan G, Harris E, Wu CH, Ansari MA, Bowden R, et al. Genomic Epidemiology Reconstructs the Introduction and Spread of Zika Virus in Central America and Mexico. Cell Host Microbe. 2018;23(6):855-864.e7.

191. Gradner SD, Field AM, Coleman DV, and Hulme B. New human papovavirus (B.K.) isolated from urine after renal transplantation. Lancet 1971; i: 1253-1257

192. Padgett BL, Walker DL, ZuRhein GM, Eckroade RJ, and Dessel BH. Cultivation of papova-like virus from human brain with progressive multifocal leukoencephalopathy. Lancet 1971.

193. Jin L, Gibson PE, Knowles WA, Clewley JP. BK virus antigenic variants: sequence analysis within the capsid VP1 epitope. J Med Virol 1993; 39: 50-56.

194. Jin L, Pietropaolo V, Booth J C, Ward K H & Brown D W. Prevalence and distribution of BK virus subtypes in healthy people and immunocompromised patients detected by PCRrestriction enzyme analysis. Clin Diag Virol 1995; 3: 285–295.

195. Ikegaya H, Saukko PJ, Tertti R, Metsarinne KP, Carr MJ, Crowley B, Sakurada K, Zheng H-Y, Kitamura T and Yogo Y. Identification of a genomic subgroup of BK polyomavirus spread in European populations. J Gen Virol. 2006;87:3201-3208.

196. Nishimoto Y, Zheng HY, Zhong S, Ikegaya H, Chen Q, Sugimoto C, Kitamura T, Yogo Y. An Asian origin for subtype IV BK virus based on phylogenetic analysis. J Mol Evol 2007; 65(1):103-111.

197. Yogo Y, Kitamura T, Sugimoto C, Ueki T, Aso Y, Hara K, and Taguchi F. Isolation of a Possible archetypal JC virus DNA sequence from nonimmunocompromised individuals. J Virol 1990; 64: 3139-3143.

198. Sugimoto C, Kitamura T, Guo J, AL-Ahdal MN, Shchelkunov SN, Otova B, Ondrejka P, Chollet J-Y, El-Safi S, Ettayebi M, Gresenguet G, Kocago T, Chaiyarasamee S, Thant KZ, Thein S, Moe K, Kobayashi N, Taguchi F, and Yogo Y. Typing of urinary JC virus DNA offers a novel means of tracing human migrations. Proc Natl Acad Sci USA 1997; 94: 9191–9196.

199. Agostini HT, Ryschkewitsch CF, Baumhefner RW, Tourtellotte WW, Singer EJ, Komoly S, Stoner GL. Influence of JC virus coding region genotype on risk of multiple sclerosis and progressive multifocal leukoencephalopathy. J Neurovirol 2000;6: S101–S108.

200. Allander T, Andreasson K, Gupta S, Bjerkner A, Bogdanovic G, Persson MAA, Dalianis T, Ramqvist T, and Andersson B. Identification of a third human polyomavirus. J Virol 2007; 81: 4130–4136.

201. Gaynor AM, Nissenm D, Whiley DM, Mackay IM, Lambert SB, Wu G, Brennan DC, Storch GA, Sloots TP, Wang D. Identification of a novel polyomavirus from patients with acute respiratory tract infections. PLoS Pathog 2007; 3:595–604.

202. Bialasiewicz S, Rockett R, Whiley DW, Abed Y, Allander T, Binks M, Boivin G, Cheng AC, Chung J-Y, Ferguson PE, Gilroy NM, Leach AJ, Lindau C, Rossen JW, Sorrell TC, Nissen MD, and Sloots TP. Whole-Genome Characterization and Genotyping of Global WU Polyomavirus Strains. J Virol 2010; 84: 6229-6234.

203. Feng H, Shuda M, Chang Y, and Moore PS. Clonal integration of a polyomavirus in human Merkel cell carcinoma. Science 2008; 319: 1096–1100.

204. Calvignac-Spencer S, Feltkamp MCW, Daugherty M, Moens U, Ramquist T, Johne R, Ehlers B. Polyomaviridae Study Group of the International Committee on Taxonomy of Viruses. A taxonomy update for the family Polyomaviridae. Arch Virol 2016; 161: 1739-1750.

