Infection, Genetics and Evolution xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Review

Recombination in viruses: Mechanisms, methods of study, and evolutionary consequences

Marcos Pérez-Losada [a,b], Miguel Arenas [c], Juan Carlos Galán [d,e], Ferran Palero [e,f], Fernando González-Candelas [e,f,*]

[a] CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Campus Agrário de Vairão, Portugal
[b] Computational Biology Institute, George Washington University, Ashburn, VA 20147, USA
[c] Centre for Molecular Biology "Severo Ochoa", Consejo Superior de Investigaciones Científicas (CSIC), Madrid, Spain
[d] Servicio de Microbiología, Hospital Ramón y Cajal and Instituto Ramón y Cajal de Investigación Sanitaria (IRYCIS), Madrid, Spain
[e] CIBER en Epidemiología y Salud Pública, Spain
[f] Unidad Mixta Infección y Salud Pública, FISABIO-Universitat de València, Valencia, Spain

ARTICLE INFO

Article history:
Received 23 November 2014
Received in revised form 15 December 2014
Accepted 17 December 2014
Available online xxxx

Keywords:
Recombination
Reassortment
Linkage disequilibrium
Population structure
Recombination rate
Mutation rate

ABSTRACT

Recombination is a pervasive process generating diversity in most viruses. It joins variants that arise independently within the same molecule, creating new opportunities for viruses to overcome selective pressures and to adapt to new environments and hosts. Consequently, the analysis of viral recombination attracts the interest of clinicians, epidemiologists, molecular biologists and evolutionary biologists. In this review we present an overview of three major areas related to viral recombination: (i) the molecular mechanisms that underlie recombination in model viruses, including DNA-viruses (Herpesvirus) and RNA-viruses (Human Influenza Virus and Human Immunodeficiency Virus), (ii) the analytical procedures to detect recombination in viral sequences and to determine the recombination breakpoints, along with the conceptual and methodological tools currently used and a brief overview of the impact of new sequencing technologies on the detection of recombination, and (iii) the major areas in the evolutionary analysis of viral populations on which recombination has an impact. These include the evaluation of selective pressures acting on viral populations, the application of evolutionary reconstructions in the characterization of centralized genes for vaccine design, and the evaluation of linkage disequilibrium and population structure.

Contents

1. Introduction . . . 00
2. Mechanisms of recombination in viruses . . . 00
   2.1. Recombination in DNA viruses . . . 00
   2.2. Recombination in RNA viruses . . . 00
3. Estimation of recombination . . . 00
   3.1. Statistical tests to detect recombination . . . 00
   3.2. Inference of recombination breakpoints . . . 00
   3.3. Estimating recombination rates . . . 00
   3.4. The impact of next-generation sequencing (NGS) on the estimation of recombination . . . 00
4. The impact of recombination . . . 00
   4.1. The role of recombination with respect to mutation as a source of genetic diversity . . . 00
   4.2. The influence of recombination on selection estimation . . . 00

* Corresponding author at: Unidad Mixta Infección y Salud Pública, FISABIO-Universitat de València, C/Catedrático José Beltrán, 2, E-46980 Paterna, Valencia, Spain. Tel.: +34 963543653, +34 961925961; fax: +34 963543670.
E-mail address: fernando.gonzalez@uv.es (F. González-Candelas).

http://dx.doi.org/10.1016/j.meegid.2014.12.022
1567-1348/

4.3. The role of recombination in the design of centralized viral genes . . . 00
4.4. Population structure and recombination . . . 00
4.5. Interaction of population-level processes and recombination . . . 00
5. Conclusions . . . 00
Uncited reference . . . 00
Acknowledgements . . . 00
References . . . 00

1. Introduction

Viruses encompass an enormous variety of genomic structures. The majority of classified viruses have RNA genomes, but many have DNA genomes, and some may even have DNA and RNA at different stages in their life cycle. Attending to genome architecture, they can be either single-stranded (ss) or double-stranded (ds) and some virus families (e.g., Hepadnaviridae) contain a double-stranded genome with single-stranded regions. Finally, several virus classes (RNA genomes and some with ssDNA genomes) are classified on the basis of the nature and polarity of their mRNAs, depending on whether they have the translatable information in the same or complementary strand (denoted as plus-strand, +, or minus-strand, −). Several types of ssDNA (e.g., geminiviruses) and ssRNA (e.g., arenaviruses) viruses have genomes that are ambisense (+/−). All these genome properties are used to classify viruses into seven groups (following Baltimore, 1971): Group I, dsDNA viruses (e.g., Herpesvirus); II, ssDNA (e.g., Parvoviruses); III, dsRNA (e.g., Reoviruses); IV, (+)ssRNA (e.g., Picornaviruses); V, (−)ssRNA (e.g., Orthomyxovirus); VI, ssRNA-RT (reverse transcriptase) (e.g., Retroviruses); VII: dsDNA-RT (e.g., Hepadnaviruses). This huge biological diversity suggests multiple strategies for generating genetic diversity.

