DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS SERIES S
Volume **4**, Number **6**, December **2011**

doi:10.3934/dcdss.2011.4.1577

pp. **1577–1586**

# ALGEBRAIC MODEL OF NON-MENDELIAN INHERITANCE

Jianjun Paul Tian

Department of Mathematics
College of William and Mary
Williamsburg, VA 23187, USA

Abstract. Evolution algebra theory is used to study non-Mendelian inheritance, particularly organelle heredity and population genetics of Phytophthora infectans. We not only can explain a puzzling feature of establishment of homoplasmy from heteroplasmic cell population and the coexistence of mitochondrial triplasmy, but also can predict all mechanisms to form the homoplasmy of cell populations, which are hypothetical mechanisms in current mitochondrial disease research. The algebras also provide a way to easily find different genetically dynamic patterns from the complexity of the progenies of Phytophthora infectans which cause the late blight of potatoes and tomatoes. Certain suggestions to pathologists are made as well.

1. **Introduction.** In this article, we apply evolution algebra theory to the study of non-Mendelian genetics. As Mendelian genetics, non-Mendelian inheritance is a huge family in genetics. We focus on two particular genetic phenomena to show how evolution algebras work for them. One is organelle population genetics, and the other is Phytophthora infestans population genetics. A puzzling feature of organelle inheritance is how a homoplasmic cell population is established from a heteroplasmic cell population over cell divisions. Because concepts of algebraic transiency and algebraic persistency catch the essences of biological transitory and biological stability respectively, evolution algebras can be used to explain this feature. Since an algebra can have any number of mutants as its generators, algebraical modeling triplasmy in tissues of patients with sporadic mitochondrial disorders seems straightforward. We also study another type of uniparental inheritance of Phytophthora infectans which cause the late blight of potatoes and tomatoes. After we construct several relevant evolution algebras for the progeny populations of Phytophthora infectans, we can see different genetically dynamical patterns from the complexity of the progenies of Phytophthora infectans. We then predict the existence of intermediate transient races and the periodicity of the reproduction of biological stable races. Practically, we can suggest to stop the spread of late blight disease in a right phase. Theoretically, we can use our algebras to provide information of Phytophthora infectans reproduction to Plant pathologists.

The article is organized as follows. In section 2, we will recall the basic biology of non-Mendelian genetics and the inheritance of organelle genes. We also give a general algebraic formulation of non-Mendelian genetics. In section 3, we use evolution

2000 *Mathematics Subject Classification.* Primary: 17D99, 17D92; Secondary: 92D10.

*Key words and phrases.* Genetics, non-Mendelian inheritance, algebraic model, genetic algebras.

The first author is supported by NSF grant No. 0112050.

algebras to study the heteroplasmy and homoplasmy of organelle populations, and show that concepts of algebraic transiency and algebraic persistency catch essences of biological transitory and stability respectively. Coexistence of triplasmy in tissues of patients with sporadic mitochondrial disorders is also studied as well. In section 4, we apply evolution algebra theory to the study of asexual progenies of Phytophthora infestans.

## 2. **Non-Mendelian genetics and evolution algebras.**

2.1. **Non-Mendelian vs Mendelian inheritance.** Following Birky's articles [1][2], non-Mendelian genetics has five components contrasting to Mendelian genetics:

- **(1):** During asexual reproduction, alleles of nuclear genes do not segregate: heterozygous cells produce heterozygous daughters. This is because all chromosomes in nuclear genomes are replicated once and only once in interphase and mitosis ensures that both daughter cells get one copy of each chromosome. In contrast, alleles of organelle genes in heteroplasmic cells segregate during mitotic as well as meiotic divisions to produce homoplasmic cells. This is because in the vegetative division of the organelles, some copies of the organelle genome can replicate more than others by chance or in response to selective pressures or intrinsic advantages in replication, and alleles can segregate by chance.
- **(2):** Alleles of a nuclear gene always segregate during meiosis, with half of the gametes receiving one allele and half the other. Alleles of organelle genes may or may not segregate during meiosis; the mechanisms are the same as for vegetative segregation.
- **(3):** Inheritance of nuclear genes is biparental. Organelle genes are often inherited from only one parent, uniparental inheritance.
- **(4):** Alleles of different nuclear genes segregate independently. Organelle genes are nearly always on a single chromosome and recombination is often severely limited by uniparental inheritance or failure of organelles to fuse and exchange genomes.
- **(5):** Fertilization is random with respect to the genotype. This is the only part of Mendel's model that applies to organelle as well as nuclear genes.

