The Journal of Animal & Plant Sciences, 28(6): 2018, Page: 1868-1880
ISSN: 1018-7081

INCORPORATING POPULATION GENETICS SIMULATORS IN SCIENTIFIC PLATFORM FOR CLOUD

Rabia Shahid[1], Muhammad Kashif Hanif[1*], Wesley Brewer[2], Ramzan Talib[1], and Omar Hayat[1]

[1]Department of Computer Science, Government College University, Faisalabad, Pakistan
[2]Department of Computer Science, Pyongyang University of Science and Technology, Korea
[*]Corresponding Author e-mail: mkashifhanif@gcuf.edu.pk

ABSTRACT

Geneticists employ wide range of simulation programs for the analysis of genomic dataset. Main difficulties are to deal which include complex user interface, limited hardware capacity, complex software design, and analyzing diverse genomes. Forward time population genetics is powerful technique to identify gene underlying traits or mutations of interest. In this work, population genetics simulators Mendel's Accountant and Nemo are integrated in a cloud computing environment. Wolbachia was used for platform testing. Computational analysis of Wolbachia was performed to highlight the factors that associated with prevention of arboviral diseases.

Key words: Cloud computing, Population genetics, Mendel's Accountant, Nemo, Wolbachia, Arboviral diseases.

INTRODUCTION

Population is the amount or quantity of all types of organisms including (human, plant, animal, archaeon, bacterium and fungus) of the identical species or group that live in a specific geographical environment. Population genetics is the study of allele frequency distribution and change under the influence of evolutionary processes which include natural selection, genetic drift, mutation, and gene flow. Genetic processes work in environment with an organism as behavioral and development factors (Hopf*et al.,* 2017)

Many simulation programs have been extensively used for various analysis in the field of population genetics. However, most of these simulators do not have broad potential users. The reason is development of custom systems that address a particular problem. Most of these simulators are not user friendly and difficult to export. As a consequence, most of the simulators fall in disuse in few years. The utilization of these simulators can be enhanced if they are easily accessed, configured, and operated. Moreover, most of the users of these programs are not programmers. They require virtual environment to run complex scientific simulations and user friendly interface to interpret results of simulations correctly. There is need for middleware execution platform which provides different simulation programs on single platform. Benefit of this type of platform to rapidly deploy population genetics simulations on cloud. This platform also motivates developers to use standard protocols by using standardized input formats. Moreover, this platform should provide features like real time data modeling, usability, portability, and scalability.

Cloud computing paradigm increasingly becoming popular. Large number of research organizations seek to gain value from its unique characteristic, deployment forms, and service models (Granados*et al.,* 2017). Genome sequences of different organisms are being stored in cloud environment which is accessible to researchers for various simulation studies. Cloud computing can be defined as a model for enabling ubiquitous, on demand, and convenient network used to access a shared pool of configurable computing resources. These resources provide provision of nominal configuration, minimal interaction with service provider, and less management efforts. Cloud computing provides many advantages including massive parallel execution of advanced data processing, scalability, efficiency, low cost, and data storage capacity. It also supports full range of advance computational analytics and data integration (Sharma and Jha 2015).

In this work, we present an integrated cloud based scientific platform (Brewer*et al.,* 2015), which helps to reduce setup procedure and learning curve. Researchers only needs to be familiar with one system. It also allows users to easily simulate and compare results from multiple population genetics simulators. Population genetics simulation software Mendel's Accountant and Nemo were incorporated in cloud based scientific platform. Furthermore, computational analysis of Wolbachia was performed to test the environment.

The rest of paper is organized in different sections. Section II discuss some closely related population genetics simulators. Section III outline framework infrastructure, design, and configuration. Section IV describes detail simulation of Wolbachia population on scientific platform and their role in prevention strategies of arboviral diseases. Section V

discusses results. At the end, we conclude the outcomes with future directions.

Population genetics simulators: Computer based simulations have been extensively used to evaluate and validate the influence of statistical approaches for the analysis of population genetics. Due to mounting computational influence with diverse proficiencies and capabilities, simulation programs are becoming progressively useful and powerful. Large number of simulation programs, software packages, frameworks have been used for population genetics. In this section, widely used forward time population genetics simulators are discussed.

Skelesim is population genetic simulator based on an extensible framework in R language (Parobek*et al.,* 2017). This simulation program support to choose appropriate models for simulations, setting accurate parameters, organizing all output data, and plot graphs from the summary of genetical statistics.Skelesim implements forward time and coalescent models that are accessible in the RMetaSim and in FastSimCoal2 engines. These engines are helpful to produce null distributions under the diversity of many demographic conditions for different statistics of population genetic (Parobek*et al.,* 2017).

SLiM2 is forward time population genetic simulation software (Haller and Messer 2016). This simulator has diversity of modeling techniques for composite evolutionary conditions. It delivers advanced control over maximum features of evolutionary conditions and scenarios by using R alike scripting language called Eidos. Simulation engine of SLiM2 is very optimized to qualify modeling technique of whole chromosomes for big population size. This simulation tool has user friendly graphical interface, collaborating runtime controller, and dynamic visualization features (Haller and Messer 2016).

Fwdpp is a C++ based population genetic simulation library (Thornton 2014). It supports to develop new forward time simulations under fitness and arbitrary mutation models. This library has a combination of truncated memory overhead, speed, and flexibility in modeling that are not currently offered in most of forward time simulators. This library is used for analysis of large populations and rapid implementation of innovative models (Thornton 2014).

Mendel's Accountant is forward time simulation tool that allows modeling of complex biological models (Baumgardner*et al.,* 2008). It is suitable for simulation of sexual or asexual reproduction of haploid or diploid organisms. Mendel's Accountant program can be used to build and track a single population of millions of mutations. It provides analysis of average mutation per individual, fitness and population size history, distribution of deleterious accumulated mutations, distribution of beneficial accumulated mutations, near neutral beneficial effects, near neutral deleterious effects, selection threshold history, selection effect, allele frequencies, total number of variants vs allele frequencies, and initial contrasting alleles (Baumgardner*et al.,* 2008).

Nemo is flexible genetically explicit framework for forward time population genetics (Guillaume and Rougemont 2006). It is used to study evolution by using genetic markers, phenotypic traits, and traits of life history in the meta population. Nemo implements different lifecycle and evolutionary features. This simulation program specifically used for computational analysis of species interactions between parasite and their hosts. Nemo is based on flexible model of meta population which allows specific patch carrying capacity, diffusion rate (diffusion matrix), random extinction, and demographic randomness within the framework. Large number of simulations can be run from a single parameter file with multiple parameter values. Many complex cases of evolution and population can be easily modeled on Nemo by providing time varying parameter values (Guillaume and Rougemont 2006).

SCIENTIFIC PLATFORM FOR CLOUD

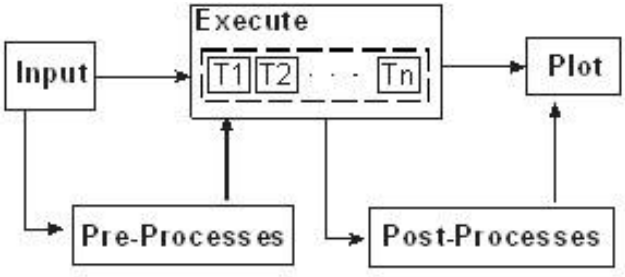

Fig.1. SPC infrastructure

There exists no platform which integrate different population genetics simulators. In this work, we.present a scientific platform for cloud (SPC) which takes advantage of cloud computing to simulate different population genetics programs. The features of the SPC platform are user friendly interface, job scheduling, job execution, job monitoring, robustness, easy migration of applications, and plotting of the results.

