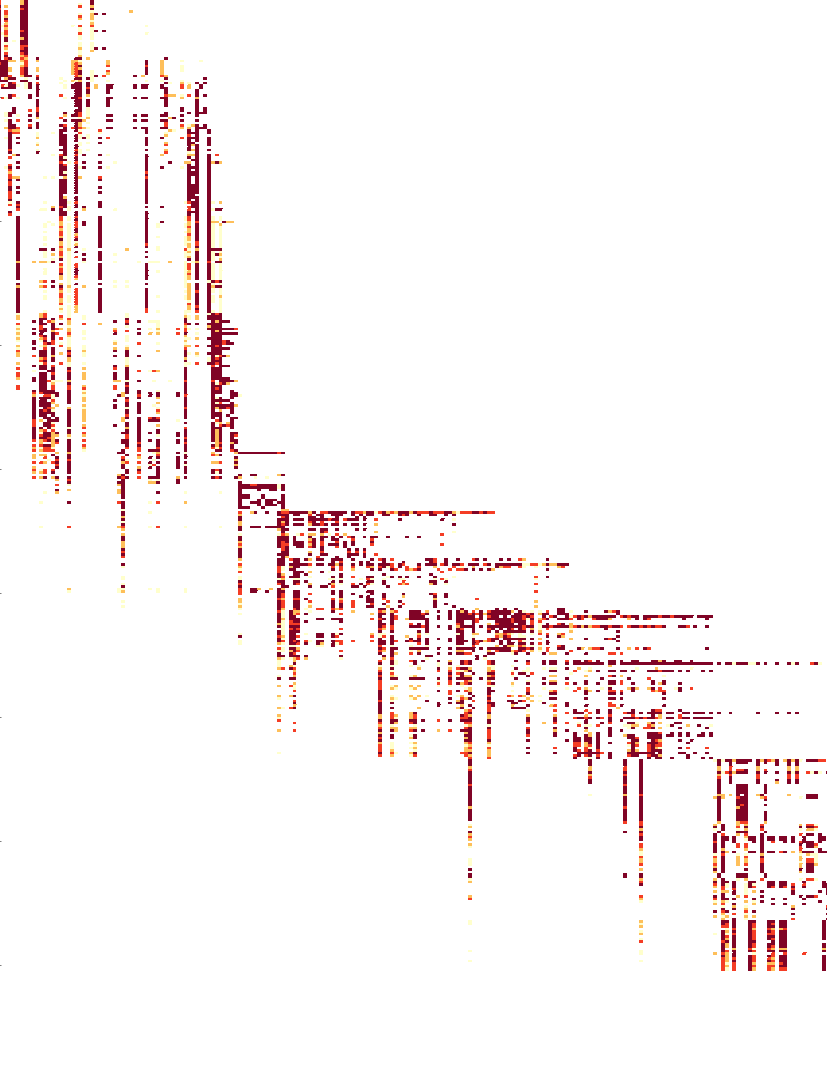
Supplementary Materials

**Figure S1. Heatmap of the A/H3N2 hemagglutination inhibition assay titer measurements.** Darker color intensity represents higher titer and white represents missing data. Each cell represents the titer between a virus (row) and a serum (column). Viruses are sorted by year from 1968 (first row) to 2011 (last row). Serum is also sorted by year from 1968 (leftmost column) to 2011 (rightmost column).

Macintosh HD:Users:charles:Documents:research:antigenic:GenoPheno:antigenic-clustering:manuscript:figures:lineart:swapIandMuNeighbor.pdf

**Figure S2. Schematic diagram of Proposal 4 - Swap Ii and μi with a neighbor**.

In this phylogeny, the brown active node is chosen to perform the swap-with-neighbor proposal. In this case, with a neighborhood step-size of 2, the set of candidates consist of all the nodes within the green neighborhood.

Macintosh HD:Users:charles:Documents:research:antigenic:GenoPheno:antigenic-clustering:manuscript:figures:lineart:walkAndBalance.pdf

**Figure S3. Proposal 5 - Performs a random walk on an active μi and balance other active μj.**

In this antigenic map, μ1 is now changed by δ, which causes the antigenic location of viruses in the brown cluster to change. To retain the absolute location of viruses in other clusters, μ2 and μ3 are adjusted by subtracting δ .

