Phylodynamics Simulation of Phylogenetic Trees

Hsiang-Yu Yuan

Date: July 22, 2013

Version: 1.2

Date: Mar 14, 2014

Version: 3

Section:

Background

System Architecture

System Design

Package Descriptions

Program Flow

Background:

This document describes the system design and program flow of the new project based on the previous project entitled *critical analysis of binding avidity hypothesis*. The previous project ends in April, 2013 with main figures generated for thesis work. The files for the old project are mainly located in Z:/Projects/Binding\_Phy, Binding\_Evo\_Tree, etc.

System Architecture

BindingSim

InferPhylo

BindingAvidity

make\_infection\_tree

Mbindingavidity

trees

Viruses

dat\_x\_trans\_tmp.m

DataTLSIR.mat

virus\_traits.mat

Size of total population?

Initial infected individuals?

Initial binding avidity?

System Design

main\_binding\_fromeq.m

*make\_infection\_tree*

*Mbindingavidity*

Nexus, Newick, mat

Tree

main\_simulate\_genealogy\_indiv\_binding

*copy*

dat\_x\_trans\_tmp.mat

virus\_traits.mat

indiv\_treeData.mat

tree

Traits

viruses

*tau\_leap*

virus\_traits.mat

branch2pair

cal\_meanbding.m

*from selected*

*from whole*

calNetChargePhylogeny

Mean\_binding.mat

mean netcharge

plot\_simulation.m

V1.2

V3 (May 14, 2014)

Rho- = 1-1/R0

R0 = f x g x N

Package Descriptions:

* BindingAvidity: Phylodynamic simulation of viral infection using mixture of compartmental and Individual-based model.
* InferPhylo: Inferring phylogenetic trees based on finding MRCA in each lineage.

Program Flow

1. Edit parameter file params\_std\_reinfect\_larg.dat in dat/ folder (N=3x10^6)
2. Run Main\_binding\_fromeq.m
   1. Generate SIR, Traits, and Viruses.

Viruses: dat\_x\_trans\_tmp.m

SIR: DataTLSIR.mat

Traits: virus\_traits.mat

1. Copy dat\_x\_trans.tmp.mat and virus\_traits.mat to make\_infection\_tree/ folder
2. Generate viral phylogeny (n = 300)
   1. Make\_Infection\_Tree (main\_simulate\_genealogy\_indiv\_binding.m)

Input: {'births', 'deaths', 'parent'}

Output: outfile\_treeData (indiv\_genealogy\_300.tree)

indiv\_treeData

indiv\_genealogy

What are the label for each nodes?

1. branch2pair
2. Calculate netcharge distribution of virues.
   1. [netcharge infectK] = calNetChargePhylogeny('20130718/dat\_x\_trans\_tmp.mat','20130718/virus\_traits.mat');

VirusesArray(vid,:) = [vid, birth, death, parent, infectionK, beta, initialV, currentV];

BuildTree\_indiv\_v2.m

External node label:

i=1->300

names{i} = strcat('sample', int2str(i), '\_' ,num2str(indiv\_sampled(i)), '\_', num2str(seq\_times(i)));

% seq\_times are the death times of each individuals

Internal node label:

I=301->599

names{n\_tot\_samples + loc\_b} = strcat('node', int2str(loc\_b), '\_', num2str(coal\_parent) ,'\_' ,num2str(timeOfCoalescence)); % add to name

write newick tree

tree = phytree(b,d, names);

where the sizes of each variables are

b = 299

d = 599

names = 599

BuildTree\_indiv\_nexus.m

Create a new file BuildTree\_indiv\_nexus.m

The tree will be produced in nexus file format, which can be read by figtree.

\*Nexus tree contains antigenic changes.

How to model antigenic change?

To be continued..