A. Infrastructure: SPC architecture is mainly based on Input execute Plot style software systems (Fig.1). SPC supports INI, name list, and XML input formats. Different flags can be specified on command line. Whenever the user submits a task, execution phase reads the input file and maps all the parameters to command line inputs. Post processing phase computes the subset of simulated data and convert it into Python format. The data in Python format can be used as input to plotting libraries. SPC supports two plotting libraries JQuery and Matplotlib.

SPC provides portability support to simulators. Configuration of new application on SPC requires name, description, input file format, database configuration, default input file, HTML input form template, upload and test binary file, and setup plot. Nemo configuration steps for SPC (Fig.2). SPC was designed to allow geneticists to easily upload anexecutable and sample input file on shared infrastructure and receive a transferable output file.

Fig.2. Nemo configuration on SPC.

B. Design: Design goals meets in SPC by using model view template (MVT) as shown in (Fig. 3).

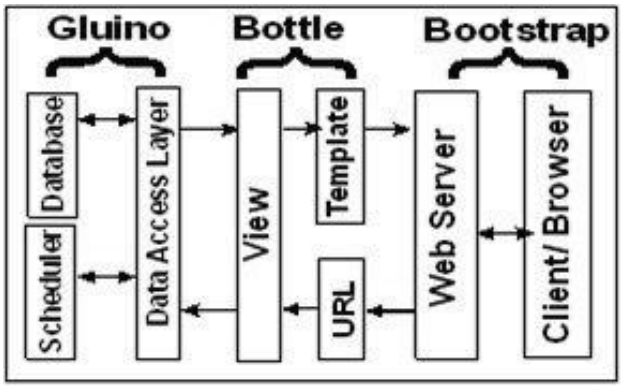

Fig.3. MVT architecture of SPC

- Bootstrap (getbootstrap.com) was used for frontend framework to build responsive projects on the web.
- Bottle (bottlepy.org) was used for micro web server side framework to map URL routes to Python methods with powerful template system.
- Gluino was used for backend framework that ported from web2py to get support of different databases like PostgreSQL, SQLite, MySQL, Oracle and many others. It is also used for jobs scheduling.
- Bottle is not full stack framework but it is easily extended by using third party plugins to *get al*l features that full stack support. Third party plugins provide features of session management, object relational mapping, and flash messages, etc. Bottle combined with separate data access layer and job scheduler.

SPC store all information about apps, jobs, users, and plots in database. SPC implemented uni and priority based multiprocessor scheduler for spawning new processes, which repeatedly polls the database every second and execute any job in front of queue. Jobs are submitted to a job table in database which maintains job states. Jobs states can be C,Q,R and X for completed, waiting in queue, running, and terminated jobs, respectively.

Design goals accomplished in scientific platform are automatic web interface configuration, execution and scheduling of jobs, managing simulations, plotting interface library, handling multiple users on shared infrastructure, facility of easy deployment on Google App Engine, Computer Engine, Amazon EC2, and on RedHatOpenShift.

C. Incorporation

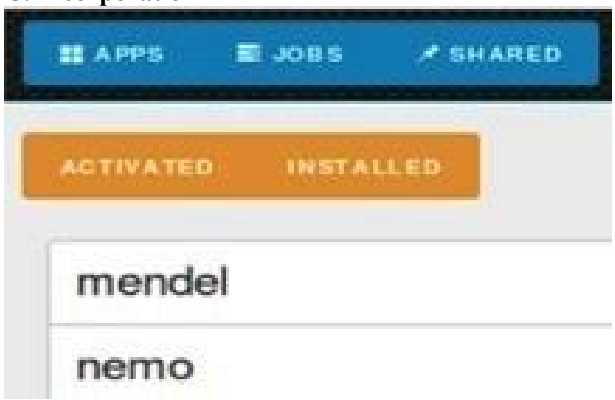

Fig.4. SPC interface after incorporation.

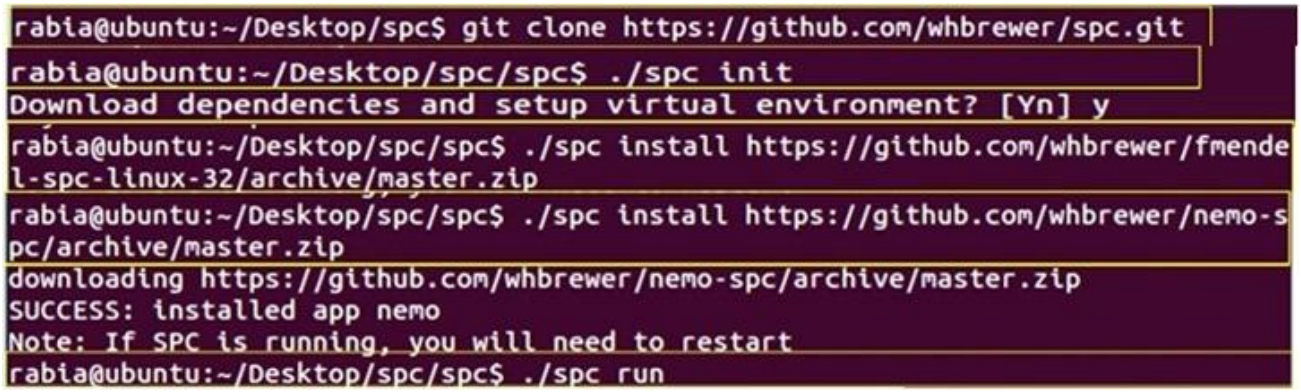

Fig.5. Incorporation steps of Mendel's Accountant and Nemo on SPC.

SPC is python based open source software. It can be easily accessed, and latest version of software can be downloaded from the give link https://github.com/whbrewer/spc. Fig.4 illustrate incorporation steps of population genetics simulators on SPC. Python 2.7 standard library was used for the establishment of scientific platform on Linux. It provides compatibility support for incorporation of additional modules or other applications on this platform.

Fig. 5 shows interface of SPC after incorporation of Mendel's Accountant and Nemo. Table1 displays the important dependencies with version for SPC.

Table1. Dependencies for SPC.

Parameters	Version	Description
Bottle	0.12.13	Bottle is simple, fast, and lightweight micro web framework for Python.
Beaker	1.9.0	Beaker is caching and web session Python library that includes webserver gateway interface for web applications.
PyYaml	3.12	PyYAML is a YAML emitter and parser for Python.
Psutil	5.2.2	Psutil is cross platform library for system and process utilities in Python.
pytoml	0.1.13	Parser for Tom's Obvious Minimal Language in Python.
DockerPy	1.1.0	Python library for Docker Engine programming interface.
Boto	2.6.0	Python based software development kit for Amazon Web Services.
Requests	2.9.1	Hypertext Transfer Protocol library for Python.
WebTest	2.0.28	Helper to test web server gateway interface applications.
CherryPy	3.2.2	Object oriented web framework.

Wolbachia population simulation: Wolbachia is genus of gram negative bacterial endosymbiont which infects invertebrates and some nematodes. Recently, Wolbachia has gained prominence specially for the prevention of arboviral diseases (Jiggins 2017). Arboviruses transfer through arthropod vectors. Arboviruses include Zika(Dutra*et al.,* 2016), Dengue (Yixin*et al.,* 2015), Chikungunya (Van den Hurk*et al.,* 2012), etc.

Mosquitoes borne viruses transmission for human diseases has been control by release of genetically modified mosquitoes which is infected with strain of Wolbachia that originally isolated from Drosophila flies. Naturally, Aedesaegypti and Aedesalbopictus mosquitoes are not infected with its own strain of Wolbachia. Wolbachia strain that extracted from Drosophila simulans is beneficial to block the transmission of arboviruses (Von Seidlein*et al.,* 2017). This controlling strategy is currently being tested, implemented, and verified in definitive trails against Zika in Colombia (Von Seidlein*et al.,* 2017), Brazil, and against dengue in Southeast Asia. Mosquitoes reproduction is controlled by introducing strain of Wolbachia (extracted from Drosophila simulans) into male mosquitoes and as a result sterilized eggs produced from mating. This method grasps some potentials against Aedesaegypti and some other similar vectors with concentrated density of wild populations with the release of the sterilized males (Benelli*et al.,* 2016). Sterile mosquito technique used in Mexico, Central and North America and proposed to be adopted for Aedesemphasisis in Brazil (Reis*et al.,* 2017).