205. Martel-Jantin C, Filippone C, Cassar O, Peter M, Tomasic G, Vielh P, Brière J, Petrella T, Aubriot-Lorton MH, Mortier L, Jouvion G, Sastre-Garau X, Robert C, Gessain A. Genetic variability and integration of Merkel cell polyomavirus in Merkel cell carcinoma. Virology 2012; 426: 134–142.

206. Katano H, Ito H, Suzuki Y, Nakamura T, Sato Y, Tsuji T, Matsuo K, Nakagawa H, Sata T. Detection of Merkel cell polyomavirus in Merkel cell carcinoma and Kaposi's sarcoma. J. Med. Virol. 2009; 81: 1951–1958.

207. Schowalter RM, Pastrana DV, Pumphrey KA, Moyer AL, Buck CB. Merkel cell polyomavirus and two previously unknown polyomaviruses are chronically shed from human skin. Cell Host Microbe 2010; 7: 509–515.

208. Hashida Y, Higuchi T, Matsuzaki S, Nakajima K, Sano S, Daibata M. Prevalence and genetic variability of human polyomaviruses 6 and 7 in healthy skin among asymptomatic individuals. J Infect Dis 2018;217: 483-493.

209. Yogo Y, Sugimoto C, Zheng HY, Ikegaya H, Takasaka T, Kitamura T. JC virus genotyping offers a new paradigm in the study of human populations. Rev Med Virol 2004; 14: 179–191.

210. Ho J, Jedrych JJ, Feng H, Natalie AA, Grandinetti L, Mirvish E, Crespo MM, Yadav D, Fasanella KE, Proksell S, Kuan SF, Pastrana DV, Buck CB, Shuda Y, Moore PS, Chang Y. Human polyomavirus 7-associated pruritic rash and viremia in transplant recipients. J Infect Dis 2015; 211:1560-1565.

211. Rascovan N, Monteil Bouchard S, Grob JJ, Collet-Villette AM, Gaudy-Marqueste C, Penicaud M, Lepidi H, Raoult D, Desnues C. Human polyomavirus-6 infecting lymph nodes of a patient with an angiolymphoid hyperplasia with eosinophilia or Kimura disease. Clin Infect Dis 2016; 62:1419–1421.

212. van der Meijden E, Janssens RWA, Lauber C, Bavinck JNB, Gorbalenya AE, Feltkamp MCW. Discovery of a new human polyomavirus associated with Trichodysplasia spinulosa in an immunocompromized patient. Plos Pathogens 2010; 6: e1001024

213. Siebrasse EA, Reyes A, Lim ES, Zhao G, Mkakosya RS, Manary MJ, Gordon JI and Wang D. Identification of MW Polyomavirus, a novel polyomavirus in human tool. J Virol 2012; 86: 10321–10326.

214. Lim ES, Reyes A, Antonio M, Saha D, Ikumapayi UN, Adeyemi M, Stine OC, Skelton R, Brennan DC, Mkakosya RS, Manary MJ, Gordon JI, Wang D. Discovery of STL polyomavirus, a polyomavirus of ancestral recombinant origin that encodes a unique T antigen by alternative splicing. Virology 2013; 436: 295–303.

215. Korup S, Rietscher J, Calvignac-Spencer S, Trusch F, Hofmann J, Moens U, Sauer I, Voigt S, Schmuck R, Ehlers B. Identification of a novel human polyomavirus in organs of the gastrointestinal tract. Plos One 2013; 8: e58021.

216. Mishra N, Pereira M, Rhodes RH, An P, Pipas JM, Jain K, Kapoor A, Briese T, Faust PL, and Lipkin WI. Identification of a novel polyomavirus in a pancreatic transplant recipient with retinal blindness and vasculitic myopathy. J Infect Dis 2014; 210: 1595-1599.

217. Read AF, Mackinnon MJ. Pathogen evolution in a vaccinated world.In Evolution in Health and Disease. Oxford University press. 2007, 139-152pp.

