ABSTRACT
By using selection mapping, we can better understand the mutational landscapes of influenza A H3 hemagglutinin, HIV-1 reverse transcriptase, and HIV-gp120, which are not directly associated with the identification of amino acids at these codons but are considered positive selection in previous methods (e.g., dN/dS).

INTRODUCTION
The genetic structure of viruses is highly conserved and it is known that the viruses are able to self-replicate and spread to other hosts. It is also known that some viruses have a mechanism of replication and can be used as vectors for a variety of viral diseases.

Viruses can also be used as vectors for other diseases, for example, HIV and hepatitis B. In this study, we provide an algorithm to help identify positive candidates for viral disease with a high degree of accuracy. We present a novel approach to identify positive candidates in the human genome.

The project was sponsored by the US National Institutes of Health (NIH) under grant number R01 GM105872.

In addition to the authors, all authors read and approved the final manuscript.

Author Contributions: Dr M The creation of viral quasispecies and antigenic drift. A quasispecies is a group of related viral variants that coexist in both field populations and within single infected individuals, as defined by researchers at GeneBank. The hemagglutinin (HA) envelope surface glycoprotein--the primary neutralizing determinant of influenza A- is a classic example of an antigenically drifting protein. Walter Gerhard and colleagues demonstrated that monoclonal antibodies (Abs) selects for HA escape mutants in model systems. Later, Dimmock and others showed that polycloRNA (PSA) also select the escape mutations. Genealogical analyses have been used to determine the selective advantages of viral variation through phylogenetic analyses. This includes an overabundance of replacement mutations in a viral protein, which appears to confer selective advantage to HIV-1. We have developed a "selection mapping" algorithm to overcome these limitations and identify positively selected variants of influenza A HA (H3 serotype), HIV-1 reverse transcriptasE (RT), and HIV-120 gp120, by testing each observed replacement mutation at each codon.

CONCLUSION
We have developed an algorithm for mapping the positively selected mutations of viral quasispecies using sequence data. This method was used to map the positive selected variants (for example, those of influenza A HA, HIV-1 RT and HIV-120), as well as other obvious examples of selective mapping targeting viral subpopulations under different selective pressures: For example we used selection mapping of HIV isolates with different cellular tropisms; another application of this principle is possible for HIV breakthrough infections where we determined whether vaccines prevented the HIV quasivalent vaccine substances from a greater distance