SPC can be used to find factors association for prevention of arboviral diseases. This section discusses Wolbachia population analysis using Mendel's Accountant and Nemo on SPC. Mendel's Accountant and Nemo were used to create virtual population of Wolbachia.

Mendel's Accountant created possible mutations with precisely quantified characteristics including: frequency distribution, fraction of beneficial mutations, range of mutation effects, and ratio of dominants to recessives. Mutations were selected randomly from the pool of mutations and assigned erratically to the novel offspring in respective generations.

Table2. Important parameters used for simulation of Wolbachia in Mendel's Accountant and Nemo.

Parameters	Values
Mendel's Accountant	
Mutation Rate	0.867
Fraction Favorable Mutation	0.0001
Reproductive Rate	2.0
Population Size	1000
Number Generations	500
Genome Size	98000000
Impact Mutation Fraction (High)	0.001
Impact Mutation Threshold (High)	0.1
Fitness Gain (Max)	0.01
Heritability	0.5
Haploid Chromosome Number	4
Migration Generations	10
Nemo	
Generation	500
Patch Number	50
Patch Capacity	20
Extinction Rate	0.05
Mean Fecundity	15
Mating Proportion	0.8
Dispersal Cost	0.2
Neutral Loci	20
Neutral Mutation Rate	0.0001
Neutral Recombination Rate	0.5
Deleterious Loci	100
Deleterious Mutation Rate	0.0001
Deleterious Recombination Rate	0.5
Deleterious Effect Mean	0.05
Dispersal Mutation Rate	0.001
Dispersal Mutation Mean	0.2

These mutations were constructed on the basis of quantified average mutation rate (Poisson Distribution). Nemo was used to study evolutionary effect of phenotypic traits of Wolbachia endosymbiont. Population from generation 0 to generation 500 was considered through successive life cycle iterations. Nemo offers population models (island and lattice), traits (neutral markers, deleterious mutations and more), life cycle events (dispersal, mating, aging, selection, etc.) and saving statistics and data as output. Selection of Wolbachia was done on its phenotype or traits values. Only limited number of Wolbachia traits were selected using viability trait selection procedure.

Mendel's Accountant created possible mutations with precisely quantified characteristics. Mutations were selected randomly from the pool of mutations and allocated erratically to the novel offspring in respective generations. These mutations were constructed on the basis of quantified average mutation rate (Poisson Distribution). Nemo was used to study evolutionary effect of phenotypic traits of Wolbachia endosymbiont. Population from generation 0 to generation 500 was considered through successive life cycle iterations. Selection of Wolbachia was done upon its phenotype or traits values. Wolbachia population dynamically changed during simulation that allowing patch fission/fusion and population bottlenecks.

Genetic fitness of individuals considered as 1.0. Genetic fitness was adjusted by negative and positive effects of mutations. On the basic of phenotypic fitness, Mendel's Accountant applied different selection schemes i.e., truncation, partial truncation, strict proportional probability, and unrestricted probability schemes. For analysis, we considered only unrestricted probability selection scheme that was applied on mating pool to eliminate a quantified fraction of the individuals. Elimination of fraction factor was based on average offspring per female (fertility) of the population. Selection scheme rejects some offspring per female. Selection scheme was applied to keep persistent population size of individuals. Gametes were extracted from individuals that subsist form the selection procedure. Replicating individuals were paired off randomly and their were merged to generate the succeeding generations of individuals.

Some factors that were considered for analysis are deleterious selection trait, direct selection model, monogamy mating system, island model with propagule pool migration, and dispersal model. Table II provides important parameters of Mendel's Accountant and Nemo that were used for simulation of Wolbachia, respectively.

RESULTS AND DISCUSSION

SPC provides powerful capabilities for researcher to virtually perform migration, mating, offspring creation, and selection procedure in precise and detailed way on a single platform. Researchers can easily explore these procedures how they actually work. For platform testing detailed analysis on Wolbachia was performed to check the authentication and validity of platform. Moreover, this platform is beneficial for researchers for the detailed analysis of different vectors. Population genetics models highly depends on the mutations that are implicit to change fitness. Mutations effect on organism vary because of genomic information and types of mutations (beneficial, deleterious, neutral).

Near neutral mutations occur at high frequencies and have higher impact on fitness. Previous studies have shown that slight mutations with larger effect, and larger mutations with smaller effect. Mendel's Accountant used Weibull function for tracking fitness distribution effect.

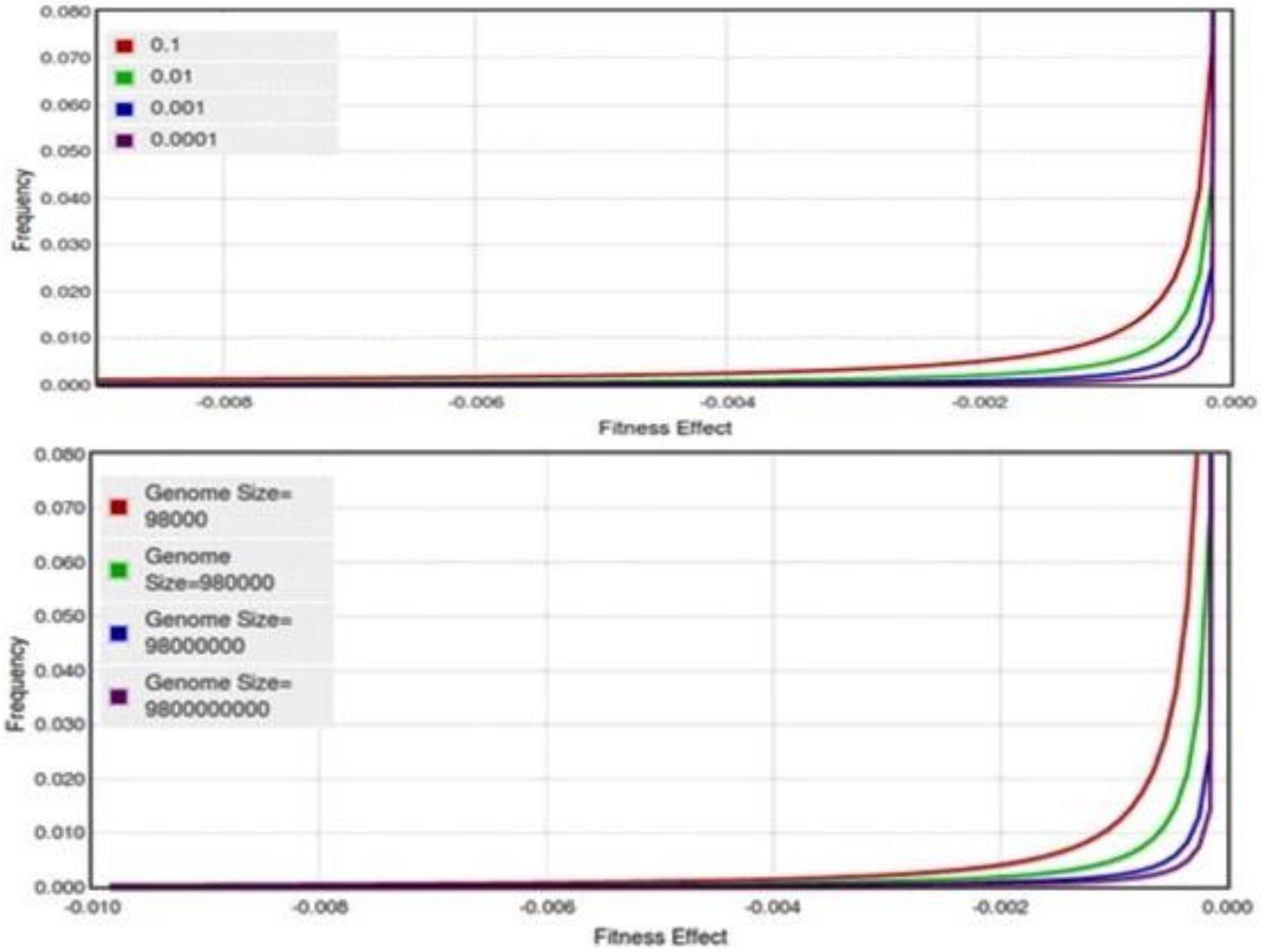

Fig.6. (a) Fitness distribution effect to change in the fraction of high impact mutations. (b) fitness distribution effect to change in haploid genome size.