218. Ypma RJF, W. Marijn van Ballegooijen WM, Wallinga J. Relating Phylogenetic Trees to Transmission Trees of Infectious Disease Outbreaks. 2013. Genetics, 195, 1055–1062.

219. Wohl S, Schaffner SF, Sabeti PC. Genomic Analysis of Viral Outbreaks. Annu Rev Virol. 2016;3:173-195.

220. Kühnert D, Wu CH, Drummond AJ. Phylogenetic and epidemic modeling of rapidly evolving infectious diseases. Infection, Genetics and Evolution 2011. 11, 1825–1841.

221. Geoghegan JL, Edward C. Holmes EC. Evolutionary Virology at 40. Genetics, 2018. 210, 1151–1162.

222. Ojosnegros S, Beerenwinkel N. Models of RNA virus evolution and their roles in vaccine design. Immunome Research 2010, 6:S5

223. World Health Organization. WHO recommendations on the composition of influenza virus vaccines http://www.who.int/influenza/vaccines/virus/recommendations/en/

224. Houser K, Subbarao K. Influenza vaccines: challenges and solutions. Cell Host Microbe 2015;17(3):295-300.

225. Carrat F, Flahault A. Influenza vaccine: the challenge of antigenic drift. Vaccine. 2007;25:6852–62.

226. Castro-Nallar E., Pérez-Losada M, Burton GF, Crandall KA.The evolution of HIV: Inferences using phylogenetics. Mol Phylogenet Evol. 2012. 62: 777–792.

227. Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, et al. HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. Science 2014; 346:56–61.

228. Rambaut A, Posada D, Crandall KA, Holmes EC. The causes and consequences of HIV evolution. Nat Rev Genet 2004; 5:52–61.

229. Cohen YZ, Dolin R. Novel HIV vaccine strategies: overview and perspective. Ther Adv Vaccines 2013; 1(3) 99–112.

230. Rappuoli R, Bottomley MJ, D'Oro U, Finco O, De Gregorio E. Reverse vaccinology 2.0: Human immunology instructs vaccine antigen design. JEM 2016; 213 (4): 469

231. Schweitzer A, Horn J, Mikolajczyk RT, Krause G, Ott JJ. Estimations of worldwide prevalence of chronic hepatitis B virus infection: a systematic review of data published between 1965 and 2013. The Lancet 2015;386: 1546–1555.

232. Tatematsu K, Tanaka Y, Kurbanov F, Sugauchi F, Mano S, Maeshiro T, Nakayoshi T, Wakuta M, Miyakawa Y, Mizokami M. A genetic variant of hepatitis B virus divergent from known human and ape genotypes isolated from a Japanese patient and provisionally assigned to new genotype J. Journal of Virology 2009; 83:10538–10547.

233. Kostaki EA, Karamitros T, Stefanou G, Mamais I, Konstantinos Angelis K, Hatzakis A, Kramvis A, Paraskevis D. Unravelling the history of hepatitis B virus genotypes A and D infection using a full-genome phylogenetic and phylogeographic approach. eLife 2018;7:e36709

234. Awerkiewa S, D¨aumer M, Reiser M, Wend UC, Pfister H, Kaiser R, Willems WR, Gerlich WH. Reactivation of an occult hepatitis B virus escape mutant in an anti-HBs positive, anti-HBc negative lymphoma patient. Journal of Clinical Virology 2007;38, 83–86

235. Kennedy PGE, Gershon AA. Clinical Features of Varicella-Zoster Virus Infection. Viruses 2018; 10, 609.

236. Chow VT, Tipples GA, Grose C. Bioinformatics of varicella-zoster virus: Single nucleotide polymorphisms define clades and attenuated vaccine genotypes. Infect Genet Evol. 2013;18: 351–356.

237. Sengupta N, Taha Y, Scott FT, Leedham-Green ME, Quinlivan M, Breuer J. Varicella-zoster-virus genotypes in East London: a prospective study in patients with herpes zoster. J Infect Dis. 2007;196:1014-20.

238. Didelot X, Walker AS, Peto TE, Crook DW, Wilson DJ. Within-host evolution of bacterial pathogens. Nat Rev Microbiol. 2016;14:150-62.