Viruses undergo genetic change by several mechanisms, including point mutation (the ultimate source of genetic variation) and recombination. In general, RNA viruses have smaller genomes than DNA viruses, probably as consequence of their higher mutation rates (Holmes, 2003). The reason for this inverse relationship between genome size and mutation rate is arguably the incapability of large RNA viruses to replicate without generating lethal mutations (Belshaw et al., 2007; Sanjuan et al., 2010). In contrast, DNA viruses generally have larger genomes because of the higher fidelity of their replication enzymes. ssDNA viruses are an exception to this rule, since the mutation rates of their genomes can be as high as those seen in ssRNA genomes (Duffy and Holmes, 2009).

Recombination occurs when at least two viral genomes co-infect the same host cell and exchange genetic segments. Different types of viral recombination are recognized based on the structure of the crossover site (Austermann-Busch and Becher, 2012; Scheel et al., 2013). Homologous recombination occurs in the same site in both parental strands, while non-homologous or illegitimate recombination (Lai, 1992) occurs at different sites of the genetic fragments involved, frequently originating aberrant structures (Galli and Bukh, 2014). A particular type of recombination, known as shuffling or reassortment, occurs in viruses with segmented genomes, which can interchange complete genome segments, giving rise to new segment combinations. Illegitimate recombination is relatively infrequent in RNA viruses, but in DNA viruses can occur at much higher frequencies than homologous recombination (Robinson et al., 2011). In addition, the exchange of genetic material between viruses is usually non-reciprocal, meaning the recipient of a genome portion does not act as donor of the replaced portion in the original source. In this respect, the term recombination does not have the same meaning in viruses that it does in diploid, sexually reproducing organisms wherein the exchange of genetic material between chromatids in the first meiosis division is reciprocal. Viral recombination could be more appropriately denoted as "gene conversion," but the term has become so widely used that it is pointless to change it.

Recombination is a widespread phenomenon in viruses and can have a major impact on their evolution. Indeed, recombination has been associated with the expansion of viral host ranges, the emergence of new viruses, the alteration of transmission vector specificities, increases in virulence and pathogenesis, the modification of tissue tropisms, the evasion of host immunity, and the evolution of resistance to antivirals (Martin et al., 2011a,b; Simon-Loriere and Holmes, 2011). The frequencies of recombination vary extensively among viruses. Recombination seems highly frequent in some dsDNA viruses, such as α-Herpesviruses, where recombination is intimately linked to replication and DNA repair (Robinson et al., 2011; Thiry et al., 2005) and can prevent the progressive accumulation of harmful mutations in their genomes (i.e., prevent the mutational meltdown and increase fitness). In fact, homologous recombination seems to be the most frequent mechanism in this group of viruses, although illegitimate recombination has also been observed (Robinson et al., 2011; Thiry et al., 2005, 2006). DNA recombination can also facilitate access to evolutionary innovations that would otherwise be inaccessible by mutation alone, although the viability of DNA recombinants will depend on how severely recombination disrupts genome architecture and the interaction networks (Lefeuvre et al., 2009).

Among RNA viruses, recombination is particularly frequent in Retroviruses (ssRNA-RT), most notably in HIV, where the rate of recombination per nucleotide exceeds that of mutation (Jetzt et al., 2000). In contrast, recombination occurs at variable frequencies in (+)ssRNA viruses, with some families showing high rates (e.g., Picornaviridae), while others see only occasional (e.g., Flaviviridae) or nonexistent (e.g., Leviviridae) occurrence. Current data also indicate that recombination is even far less common in (−)ssRNA viruses (Holmes, 2009). Similarly, there is extensive variation in the rate of re-assortment among RNA viruses. Many segmented viruses, such as Hantaviruses, Lassa virus and Tenuiviruses, exhibit relatively low levels of re-assortment, while others show high rates (e.g., influenza A virus, rotavirus A and Cystoviruses) (Simon-Loriere and Holmes, 2011).

The evolutionary reasons for the occurrence of recombination in RNA viruses are less clear. Since RNA viruses exhibit high mutation rates and large population sizes, it is more likely that these factors, rather than recombination, drive their evolutionary fate, as they regularly produce advantageous mutations and protect themselves from the accumulation of deleterious ones. There is little compelling evidence supporting the idea that recombination enhances genome repair (except in viruses with diploid or pseudodiploid genomes, such as HIV) or has evolved as a form of sexual reproduction, allowing the more efficient purging of deleterious mutations and accelerating the generation of advantageous genetic combinations (Holmes, 2009; Simon-Loriere and Holmes, 2011). Instead, prior hypotheses contend that mechanistic constraints associated with particular genome structures and viral life histories are the major elements shaping recombination rates in RNA viruses. Consequently, RNA recombination and re-assortment can be considered

Please cite this article in press as: Pérez-Losada, M., et al. Recombination in viruses: Mechanisms, methods of study, and evolutionary consequences. Infect. Genet. Evol. (2014), http://dx.doi.org/10.1016/j.meegid.2014.12.022