While most of heredity of nuclear genes obeys Mendel's laws, the inheritance of organelle is not Mendelian. Vegetative segregation is the most general characteristic of the inheritance of organelle genes, occurring in both mitochondria and chloroplasts in all individuals or clones of all eukaryotes. In other words, uniparental inheritance is a major means of genetic transmission.

2.2. **Algebraic formulation of non-Mendelian inheritance.** Let us consider a population of organelles in a cell or a cell clone, and suppose that there are $n$ different genotypes in this organelle population. Denote these genotypes by $G_1$, $G_2$, $\ldots$, $G_n$. By non-Mendelian inheritance component (3), the crossing of genotypes is impossible since it is uniparental inheritance. Mathematically, we take

$$G_i \cdot G_j = 0,$$

for $i \neq j$. By non-Mendelian inheritance component (2), alleles of organelle genes may or may not segregate during meiosis following vegetative segregation, so that the frequency of each gene in the next generation can vary. From non-Mendelian

inheritance component (4), intramolecular and intermolecular recombination within a lineage provides evidence that one organelle genotype can produce other different genotypes. Therefore, by the component (5), we mathematically define,

$$G_i^2 = \sum_{i=1}^{n} \alpha_{ij} G_j,$$

where $\alpha_{ij}$ is positive number that can be interpreted as the rate of genotype $G_j$ produced by genotype $G_i$. We then have an algebra over real number field defined by generators $G_1$, $G_2$, ..., $G_n$ which are subject to these relations [3].

In [3], a general theory has been established for this type of algebra. We will use this theory to study concrete examples in the rest of the paper. As to algebras for Mendelian genetics, the reader may refer [4] [5][6][7].

## 3. **Algebras of organelle population genetics.**

3.1. **Heteroplasmy and homoplasmy.** Organelle population geneticists are usually concerned about that there are two different phenotypes or genotypes: homoplasmic and heteroplasmic. Let us denote a heteroplasmic cell by $G_0$, and two different types of homoplasmy by $G_1$ and $G_2$. Suppose $G_1$ and $G_2$ are mutant and wild-type respectively. From organelle inheritance, we know that heteroplasmic parent can produce both heteroplasmic progeny and homoplasmic progeny, and homoplasmic parent can only produce homoplasmic progeny with the same type if mutation is not considered. We then have the following reproduction relations.

$$G_0^2 = \pi G_0 + \alpha G_1 + \beta G_2, \tag{1}$$

$$G_1^2 = G_1, \tag{2}$$

$$G_2^2 = G_2; \tag{3}$$

and for $i \neq j$, $i, j = 0, 1, 2$,

$$G_i \cdot G_j = 0; \tag{4}$$

where $\pi$, $\alpha$, $\beta$ are all positive real numbers. These numbers can be taken as the segregation rates of corresponding types. For a specific example, these coefficients can be determined by modified Wright-Fisher model.

Thus, we have an evolution algebra, denoted by $A_h$, generated by $G_0$, $G_1$ and $G_2$ and subject to the defining relations (1)-(4).

From evolution algebra theory [3], algebraic generator $G_0$ is transient, $G_1$ and $G_2$ are persistent. Because $G_1$ and $G_2$ do not communicate, $A_h$ has two simple subalgebras generated by $G_1$ and $G_2$ respectively. Biologically, $G_0$ is transitory as N. W. Gillham pointed out [8]. $G_1$ and $G_2$ are of stable homoplasmic cell states. By transitory, biologists mean that the cells of transitory are not stable, they are just transient phases, and they will disappear eventually after certain cell generations. This property is imitated by algebraic transiency. By biological stability, we mean it is not changeable over a period of time, and it is kept the same from generation to generation. This property is imitated by algebraic persistency.