Macintosh HD:Users:charles:Documents:research:antigenic:GenoPheno:antigenic-clustering:manuscript:figures:lineart:swapAndFlipNeighbor.pdf

**Figure S4. The candidate of nodes to be flipped in Proposal 7 - Swap with a neighbor and flip another neighbor.** In the second operation that flips the indicator status of another neighbor, the list of nodes to consider include neighbors (green neighborhood) from the original pivot node (green node) and the neighbors (blue neighborhood) of the newly chosen node (diamond).

Macintosh HD:Users:charles:Documents:research:antigenic:GenoPheno:antigenic-clustering:manuscript:figures:custom:simulationNoiseLog2Titer.pdf

**Figure S5. Simulation of log2 HI titer consists of measurement noises.** For a virus i and a reference virus for serum j that belongs to the same antigenic cluster, the distribution of measurements follows *N*(10, σ2). For each antigenic transition between the two viruses, the mean measurement decreases by 2 units.

|  |
| --- |
| Macintosh HD:Users:charles:Documents:research:antigenic:GenoPheno:driver:clustering:analysisManuscript7-31-2015:H3N2:mu:mu-seed-Euclidean-mds0_1.pdf |
| Macintosh HD:Users:charles:Documents:research:antigenic:GenoPheno:driver:clustering:analysisManuscript7-31-2015:H1N1:mu:mu-seed-Euclidean-mds0_3.pdf |

**Figure S6. Distribution of the estimated Euclidean distance of antigenic jump size μ between neighbor seed strains in A/H3N2 [top] and A/H1N1 [bottom] analysis with their default κ, conditional on MCMC samples that have antigenic jump.**

/Users/charles/Documents/research/antigenic/GenoPheno/driver/clustering/analysisManuscript7-31-2015/H3N2/mu/mu-seed-Euclidean-mds0\_1.pdf

/Users/charles/Documents/research/antigenic/GenoPheno/driver/clustering/analysisManuscript7-31-2015/H1N1/mu/mu-seed-Euclidean-mds0\_3.pdf

**Sensitivity Analyses**

|  |
| --- |
| Macintosh HD:Users:charles:Documents:research:antigenic:GenoPheno:driver:clustering:analysisManuscript7-31-2015:H1N1:MCMC-diagnostics:H1N1-0.1-outGelmanRubinPlot.pdf |
| Macintosh HD:Users:charles:Documents:research:antigenic:GenoPheno:driver:clustering:analysisManuscript7-31-2015:H1N1:MCMC-diagnostics:H1N1-0.3-outGelmanRubinPlot.pdf |

**Figure S7. Distribution of Gelman-Rubin statistics of the estimated pi from the 10 replicates for A/H1N1 at κ = 0.1 (top) and at 0.3 (bottom) .**

/Users/charles/Documents/research/antigenic/GenoPheno/driver/clustering/analysisManuscript7-31-2015/H1N1/MCMC-diagnostics/H1N1-0.1-outGelmanRubinPlot.pdf

/Users/charles/Documents/research/antigenic/GenoPheno/driver/clustering/analysisManuscript7-31-2015/H1N1/MCMC-diagnostics/H1N1-0.3-outGelmanRubinPlot.pdf

|  |  |
| --- | --- |
| **κ = 0.05Macintosh HD:Users:charles:Documents:research:antigenic:GenoPheno:driver:clustering:analysisManuscript7-31-2015:H3N2:MCMC-diagnostics:H3N2-0.05-outGelmanRubinPlot.pdf** | **κ = 0.1**  Macintosh HD:Users:charles:Documents:research:antigenic:GenoPheno:driver:clustering:analysisManuscript7-31-2015:H3N2:MCMC-diagnostics:H3N2-0.1-outGelmanRubinPlot.pdf |
| **κ = 0.15Macintosh HD:Users:charles:Documents:research:antigenic:GenoPheno:driver:clustering:analysisManuscript7-31-2015:H3N2:MCMC-diagnostics:H3N2-0.15-outGelmanRubinPlot.pdf** |  |

**Figure S8. Distribution of Gelman-Rubin statistics of the estimated pi from the 10 replicates for A/H3N2 at different κ.**