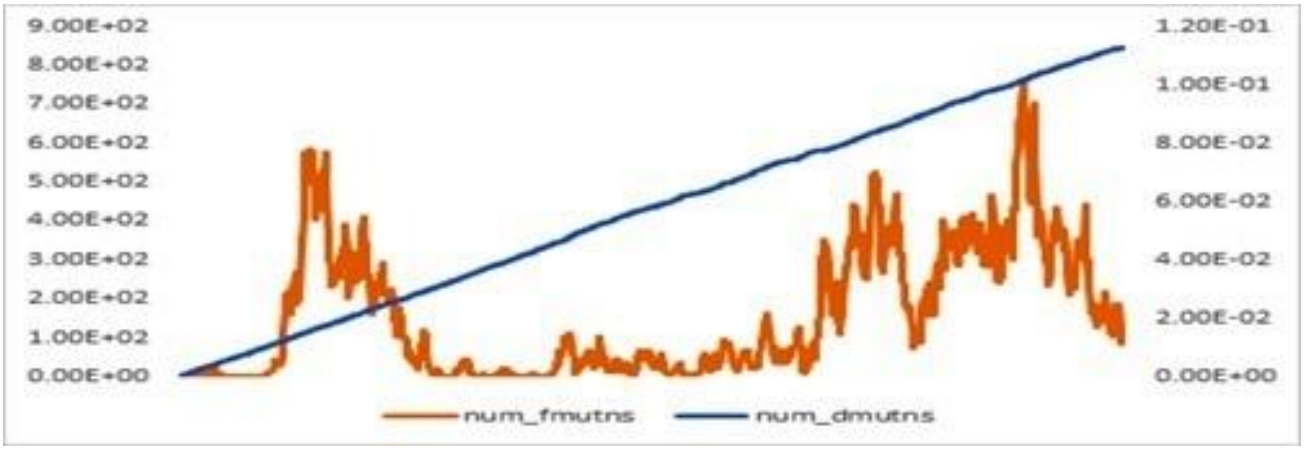

Fig. 7. Deleterious and favorable mutation effect (Mendel's Accountant).

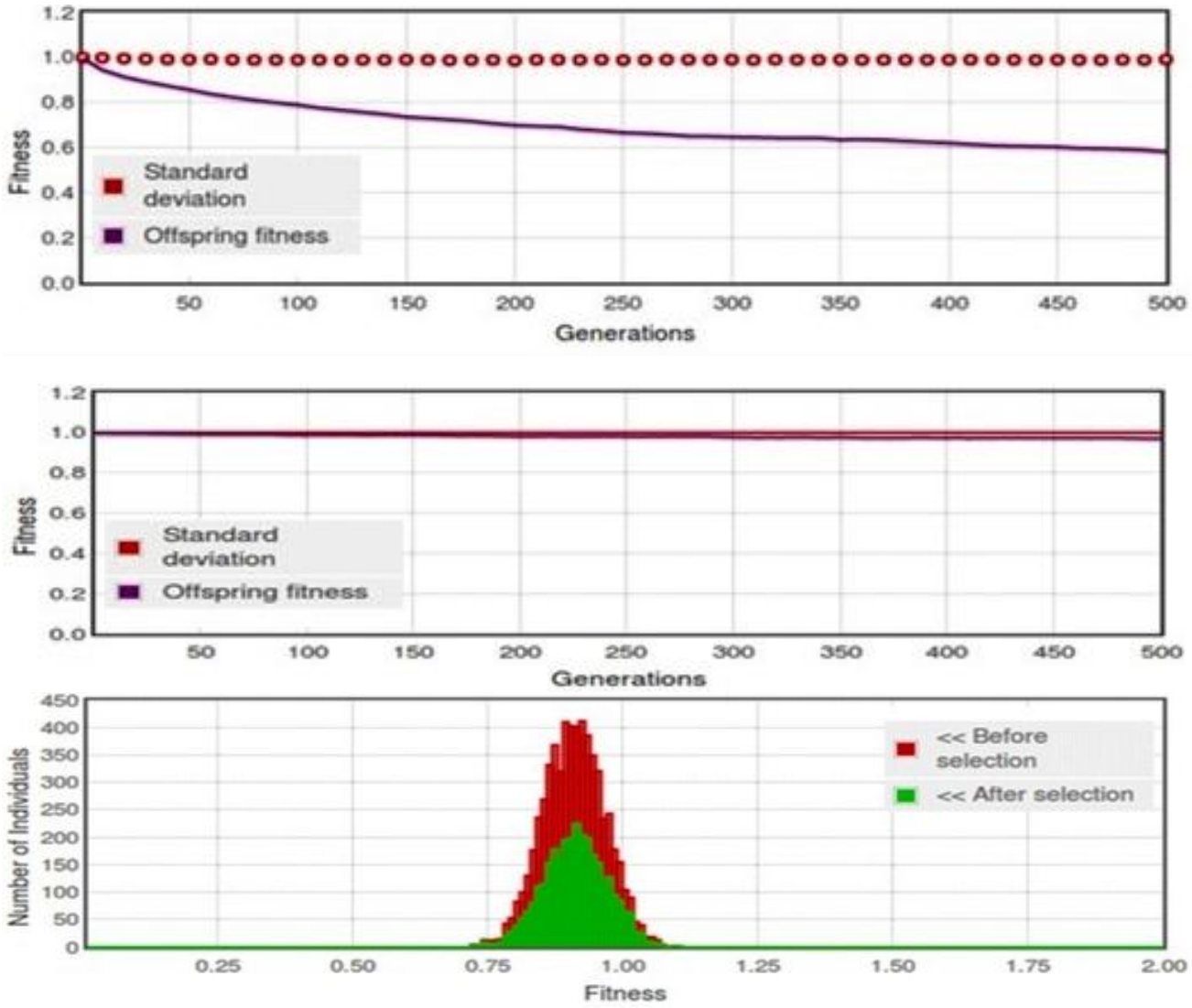

Fig. 8.(a) Fitness effect before selection procedure. (b) Fitness effectafter selection procedure. (c) Number of individual's before and after selectionprocedure.