The puzzling feature of organelle heredity is that heteroplasmic cells eventually disappear and the homoplasmic progenies are observed. The underlying biological mechanisms are still unknown. Actually, it is a intensive research field currently since it is related to aging and many other diseases caused by mitochondrial mutations [9],[10]. However, we can mathematically understand this phenomenon from evolution algebra theory. Because $G_0$ is transient, $G_1$ and $G_2$ are persistent, $A_h$

can eventually have two simple subalgebras. These two subalgebras are of zero-th in the hierarchy of this algebra, and thus they are stable. The subalgebra generated by $G_1$ is homoplasmic and mutant; the subalgebra generated by $G_2$ is homoplasmic and wild-type. Moreover, from a formula in [3], the mean time $T_h$ to reach these homoplasmic progeny is given by

$$T_h = \frac{1}{1-\pi}.$$

If we consider a mutant to be lost, say $G_2$ will be lost, we have the following several ways to model this phenomenon. The algebraic generator set still is $\{G_0, G_1, G_2\}$.

First, we think that $G_2$ disappears in a dramatic way, that is

$$G_2^2 = 0.$$

Other defining relations are (1), (2) and (4). Thus, the evolution algebra we get here is different from $A_h$. It has one non-trivial simple subalgebra corresponding to homoplasmic progeny generated by $G_1$.

Second, we consider that $G_2$ gradually mutates back to $G_1$, that is

$$G_2^2 = \eta G_1 + \rho G_2,$$

where $\eta$ is not zero and could be 1. And other defining relations are (1), (2) and (4). Although we eventually have one simple subalgebra by this relations, the evolution path is different.

Third, we consider that $G_2$ always keeps heteroplasmic property, that is

$$G_2^2 = \eta G_0 + \rho G_2.$$

The other defining relations are still (1), (2) and (4). Eventually, we have homoplasmic progenies which all are $G_1$. That is the only simple subalgebra generated by $G_1$.

In conclusion, we have four different evolution algebras derived from the study of homoplasmy. They are not the same in the skeleton classification of algebras. Therefore, their dynamics, which is actually genetic evolution processes, are different. However, we need to look for what are the biological evidences for defining these different algebras. In Ling et al [10], several hypothetical mechanisms were put forward for establishment of homoplasmy, and their hypothetical mechanisms exactly correspond to four different algebraic structures we get.

3.2. **Coexistence of triplasmy.** In mitochondrial genetics, if we consider different genotypes of mutants instead of just two different phenotypes of homoplasmy and heteroplasmy, we will have higher dimensional algebras that contain more genetic information. Recently, in Tang et al [11], it studies the dynamical relationship among wild-type and rearranged mtDNAs.

Large-scale rearrangements of human mitochondrial DNA (including partial duplications and deletion) are found to be associated with a number of human disorders, including Kearns-Sayre syndrome, progressive external ophthalmoplegia, Pearson's syndrome, and some sporadic myopathies. Each patient usually harbors a heteroplasmic population of wild-type mitochondrial genomes (wt-mtDNA) together with a population of a specific partially deleted genome ($\Delta$-mtDNA) in clinically affected tissues. These patients also harbor a third mtDNA species, a partial duplication (dup-mtDNA), as well. To study the dynamical relationship among these genotypes, authors of paper [11] cultured cell lines from two patients. After a

long-term (6 months, 210-240 cell divisions) culture of homoplasmic dup-mtDNAs from one patient, they found the culture contained about 80% dup-mtDNA, 10% wt-mtDNA, and 10% $\Delta$-mtDNA. After a long-term culture of the heteroplasmic which contains wt-mtDNA and $\Delta$-mtDNA from the same patient, they did not find any new cell species, although there were the fluctuations of percentages of these two cell populations. From this same patient, after cultured two year of $\Delta$-mtDNA cell line, they did not find any new cell species. Now, let's formulate this genetic dynamics as an algebra.