/Users/charles/Documents/research/antigenic/GenoPheno/driver/clustering/analysisManuscript7-31-2015/H3N2/MCMC-diagnostics/H3N2-0.05-outGelmanRubinPlot.pdf

/Users/charles/Documents/research/antigenic/GenoPheno/driver/clustering/analysisManuscript7-31-2015/H3N2/MCMC-diagnostics/H3N2-0.1-outGelmanRubinPlot.pdf

/Users/charles/Documents/research/antigenic/GenoPheno/driver/clustering/analysisManuscript7-31-2015/H3N2/MCMC-diagnostics/H3N2-0.15-outGelmanRubinPlot.pdf

|  |  |
| --- | --- |
| Macintosh HD:Users:charles:Library:Group Containers:UBF8T346G9.Office:msoclip:clip_image005.png1 | Macintosh HD:Users:charles:Library:Group Containers:UBF8T346G9.Office:msoclip:clip_image006.png2 |
| Macintosh HD:Users:charles:Dropbox:antigenic-clustering:analysisManuscript2-12-2015:H1N1:heatmap:H1N1-C3b-mds0_1_vsS4-ordered.png3 | Macintosh HD:Users:charles:Dropbox:antigenic-clustering:analysisManuscript2-12-2015:H1N1:heatmap:H1N1-C3b-mds0_1_vsS5-ordered.png4 |
| Macintosh HD:Users:charles:Dropbox:antigenic-clustering:analysisManuscript2-12-2015:H1N1:heatmap:H1N1-C3b-mds0_1_vsS6-ordered.png5 |  |

**Figure S9. Comparing the clustering patter between the MCC tree (upper triangle) and sampled tree 1-5 (lower triangle) for A/H1N1.**

/Users/charles/Documents/research/antigenic/GenoPheno/driver/clustering/analysisManuscript1-21-2015/H1N1/heatmap/H1N1-C3b-**mds0\_1**\_vsS2-ordered.png

 Users/charles/Documents/research/antigenic/GenoPheno/driver/clustering/analysisManuscript1-21-2015/H1N1/heatmap/H1N1-C3b-mds0\_1\_vsS3-ordered.png

|  |  |
| --- | --- |
| achine generated alternative text: 」 幸 明 必 。 ー  0  ・ 厰 、 当  010  開 9  7 〕 8  8 村 2 〕 8  . 司 嗣 当 圷 当 一 一 ~ ′ 当 司 社 小 一 当 凱  38 ′ 2005  : 了 省 、  2003  爿 当 い 司 爿 当 司 司 当 当 当 当 当  ・ 事 い 環 、 ギ 当 ・  ~ 旧 」 れ 物 温 彗 れ 0 囃 弭 眦 ~ い 呶 ( 轟 耻 囃 れ 上 、  崕 箋 瞻 區 幇 を 。 。 糴 ) ~ 黐 爾  ル に ー  ・ ぶ 磆 ・  当 4 鰲 」 キ ト き を - : ド  パ こ : - 墾 ま ぎ 隴  : 嚶 巻  : = を 当 い  ょ 蠧 : 尊 鼇 噐 鑿 駲 : 蕊  ま ま ~ び  4!MS,'V  - 山 pe イ | Macintosh HD:Users:charles:Dropbox:antigenic-clustering:analysisManuscript2-12-2015:H3N2:heatmap:H3N2-C3b-mds0_1_vsS3-ordered.png |
| Macintosh HD:Users:charles:Dropbox:antigenic-clustering:analysisManuscript2-12-2015:H3N2:heatmap:H3N2-C3b-mds0_1_vsS4-ordered.png | Macintosh HD:Users:charles:Dropbox:antigenic-clustering:analysisManuscript2-12-2015:H3N2:heatmap:H3N2-C3b-mds0_1__vsS5-ordered.png |
| Macintosh HD:Users:charles:Dropbox:antigenic-clustering:analysisManuscript2-12-2015:H3N2:heatmap:H3N2-C3b-mds0_1__vsS6-ordered.png |  |

**Figure S10. Comparing the clustering patter between the MCC tree (upper triangle) and sampled tree 1-5 (lower triangle) for A/H3N2.**

/Users/charles/Documents/research/antigenic/GenoPheno/antigenic-clustering/summarizedResults/H3N2/heatmap/H3N2-C3b-mds0\_1\_vsS2-ordered.png