Fig.6(a) shows the distribution function of fitness effect to change in the fraction of high impact mutations. Assumed values for high impact mutations are following: population size 1000, generations 500, frequency range 0.000 to 0.080, genome size= $9.8x10^9$, and fraction of exceeding threshold mutations 0.1 to 0.0001. Near neutral mutations always occur at higher frequencies. We assumed initial fitness value as 1.0. Minimum absolute value of fitness and fraction of mutations highly influence on the distribution of fitness for random mutations. We considered same values for deleterious and favorable mutations. Inverse of haploid genome size must be taken for finding minimum absolute value of fitness effect. Genome size measured in numbers of nucleotides. For example, for the Wolbachia genome $G=9.8x10^7$, for the computation of the case of deleterious mutations d(1)=1/G. Therefore, for larger genome size the least value must nearby to zero, and for smallest genome size the effect of minimum absolute value must be greater. In smaller genome on average, each nucleotide highly influences on the organism's fitness. Fig.6(b) represent changes in haploid genome size and their response on fitness distribution function. Considered values are population size 1000, generations 500, frequency range 0.000 to 0.080, genome sizes= 9.8x104 to $9.8x10^9$, and fraction of exceeding threshold mutations 0.001. These graphs only show small portion of effects on fitness distribution. Mostly mutations have shown their nearly zero effect. Vertical scale is the frequency of mutations per unit fitness effect. Note that mostly cases displayed larger trend of near neutral mutations so, influence of these mutations on fitness must be smaller.

In case of average mutations per individual, favorable mutation rate per offspring is considered 0.01.

We also considered size of population 3000 which is constant during reproduction, and reproduction rate before selection procedure is considered as 6 offspring per female. After selection procedure only 2 offspring are compulsory to sustain the constant size of population. In each generation, 2/3 of the offspring are carefully chosen away that do not reproduce for next generations. We choose hundred new mutation per offspring. The value of this favorable mutation is 100 times greater than expected value. Reason of choosing higher favorable rate is to easily analyzed mutation effect.

Fig. 7 shows the visual representation of favorable and deleterious mutations effect that run for 3000 population size and 500 number of generations. This shows that the effect of deleterious mutations per individual for whole population seems to be constant and that is nearby the product of elapsed generations and numeral of new mutations per offspring (100). This indicates that the average number of mutations per individuals that do not reproduce or survive must be close to the average mutations per individuals that do. Note that the increase in deleterious mutations effect is nearly constant. By natural selection, these mutations effect highlighted the fact that majority of large amount of mutations have smallest influence on fitness. Hence, the frequencies of vast majority of favorable and deleterious mutations cannot be altered during natural selection procedure phase. Because of insignificant favorable mutations, statistical fluctuations in accumulation rate are must be obvious. Increase in number of favorable mutations are relatively constant. But due to the small number of mutations the statistical discrepancies are more prominent. This type of distinction in mutations clearly represent the consequence that almost all favorable mutations and vast number of deleterious mutations are unelectable. Their effects per individual for selection are too small to detect. Selection procedure directly acts on environmental conditions and on cumulative phenotypic fitness.

Specifying the selection procedure within the population whose individual's fitness vary is the critical aspect of population genetics model. Intensity of fitness is mainly specified through fertility. Fertility define as mean number of offspring per female. Mostly, during the reproduction the size of population held constant. During the selection phase surplus offspring eliminates that beyond the limit of targeted population size. Selection procedure basically important because it discriminates individuals that will mate and replicates from those individuals that will not. Generally, worst phenotypes do not reproduce while best reproduce. Naturally environmental variations also play a vital role on individual fitness or survival. For adding environmental variation heritability parameter is used. Heritability is used as the ratio of genetics fitness variance to total fitness variance (environmental variance plus genetic fitness variance). For simulating Wolbachia we considered unrestricted probability selection that does not impose any restricted limit on the scaling factor. For maintain population size it selects sufficient offspring to mate and reproduce. By using this method, offspring whose fitness exceeding with scaled fitness is automatically selection for reproduction.

Fig. 8(a) illustrate the standard deviation and fitness effect without selection procedure within the specified population size as function of elapsed generations. Note that average fitness of population drops by 65% over 500 number of generations. From generations to generations, the inexorable reduction in fitness effect is the consequence of the persistent accumulation of deleterious mutation (unelectable). This simulation clearly approves the realism of genetic entropy.Fig.8(b) represents the means deleterious fitness effect and mutation rate for 500 generations. By using unrestricted probability selection procedure, the decline in means fitness is 0.09% per generation. By eliminating individuals with the maximum deleterious mutations, selection procedure is able to diminish the decline of fitness to only 9% for 500 generations. Selection intensity for this case is low, four out of six offspring are survived to mate and reproduce during the selection procedure. Fig. 8(c) shows the number of individual's before and after selection procedure. Selection is one of the most important phase during the genetic simulation.

Fig.9(a) displays the distribution effects of deleterious mutations that present in 500 generations. This figure represents the relationship of mutations frequencies and fitness effect. Red line shows the effect of mutation distribution before the selection procedure applied on population. Green bars denote the actual effects of mutations distribution after selection procedure applied on population. Change in distribution is very small that effect lesser in scale than 0.001. Only limited size of horizontal and vertical scales displayed. Effects of most mutations in the rightmost side. Mutations effect have been nominated away whose magnitude is larger than 0.01 lie on left side of the plot. Change in the effect of mutations between 0.0 and 0.002 is small and that can be negligible due to small influence. This graph shows that the high numbers of mutations are not being affected by presence or absence of selection procedure.

Fig.9(b) and 9(c) highlight the distribution effect of accumulated mutations beneficial and deleterious in the presence of selection procedure. Dominant feature is more prominent in both cases (beneficial and deleterious mutations effects). Change in recessive feature is almost zero. In case of deleterious mutation effect the change in the fraction of mutations that retained in genome size is from 0.0 to 1.0 and effect of mutational fitness degradation is from $1x10^{1}$ to $1x10^{7}$ but in case of beneficial mutation effect the change in the fraction of mutations that retained in genome size is from 0.0 to 2.0

and effect of mutational fitness degradation is from 1x10 ^{5} to 1x10 ^{7}. Logarithmic scale is used to emphasized the high impact effects of distribution of accumulated mutations. Graph is plotted to retained the effect of distribution of mutations under the impact of selection procedure comparative to the mutations effect that would occur in the absence of selection procedure. Change in the height of plot clearly reflect the action of selection procedure. Results of these mutation shows that low influence mutations are effectually unelectable. Note that mutations whose effect is smaller collectively they cause significant decline in the fitness of population.

Fig.10 display the occurrence frequency of individual that carries the same mutations in the population of 500 generations. This figure shows the relationship of number of alleles and alleles frequency. Highest frequency of mutations is only occurred in 1% of the population. No mutation of individual occurs for more than 10% of the population. This distribution effect shows the genetic drift. Rate of change of genetic drift is slow in the population of 3000 individuals that will reproduce for 500 generations.

Nemo was used to performed following analysis: demography of male or female offspring, dispersal rate, segregation mutations, no of alleles within whole population and within demes, deleterious mutation effect on heterozygosity of alleles, and genetic variance by using Nei and Chesser model. Demographical analysis on Wolbachia was performed by using Nemo simulation tool. Mutations and other environmental factors influenced on the demography of the Wolbachia. Fig.11 shows the demographical variations of Wolbachia in male or female offspring for 1000 generations. Survival rate of female offspring is more than male offspring that survive after selection procedure. These selected male or female offspring are capable of further mating and reproduction into next generations. Plot highlight the highest density region of female and male offspring. Note that at 100 generation the density of male offspring is higher than female offspring but this trend is entirely changed at 1000 generations. At 1000 generations the densityof male offspring is less than female offspring. Density offemale offspring is about 0.5% greater than male offspring. Result illustrate that the host dynamics are influenced by the mutational effects, contact pattern of individuals, and demographic construction of population. This result also shows that the male offspring are more influenced by the environmental change and mutations effect relative to female offspring. Dispersal rate of Wolbachia was simulated by using Nemo. Fig.12 represents the distribution or spreading of Wolbachia for 500 and 1000 generations. Dispersal rate of female offspring and male offspring and their mean is displayed in the plotted region. For 500 generation, at the start the dispersal rate of male offspring is more than the female offspring but after 80 generations the dispersal of rate of female offspring is more than male offspring. This trend remains constant till the 500 generations. Change in the dispersal rate is due to the demographical changes that influenced on the dispersal rate of the individuals. Note that the change in the dispersal rate is between 0.50 to 0.25 for 500 generations.