Denote triplasmic cell population by $G_0$ that contain dup-mtDNA, wt-mtDNA and $\Delta$-mtDNA, heteroplasmy that contains dup-mtDNA and wt-mtDNA by $G_1$, heteroplasmy that contains dup-mtDNA and $\Delta$-mtDNA by $G_2$, heteroplasmy that contains wt-mtDNA and $\Delta$-mtDNA by $G_3$, homoplasmy dup-mtDNA by $G_4$, homoplasmy wt-mtDNA by $G_5$, homoplasmy $\Delta$-mtDNA by $G_6$. According to the genetic dynamical relations described above, we set algebraic defining relations as follows:

$$\begin{aligned}
&G_0^2 = \beta_{00}G_0 + \beta_{01}G_1 + \beta_{02}G_2 + \beta_{03}G_3, \quad G_1^2 = \beta_{14}G_4 + \beta_{15}G_5,\\
&G_2^2 = \beta_{24}G_4 + \beta_{26}G_6, \quad G_3^2 = \beta_{35}G_5 + \beta_{36}G_6, \quad G_5^2 = \beta_{54}G_4 + \beta_{56}G_6,\\
&G_4^2 = \beta_{44}G_4 + \beta_{45}G_5 + \beta_{46}G_6, \quad G_6^2 = \beta_{64}g_4 + \beta_{65}G_5,\\
&G_i \cdot G_j = 0, \quad i \neq j, \quad i,j = 0,1,\ldots,6.
\end{aligned}$$

And the generator set is $\{G_0, G_1, \ldots, G_6\}$. This algebra has three levels in its hierarchy. On the 0-th level, it has one simple subalgebra generated by $G_4$, $G_5$ and $G_6$. These three generators are algebraic persistent. Biologically, they consist of genotypes that can be observed, and genetically stable. On the 1-st level, it has three subalgebras, and each of them has dimension 1. On the 2-nd level, there is one subalgebra generated by $G_0$. Generators on the 1-st and 2-nd levels are all algebraic transient. They are unobservable biologically.

If we have more information about the reproduction rates $\beta_{ij}$, we could quantitatively compute certain relevant quantities. For example, let's set

$$\begin{aligned}
&\beta_{00} = \beta_{01} = \beta_{02} = \beta_{03} = \frac{1}{4}, \quad \beta_{14} = \beta_{15} = \frac{1}{2}, \quad \beta_{24} = \beta_{26} = \frac{1}{2},\\
&\beta_{35} = \beta_{36} = \frac{1}{2}, \quad \beta_{44} = \frac{5}{6}, \quad \beta_{45} = \beta_{46} = \frac{1}{12}, \quad \beta_{54} = \frac{2}{3}, \quad \beta_{56} = \frac{1}{3},\\
&\beta_{64} = \frac{2}{3}, \quad \beta_{65} = \frac{1}{3}.
\end{aligned}$$

We then can compute the long-term frequencies of each genotype in the culture. Set $\Theta = \sum_{k=0}^{6} G_k$, define $\Theta^n * v = \Theta^{n-1}(\Theta v)$ for any element in this algebra. Suppose we start with a transient genotype $G_0$, as time goes to infinity, we have

$$\lim_{n\to\infty} \Theta^n * G_0 = 0.8G_4 + 0.1G_5 + 0.1G_6.$$

Therefore, to this patient, we get an algebraic structure of his mitochondrial genetic dynamics. Besides the experimental results, we can predict that there are several transient phases. These transient phases are algebraic transient elements. They are important for medical treatments. If we could have drug to stop the transition during the transient phases of mitochondrial mutations, we could help these disorder patients.

4. **Algebraic structures of asexual progenies of Phytophthora infestans.** In this section, we apply evolution algebra theory to the study of algebraic structure of asexual progenies of Phytophthora infestans based on experimental results in paper [12]. The basic biology of Phytophthora infestans and related experiments are first briefly introduced. Then we will construct evolution algebras for each race of Phytophthora infestans. Most of our biological materials are taken from paper [12]and [13].

