Index to code for simulation model and fitting.

|  |  |  |
| --- | --- | --- |
| **Simulate\_baseModel.r** (or alternatively simulate\_baseModel\_rec4.r for four siblings from recombination) | | |
| Simulates households, individuals and infections according to inputs. | | |
|  |  |  |
| calls | Simulate\_household\_distances.r | Simulates households and computes distances |
|  | Simulate\_initial\_infections.r | Simulates the initial infections |
|  |  |  |
| Reads in | Allpairs.txt | Written by simulate\_household\_distances.r |
|  | Homesteads.txt | Written by simulate\_household\_distances.r |
|  | Dist\_diff.txt | Written by simulate\_household\_distances.r |
|  | dataChrPost.txt | Position of SNP on chromosome |
|  | infHousesInitInf.txt | Written by simulate\_initial\_infections.txt |
|  | obsData.csv | Observed genotypes |
|  |  |  |
| Writes to | Hospdata.txt | Simulated infections matched to time and location of observed genotypes, ready for the calculation of the likelhood |
|  |  |  |
|  |  |  |
| **Calculate\_likelihood.r** |  |  |
| Reads in | Hospdata.txt |  |
| Writes to | Para\_loglik.txt | Calculated log likelihood for one set of inputs and seed.  For the grid search, the log likelihood for multiple scenarios must be compared. We used the university cluster to run these in parallel. |

Please note:

The code inputs the distance parameter called ‘rate’ which is equal to sigma in a half-normal distribution. To gain the mean distance used sigma\*sqrt(2/pi) (which is sigma\*0.798).