Fig.13 shows the segregation mutation effect in 500 generations. Segregation mutation refers to as new mutation. When segregation mutation entered into population its divides the whole population into subpopulation groups one who carry the mutation and one that do not. Segregation mutation only occur in one individual and it differentiate that individual from whole population. Segregation mutation may be transfer to more than one generations. The term demic refers to as the dispersal of population into and across an area that had been priory uninhabited by segregation mutation. This plot represents that only 5% of the demic population carries the segregation mutation effect in deleterious locus and 95% population is without the segregation mutation effect in deleterious locus. In natural selection, elimination of selective alleles is significant for the stabilization of the selection, through the exclusion of deleterious discrepancies that arise. Exclusion of deleterious alleles can easily be achieved through the single point segregation mutation that used as unit of selection.

Deme is a smaller group or sub population within the whole population that can spontaneously interbreed. Population habitually made of numerous demes that are isolated(partially) from each other.

We try to estimate the level of gene flow between demes for natural population. In this simulation, we considered “Island model” that proposed by Sewall Wright in 1931. This model specifically used for calculating gene flow of random migration between demes or subpopulation. Within demes, Fig.14 shows that the amount of gene flow at start is higher that gradually decrease, but after 100 generations the response of alleles seems to be stable. Without the selection, genetic variations among demes or subpopulation results form an equilibrium between gene flow and genetic drift. This equilibrium effect can easily be seen in figure after 120 generations.

Fig.15 highlighted the effect of deleterious mutation on the heterozygosity of alleles. Heterozygosity is phenomena of having different alleles at single or more consistent chromosomal loci for a trait. Terms refers to as ho: observed heterozygosity, htnei: expected total heterozygosity, and Fst: variance between population. Note that the effect of deleterious mutation on the observed heterozygosity is about 0.76, and effect on expected heterozygosity is about 0.99 for 100 generations. Moreover, effect of deleterious mutation on the variance of population is about 0.23 for 100 generations. Deleterious effect on heterozygosity is used

to measure the genetic variations within the population. This graph displays the influence effect of deleterious mutation on heterozygosity of alleles. Note that expected fitness is about 0.77 and observed fitness is about 0.51 for 500 generations. Observed fitness is 26% less than the expected fitness. Fst can easily be find out by: expected fitness (0.77) observed fitness (0.51). Hence, this result concludes that the effect of deleterious mutation is most important factor that highly influenced on the fitness of the population.

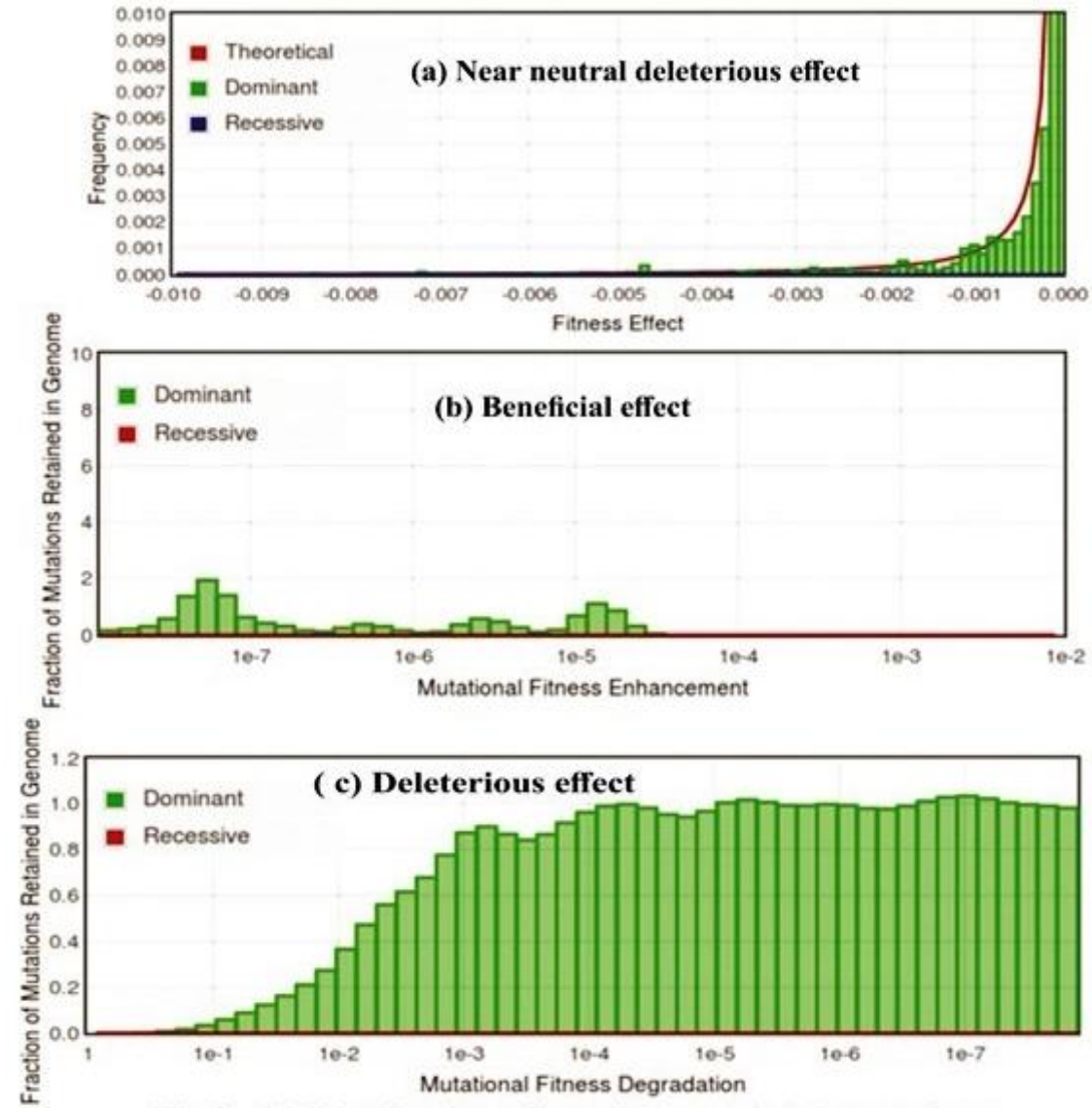

Fig.9. (a) Distribution effect of accumulated mutations (a) near neutraldeleterious. (b) Beneficial. (c) Deleterious.

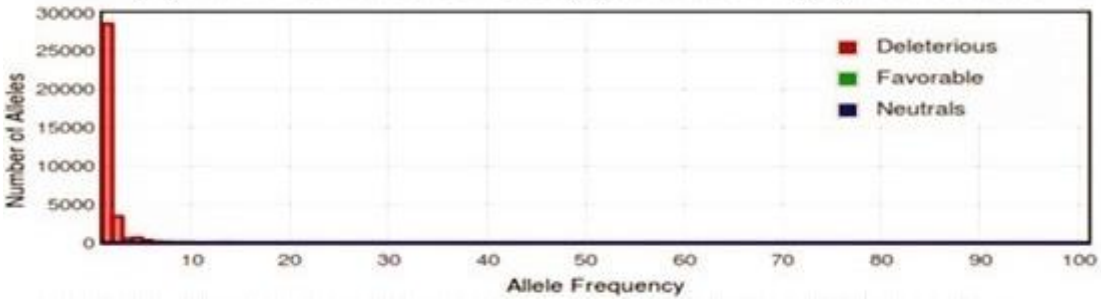

Fig. 10. Occurrence frequency of individual mutations

Wrights proposed in 1951 whose fixation indices (Fis, Fit, and Fst). According to Wrights theory these parameters are valuable for the study of genetic variance of the populations. These parameters were defined in terms of association of two uniting gametes. In natural evolutionary processes, there are numerous difficulties for finding the probability of correlations of uniting gametes, gene identities in the demes or subpopulation. In this situation, Nei's in 1977 redefined these parameters in terms of expected or observed heterozygosity and simplify the computational process. Nei's model applicable on any diploid population, with or without selection process, and with or without the presence of multiple alleles. Nei's model based on the allele frequencies of the population. Nei's model is effective for small population size.