4.1. **The basic biology of Phytophthora infestans.** The organism P. infestans (Mont.) de Bary, the cause of potato and tomato late blight, is the most important foliar and tuber pathogen of potato worldwide. Virulence variability in P. infestans populations is recognized as a major reason for the failure of race specific genes for resistance in cultivated potato management strategy. The race for P. infestans refers to possession of certain virulence factors. Isolates sharing the same virulence factors are considered to be a race that can be distinguished from other races possessing other groups of virulence factors. Characterization of isolates to different races is based on their interaction with major genes for resistance in potato. So far 11 major genes for resistance have been identified in Solanum spp [12].

In paper [12], five parental isolates of P. infestans, PI-105, PI-191, PI-52, PI-126 and PI-1, collected from Minnesota and North Dakota in 1994 to 1996, were chosen to represent different race structures. Single zoospore progenies were generated from each of the parental strains. The parental isolate PI-1 produced very small zoospores and the percent recovery of colonies was very low. Other parental isolates produced large sized zoospores and showed higher levels of developed colonies. In total, 102 single zoospore isolates were recovered, 20 isolates from isolate PI-105, 29 isolates from PI-191, 28 isolates from PI-52, 14 isolates from PI-126, and 11 isolates from PI-1. These single zoospore demonstrated different levels of variability for virulence. Although some single zoospore isolates showed the same virulence as their parental isolate, others showed lower or higher virulence than the isolate from which they were derived. Single zoospore isolates derived from PI-1 (11 isolates) were identical in virulence to their parental isolate. Single zoospore isolates derived from isolate PI-191 (29 isolates) showed low levels of variability for virulence compared with their parental isolate; 73% of these isolates (21 isolates) retained the same virulence pattern as their parent. Four isolates gained additional virulence to R8 and R9. One isolate had additional virulence to R9 which was stable. The other two showed lower virulence compared with the parental isolate. Six races were identified from the single zoospore isolates of the parental isolate PI-191.

Single zoospore isolates derived from isolate PI-126 showed higher levels of variability for virulence. Three isolates in this series gained virulence to both R8 and R9, three isolates gained additional virulence to R8, six isolates gained additional virulence to R9, and only two isolates retained the same virulence level of the parental isolate. Four races were identified within this series of isolates.

Isolates derived from the parental isolate PI-52 were highly variable for virulence. The overall trend in this series of isolates was toward lower virulence relative to the parental isolate. The total number of races identified from this parental isolate is 12.

The single zoospore progeny isolates derived from isolate PI-105 were highly variable for virulence. In this series of isolates there was a tendency for reduced virulence of the single zoospore isolates compared with their parent. 13 races were identified from this set of isolates.

4.2. **Algebras of progenies of Phytophthora infestans.** In order to mathematically understand the complexity of structure of progenies of P. infestans, we assume that there are 11 loci in genome of P. infestans corresponding to the resistant genes, denote by $\{c_1, c_2, \ldots, c_{11}\}$, and if $c_j$ functions (is expressed), the progeny resists gene $R_j$. Any non-repeated combination of these $c_j$ can forms a race mathematically. So, we have 2048 races. For simplicity, we just record a virulence part of a race by $E_i$, the complement part is avirulence. For example, $E_i = \{c_2, c_3, c_5, c_8, c_{10}\}$ represents race type $c_2c_3c_5c_8c_{10}/c_1c_4c_6c_7c_9c_{11}$. Take these 2048 races as generators set, we then have a free algebra over real number field $R$. Since reproduction of zoospore progeny is asexual reproduction, the generating relations among races are evolution algebra types. That is,

$$E_i^2 = \sum p_{ij} E_j,$$

and if $i \neq j$

$$E_i \cdot E_j = 0,$$

where $p_{ij}$ are non-negative numbers. If we interpret $p_{ij}$ as frequency, we have $\sum p_{ij} = 1$. As an example, let's look at the race PI-126P and its progenies. PI-126P has race type $E_1 = \{c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_{10}, c_{11}\}$. It has four different type of progenies:

$$\begin{aligned} E_2 &= \{c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8, c_{10}, c_{11}\}, \\ E_3 &= \{c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_9, c_{10}, c_{11}\}, \\ E_4 &= \{c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8, c_9, c_{10}, c_{11}\}, \end{aligned}$$

and $E_1$ itself. Actually, these four types of progenies are those that are biologically stable, and can be observed as outcomes of asexual evolution. These four types of progenies are persistent. There could have many transient elements that produce biologically unstable progenies. These progenies serve as intermediate transient generations, and produces stable progenies. However, the simple evolution algebra without intermediate transient generations that we can construct for race PI-126P have the following defining relations:

$$\begin{aligned} E_1^2 &= p_1 E_2 + q_1 E_3, \quad & E_2^2 &= p_2 E_1 + q_2 E_4, \\ E_3^2 &= p_3 E_1 + q_3 E_4, \quad & E_4^2 &= r_1 E_1 + r_0 E_4; \end{aligned}$$

and if $i \neq j$,

$$E_i \cdot E_j = 0.$$

If we know the frequency of the $jth$ race in the population as in paper [12], we can easily set above coefficients. For example, suppose all coefficients have the same value, 0.5, then the algebra generated by PI-126P is a simple evolution algebra. Biologically, this simple evolution algebra means that each race can reproduce other races within one population. We can also compute that the period of each generator, for each race, is 2. This means to reproduce any race itself at least needs two generations. Eventually, the frequency of races $E_1$, $E_2$, $E_3$ and $E_4$ in the whole population are $\frac{1}{3}, \frac{1}{6}, \frac{1}{6}$ and $\frac{1}{3}$ respectively. This can be done by computing

$$\lim_{n \to \infty} \Theta^n * E_1,$$

where $\Theta = \sum_{i=1}^{4} E_i$.

Now, let's assume there exists an intermediate transient generation, so there exists a transient race, $E_5$, in the developing process of progeny population of

PI-126P. We just assume $E_5$ is $\{c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_{10}, \}$. Usually, it is very difficult to observe the transient generation biologically. Our evolution algebra is now generated by $E_1$, $E_2$, $E_3$, $E_4$ and $E_5$. The defining relations are given

$$E_1^2 = p_1E_2 + q_1E_3, \quad E_2^2 = p_2E_1 + q_2E_4 + r_2E_5,$$
$$E_3^2 = p_3E_1 + q_3E_4, \quad E_4^2 = r_1E_1 + r_0E_4, \quad E_5^2 = 0,$$

and if $i \neq j$,

$$E_i \cdot E_j = 0.$$

We can verify that this evolution algebra has a simple subalgebra, which is the algebra in the above example. We can also claim that intermediate transient races will extinct, and they are not biologically stable, when the parental race is within its progeny population. Mathematically, these intermediate transient races are nilpotent elements.

The progeny population of PI-52P displays a distinct algebraic feature. There are 12 races in the progeny population of PI-52P, and the parental race is not in the population. We name these races as follows. According to paper [12]: $E_0$=$\{c_3, c_4, c_7, c_8, c_{10}, c_{11}\}$ which is parental race, and the progenies are:

$$E_1 = \{c_3, c_7, c_{10}, c_{11}\}, \quad E_2 = \{c_{10}, c_{11}\}, \quad E_3 = \{c_1, c_3, c_7, c_{10}, c_{11}\},$$
$$E_4 = \{c_3, c_{10}, c_{11}\}, \quad E_5 = \{c_1, c_2, c_3, c_{10}, c_{11}\}, \quad E_6 = \{c_2, c_4, c_{10}, c_{11}\},$$
$$E_7 = \{c_1, c_{10}, c_{11}\}, \quad E_8 = \{c_7, c_{11}\}, \quad E_9 = \{c_7, c_{10}, c_{11}\},$$
$$E_{10} = \{c_3, c_4, c_7, c_{10}, c_{11}\}, \qquad E_{11} = \{c_1, c_3, c_4, c_7, c_{10}, c_{11}\},$$
$$E_{12} = \{c_2, c_3, c_4, c_{10}, c_{11}\}.$$