239. Xu Y, Liu B, Gröndahl-Yli-Hannuksila K, Tan Y, Feng L, Kallonen T, Wang L, Peng D, He Q, Wang L, Zhang S. Whole-genome sequencing reveals the effect of vaccination on the evolution of Bordetella pertussis. Scientific Reports 2015; 5:12888.

240. Gwinn M, MacCannell DR, Khabbaz RF. Integrating Advanced Molecular Technologies into Public Health. J Clin Microbiol. 2017;55(3):703-714

241. Manara S, Pasolli E, Dolce D, Ravenni N, Campana S, Armanini F, Asnicar F,Mengoni A, Galli L, Montagnani C, Venturini E, Rota-Stabelli O, Grandi G,Taccetti G, Segata N. Whole-genome epidemiology, characterisation, and phylogenetic reconstruction of Staphylococcus aureus strains in a paediatric hospital. Genome Med. 2018 Nov 13;10(1):82.

242. Loman NJ, Constantinidou C, Christner M, Rohde H, Chan JZ-M, Quick J, Weir JC, Quince C, Smith GP, Betley JR, Aepfelbacher M, Pallen MJ. 2013. A culture-independent sequence-based metagenomics approach to the investigation of an outbreak of Shiga-toxigenic Escherichia coli O104:H4.JAMA 201; 309:1502–1510.

243. Greninger AL, Zerr DM, Qin X, Adler AL, Sampoleo R, Kuypers JM, Englund JA,Jerome KR. Rapid Metagenomic Next-Generation Sequencing during an Investigation of Hospital-Acquired Human Parainfluenza Virus 3 Infections. J Clin Microbiol.2016;55(1):177-182.

244. Greninger AL, Waghmare A, Adler A, Qin X, Crowley JL, Englund JA, Kuypers JM, Jerome KR, Zerr DM. Rule-Out Outbreak: 24-Hour Metagenomic Next-Generation Sequencing for Characterizing Respiratory Virus Source for Infection Prevention. J Pediatric Infect Dis Soc. 2017;6(2):168-172

245. Zhu Y, Zembower TR, Metzger KE, Lei Z, Green SJ, Qi C. Investigation of Respiratory Syncytial Virus Outbreak on an Adult Stem Cell Transplant Unit by Use of Whole-Genome Sequencing. J Clin Microbiol. 2017;55(10):2956-2963.

246. Amendola A, Bianchi S, Frati ER, Ciceri G, Faccini M, Senatore S, Colzani D,Lamberti A, Baggieri M, Cereda D, Gramegna M, Nicoletti L, Magurano F, Tanzi E.Ongoing large measles outbreak with nosocomial transmission in Milan, northern Italy, March-August 2017. Euro Surveill. 2017;22(33). pii: 30596

247. Brown JR, Roy S, Shah D, Williams CA, Williams R, Dunn H, Hartley J, Harris K,Breuer J. Norovirus Transmission Dynamics in a Pediatric Hospital Using Full Genome Sequences. Clin Infect Dis. 2019;68(2):222-228.

Conflict of Interest: none.

Figure legends

Fig. 1. Percentage of papers published per year from 2010 to 2019 found by typing epidemiology, phylogeny and epidemiology plus phylogeny

Fig. 2. Hippocrates, the father of epidemiology (panel A) and Charles Darwin, the father of phylogeny (panel B)

Fig. 3 The hypothetical cycle of the bioinformatics in molecular evolution.

Fig. 4 Example of phylogenetic and phylodinamic approaches

Fig. 5 Mathematical evolutionary modeling applied on protein sequence (A) and structure analysis to make a prevision on the effect of selective pressure (B).