Fig.16 shows the genetic variance analysis by using Nei and Chesser model. Parameters and their respective colors are following, ho(red): observed heterozygosity, hsnei(green): expected demicheterozygosity, htnei(blue): expected total heterozygosity, Fis(purple): genetic variance between individuals, Fit(black): genetic variance within individuals, and Fst(yellow): genetic variance between population. Note that observed heterozygosity and expected demic heterozygosity is almost same. Also, note that genetic variance within individuals and genetic variance between sub populations is nearly similar. These plotted lines show the close relationship of genetic variance within individuals between subpopulations. As we see that, observed heterozygosity and expected demic heterozygosity are smaller than the expected total heterozygosity. These lines reveal the correlation of variance in heterozygosity. Moreover, genetic variance between individuals is so smaller and it seems to be zero. This plotted graph shows the relationships that exist within the population by using statistical values and Nei and Chesser model. This graph proves that the Nei and Chesser's model is highly suitable for small population size.

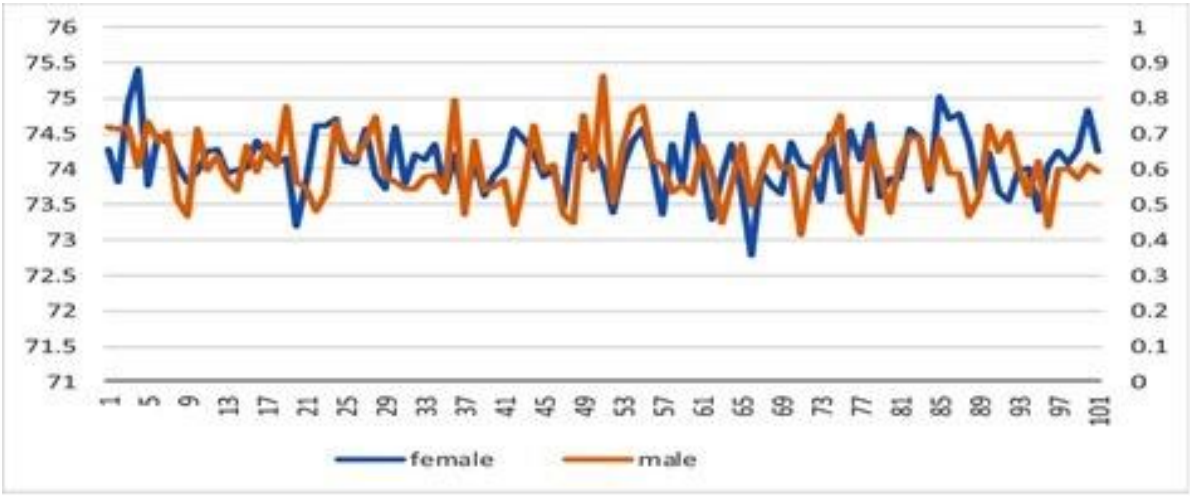

Fig. 11.Demography of male and female offspring (Nemo).

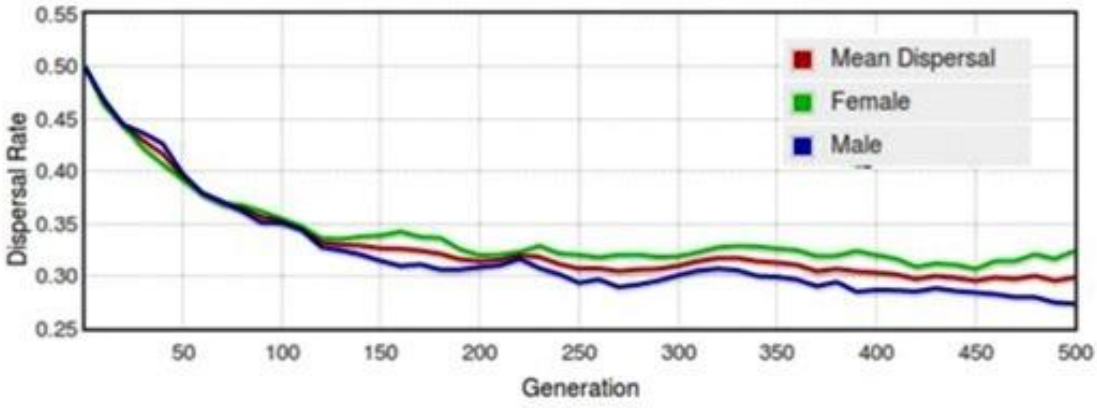

Fig. 12.Dispersal rate of male and female offspring (Nemo).

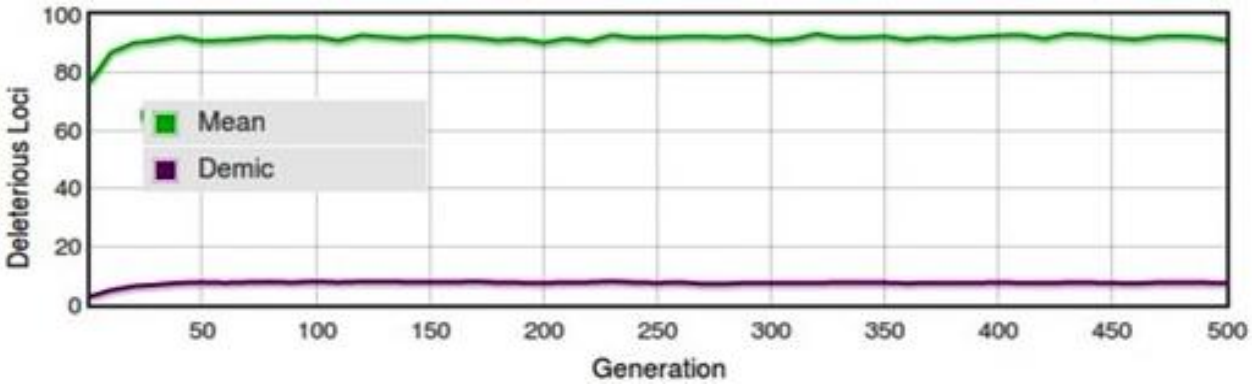

Fig. 13.Segregation mutation (Nemo).

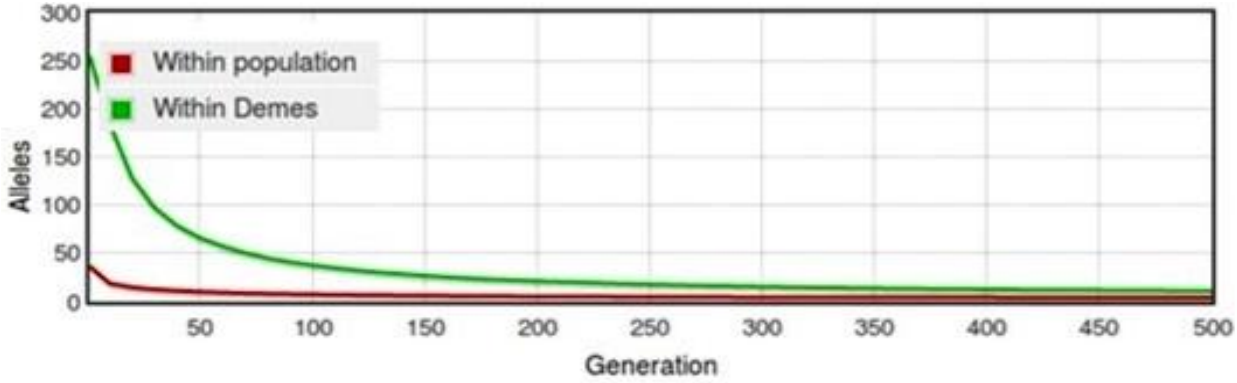

Fig. 14.No of alleles within whole population and within demes (Nemo).

