1 ODD Model - Primary simulation program

1.1 Purpose

Transposable elements are dynamic and propagate in the population unless forces of selection drive them out or balance them. Alternatively, the propagation/invasion can also be checked by the presence of piRNA like a defence mechanism in the genome.
The primary purpose of this model is to describe the spread of transposon under constraints which are defence mechanism based on piRNA, selection, recombination and presence of other transposons with different characteristics. The model tracks and traces the path taken by the transposon in population after a specified number of generations, indicating the relative success of transposable element as the mean copy number in the population or loss of it.
Ultimately, the results will be used and compared with in-lab experimental evolution results to infer the characteristics of transposons and the defence mechanism involved to stop the invasion of the transposon.

1.2 Entities, state variables, and scales

1.2.1 Scales & spatial units

Generation: This is the primary time unit of an individual simulation. Generation n corresponds to generation n of the experimental evolution system. Each generation consists of entirely new individuals derived from the previous generation with no overlap of individuals between generations.

Insertion site: This is the basic unit of genome and represents an empty cell in genome where a transposon can insert.

Genome: This is a hybrid unit which is shared by every single individual in the simulation, regardless of the generation they are in and is immutable. Genome describes the genome of the organism and contains information about recombination in the form of recombination probabilities between adjacent insertion sites. It also includes information on the selection coefficient modifier and the probability of a transposon to be inserted in the site. Lastly, the genome defines the parameter τ for each insertion site - which is the penalty on excision rate if a transposon ever inserts into the site above.

1.2.2 Agents

Individual: This is the most basic unit on which simulation operates on. An individual contains information about the piRNA content, the transposon content and the overall fitness value. Two individuals in the same generation sharing the same parameters would not be distinguishable, i.e. they lack a unique identifier.

Transposon: This is the essential active agent - meaning it can influence the simulation directly and is the object which is studied in the simulation. It contains

information about the excision rate(e), the insertion rate(i), the repair rate(r) and a unique identification number. Furthermore, each transposon can be traced back to its parent after a successful replication event.

1.2.3 Collectives

Population: This encompasses every individual present in each generation. The size of the population itself can affect the results of simulation due to factors like drift.

1.3 Process overview and scheduling

–Algorithms to do after piRNA routine is done–

1.3.1 Modules

Recombination: This module manages the recombination event in diploid individual to produce a haploid gamete for mating.

Algorithm 1: Recombination algorithm

Result: A vector containing information about haploid gamete
initialization;

Transposition This module manages the transposition event in diploid individual after successful recombination and mating events.

Algorithm 2: Transposition algorithm

Result: Two vectors and a dictionary containing information about diploid gamete, its transposon content and the transposon lineage
initialization;

Fitness This module generates the fitness in a diploid individual.

Algorithm 3: Fitness algorithm

Result: A float containing fitness value
initialization;

Copy Number Check This module calculates the transposon copy number statistics per generation.

Algorithm 4: Copy number check algorithm

Result: Two float values, containing average copy number and copy number variance

initialization;

–Flowcharts to do after piRNA routine is done - this should clarify the order–

1.4 Design concepts

Basic principles The basic principles governing the model are drift, selection and transposition. The model itself does not explicitly take account of external factors like migration or mutations(i.e. inactive transposons). Transposition influences selection, whereby excessive transposition will lead to a more robust selection and will have a purifying effect. Drift will influence smaller population sizes. These underlying concepts are discussed in the literature and are well understood at an individual level[]. The primary purpose of this model, therefore, is not to investigate the selection or drift phenomenon, but to elucidate the dynamics between transposition and the piRNA based defence against transposition in the presence of principles mentioned above.

Emergence ??? - To do -

Adaptation The transposon count in an individual will vary in successive generations, and this could be considered as an adaptive trait which influences the reproductive success of the individual. Transposons themselves are influenced by their respective positions in the genome and will likely cause a decrease in the reproductive fitness of an individual and extremely unlikely cases, a reproductive boost. Another adaptive trait is the presence of piRNA. Individuals with activated piRNA would be more resistant to transposition; the presence of such piRNA can be loosely termed as the objective.

Objectives —?— To Do (it seems that this can be merged with adaptation - the final objective would be to fix the activated piRNA cluster in the population)

Learning The learning is handled implicitly by the process of selection, where an activated piRNA cluster or a beneficial transposon insertion has a relatively higher chance to be carried to the next generation.

Prediction/Sensing —?— To Do (Learning/Prediction/Sensing should be merged)

Interaction Transposons can interact with each other in the same individual with epistatic effect on fitness. Individuals themselves can interact with each other for only the purpose of mating.

Stochasticity —To do—

Collectives The population acts as a collective and while it is not treated separately as an individual unit, it influences the final results, namely the average transposon copy number. Except for interactions within population, the populations from different simulations do not interact.

1.5 Observation

–To do after piRNA

1.6 Initialisation/Input data

–To do after piRNA