Fig. 6 The basis of molecular epidemiology

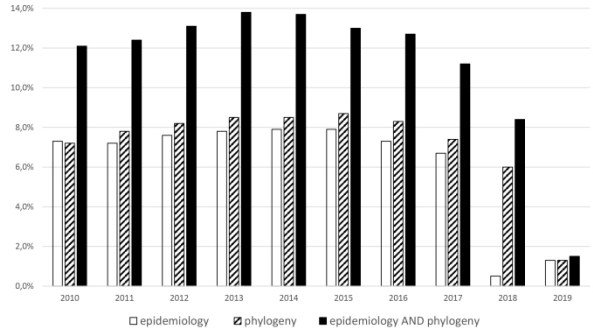

Fig. 1 Percentage of papers published per year from 2010 to 2019 found by typing epdiemiology, phylogeny and epidemiology plus phylogeny

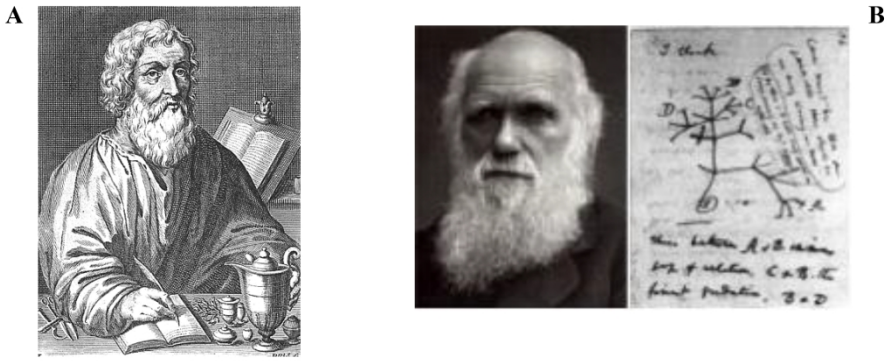

Fig. 2 Hippocrates, the father of epidemiology (panel A) and Charles Darwin, the father of phylogeny (panel B).

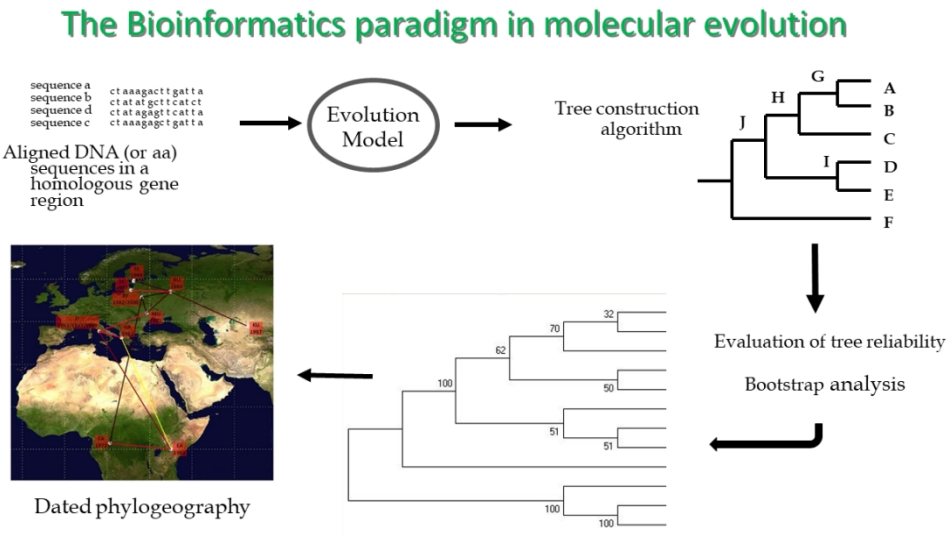

Fig. 3 The hypothetical cycle of the bioinformatics in molecular evolution.

338x190mm (96 x 96 DPI)

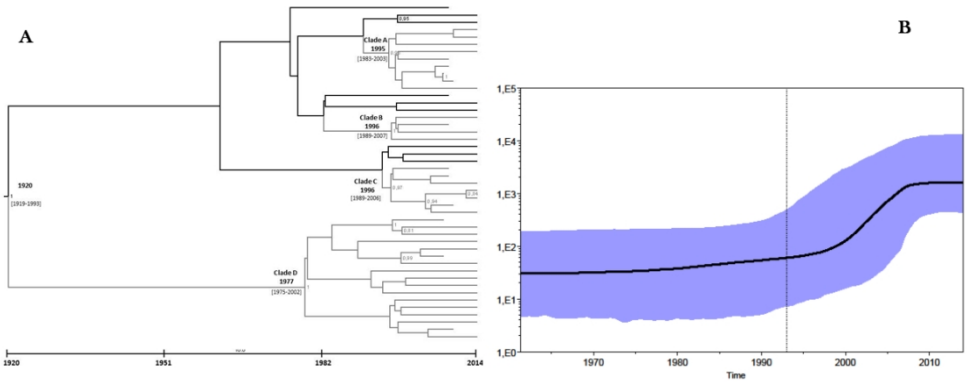

Fig. 4 Example of phylogenetic and phylodinamic approaches

338x190mm (96 x 96 DPI)