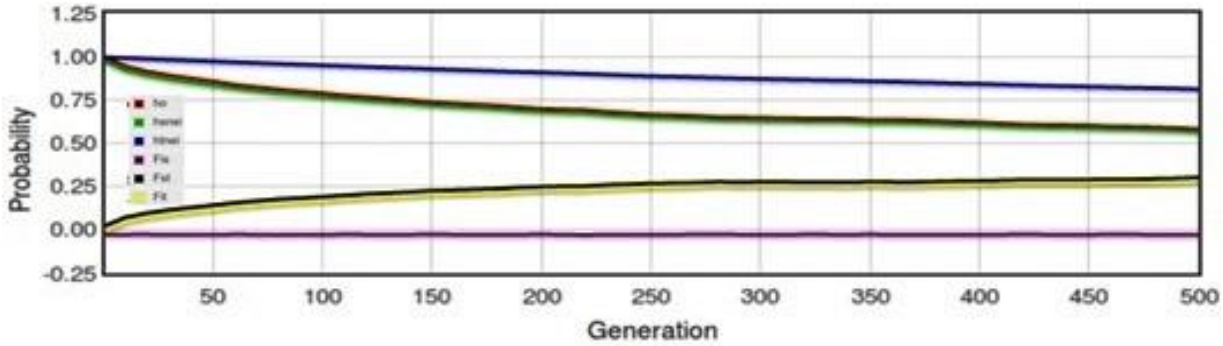

Fig. 15.Deleterious mutation effect on the heterozygosity of alleles (Nemo).

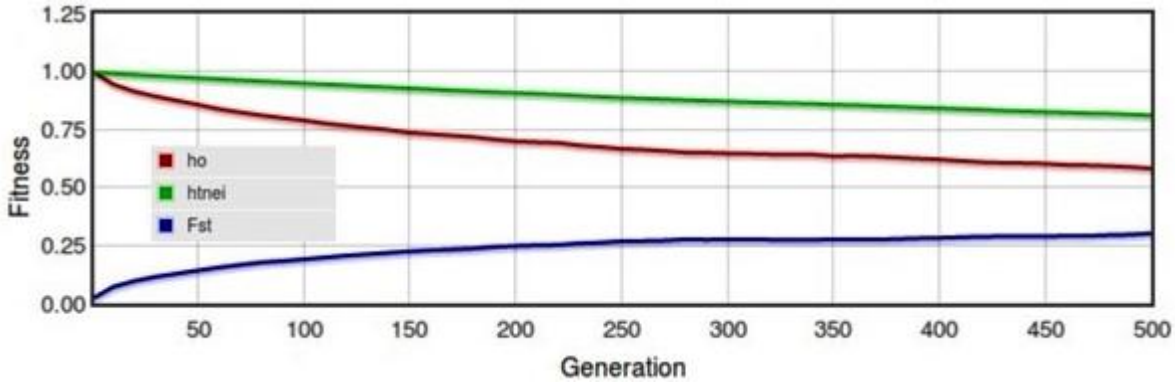

Fig. 16.Genetic variance analysis by using Nei and Chesser model (Nemo).

Conclusion: Target area of this work is development, enrichment, and incorporation of population genetic simulation tools on cloud based scientific platform. This research work highlighted different factors of Wolbachia that directly associate with prevention of arboviral diseases. Furthermore, this work also improves the features of population genetics simulation tools after incorporating them on cloud based scientific platform. However, some significant challenges need to be addressed in future including: development of advanced simulation tools and platforms for population genetics, efficient analysis of more vectors for human diseases, and advancement of prevention strategies for congenital infections.

REFERENCES

Baumgardner, J.,J. Sanford,W. Brewer,P. Gibson and W. ReMine. (2008). Mendel's Accountant: A new population genetics simulation tool for studying mutation and natural selection. Proceedings of the sixth international conference on creationism.8798.

Benelli, G.,C.L. Jeffries and T. Walker. (2016). Biological control of mosquito vectors: past, present, and future. Insects. 7(4):52.

Brewer, W.,W. Scott and J. Sanford. (2015). An Integrated Cloud Platform for Rapid Interface Generation, Job Scheduling, Monitoring, Plotting, and Case Management of Scientific Applications. Cloud Computing Research and Innovation (ICCCRI), 2015 International Conference on.156165.

Dutra, H.L.C.,M.N. Rocha,F.B.S. Dias,S.B. Mansur,E.P. Caragata and L.A. Moreira. (2016). Wolbachia blocks currently circulating Zika virus isolates in Brazilian Aedes aegypti mosquitoes. Cell host & microbe. 19(6):771774.

Granados, M.P.,Y. Joly and B. Knoppers. (2017). PublicPrivate Partnerships in CloudComputing Services in the Context of Genomic Research. Frontiers in medicine. 4:3.

Guillaume, F. and J. Rougemont. (2006). Nemo: an evolutionary and population genetics programming framework. Bioinformatics. 22(20):25562557.

Haller, B.C. and P.W. Messer. (2016). SLiM 2: flexible, interactive forward genetic simulations. Molecular biology and evolution. 34(1):230240.

Hopf, T.A.,J.B. Ingraham,F.J. Poelwijk,C.P. Schärfe,M. Springer,C. Sander and D.S. Marks. (2017). Mutation effects predicted from sequence covariation. Nature biotechnology. 35(2):128135.

Jiggins, F.M., (2017). The spread of Wolbachia through mosquito populations. PLoS biology. 15(6):e2002780.

Parobek, C.M.,F.I. Archer,M.E. DePrenger Levin,S.M. Hoban,L. Liggins and A.E. Strand. (2017). skelesim: an extensible, general framework for population genetic simulation in R. Molecular ecology resources. 17(1):101109.

Reis, N.N.,A.L. da Silva,E.P.G. Reis,F.C. e Silva and I.G.N. Reis. (2017). Viruses vector control proposal: genus Aedes emphasis. The Brazilian J. Infectious Diseases.

Sharma, A.K. and R.K. Jha. (2015). Cloud Computing. Expansion, Impact and Challenges of IT & CS. 103.

Thornton, K.R., (2014). A C++ template library for efficient forwardtime population genetic simulation of large populations. Genetics. 198(1):157166.

Van den Hurk, A.F.,S. HallMendelin,A.T. Pyke,F.D. Frentiu,K. McElroy,A. Day,S. Higgs and S.L. O'Neill. (2012). Impact of Wolbachia on infection with chikungunya and yellow fever viruses in the mosquito vector Aedes aegypti. PLoS neglected tropical diseases. 6(11):e1892.

Von Seidlein, L.,A.S. Kekulé and D. Strickman. (2017). Novel Vector Control Approaches: The Future for Prevention of Zika Virus Transmission? PLoS medicine. 14(1):e1002219.

Yixin, H.Y.,A.M. Carrasco,F.D. Frentiu,S.F. Chenoweth,N.W. Beebe,A.F. van den Hurk,C.P. Simmons,S.L. O'Neill and E.A. McGraw. (2015). Wolbachia reduces the transmission potential of dengueinfectedAedesaegypti. PLoS neglected tropical diseases. 9(6):e0003894.