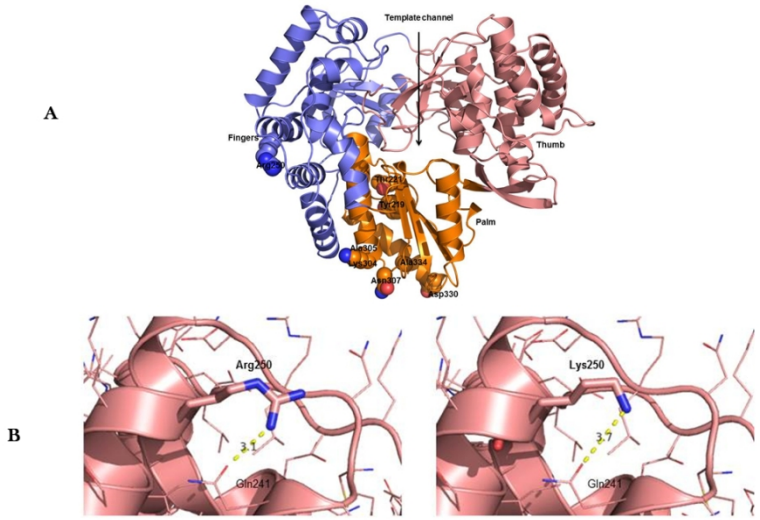

Fig. 5 Mathematical evolutionary modeling applied on protein sequence (A) and structure analysis to make a prevision on the effect of selective pressure (B).

338x190mm (96 x 96 DPI)

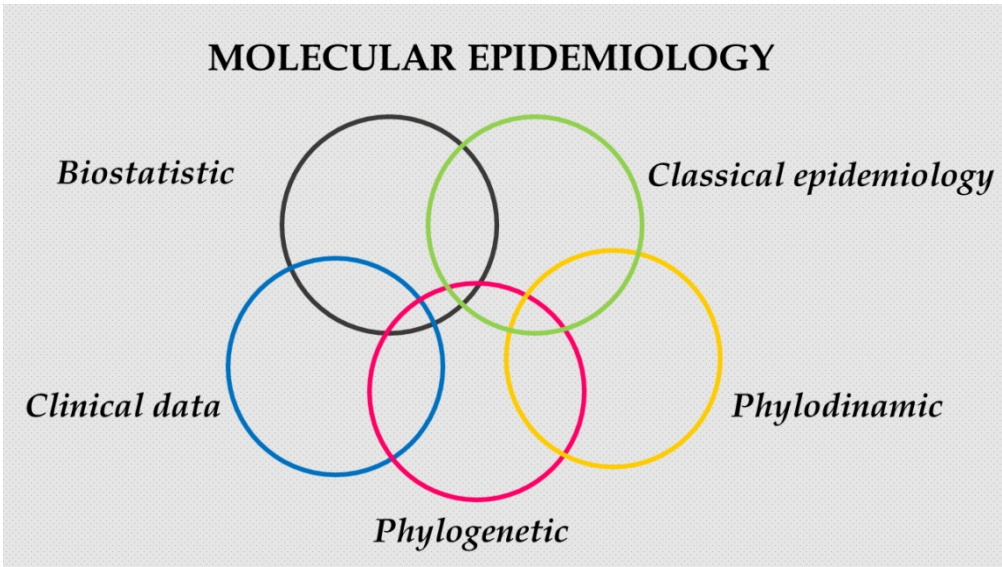

Fig. 6 The basis of molecular epidemiology

338x190mm (96 x 96 DPI)