Thus, our evolution algebra is generated by $E_0$, $E_1$, …, $E_{12}$. Although we need the detailed biological information for frequencies of each race in progeny population, $E_0$ must be a transient generator, an intermediate transient race in the progeny population. All other generators must be persistent generators, biologically stable races that can be observed in experiments. For illustration, we give the defining relation set as follows:

$$E_0^2 = \sum_{i=1}^{12} \frac{1}{12}E_i, \qquad E_1^2 = \frac{1}{2}E_1 + \frac{1}{2}E_2,$$

for $2 \leq j \leq 11$,

$$E_j^2 = \frac{1}{3}E_{j-1} + \frac{1}{3}E_j + \frac{1}{3}E_{j+1},$$

and for $j = 12$,

$$E_{12}^2 = \frac{1}{2}E_{11} + \frac{1}{2}E_{12};$$

and if $i \neq j$,

$$E_i \cdot E_j = 0.$$

This algebra is not simple. But it has a simple subalgebra generated by $\{E_1, E_2, \ldots, E_{12}\}$. We know that this subalgebra forms a progeny population of parental race PI-52P. This subalgebra is aperiodic, which means each race in progeny population can reproduce itself in the next generation. By computing

$$\lim_{n\to\infty} \Theta^n * E_0,$$

we get that in the progeny population, frequency of parental race $E_0$ is 0, frequency of race $E_1$ and $E_{12}$ is 5.88%, frequency of race $E_2$, $E_3$, …, $E_{11}$ is 8.82%. This is asymptotic behavior of the evolution operator.

Now we consider some intermediate transient races, biological unstable races. Suppose we have two such races, $E_\alpha$ and $E_\beta$. Theoretically, there are many ways to build an evolution algebra with these two transient generators based on the above algebra with biology information; each way will carry different biological evolution information. Here, we choose the following way to construct our evolution algebra.

The generator set now is $\{E_\alpha, E_\beta, E_0, E_1, \ldots, E_{12}\}$. The defining relation set is taken to be

$$E_0^2 = pE_\alpha + qE_\beta, \quad E_\alpha^2 = \sum_{i=1}^{12} \frac{1}{12} E_i, \quad E_\beta^2 = \sum_{i=1}^{12} \frac{1}{12} E_i, \quad E_1^2 = \frac{1}{2}E_1 + \frac{1}{2}E_2,$$

for $2 \leq j \leq 11$,

$$E_j^2 = \frac{1}{3}E_{j-1} + \frac{1}{3}E_j + \frac{1}{3}E_{j+1},$$

and for $j = 12$

$$E_{12}^2 = \frac{1}{2}E_{11} + \frac{1}{2}E_{12};$$

and if $i \neq j$,

$$E_i \cdot E_j = 0.$$

Although this new algebra is not simple, it has a simple subalgebra that forms progeny population. Two unstable races will eventually disappear through producing other races. Whatever the values of $p$ and $q$ are, we eventually get the same frequency of each race in the population as that in the simple algebra above, except that $E_\alpha$ and $E_\beta$ have 0 frequency.

There is a trivial simple algebra generated by race PI-1P. If we denote PI-1P by $E_{-1}$, the progeny population is generated by $E_{-1}$ which is subject to $E_{-1}^2 = E_{-1}$.

In paper [12], there are 5 different parental races in Minnesota and North Dakota from 1994 to 1996. If we want to study the whole structure of P. infestans population there, we need to construct a big algebra which is reproduced by 5 parental races, PI-105P, PI-191P, PI-52P, PI-126P and PI-1P. This algebra will have 5 simple subalgebras which corresponds to the progeny subpopulations produced by 5 parental races. We also need to compute the frequency of each progeny subpopulation. This way, we encode the complexity of structure of progenies of P. infestans into an algebra.

Let's summarize what the evolution algebras can provide to plant pathologists theoretically.

- **(1):** Evolution algebra theory can predict the existence of intermediate transient races. Intermediate transient races correspond to algebraic transient elements. They are biologically unstable, and will extinct or disappear by producing other races after a certain period of time. If we can catch the intermediate transient races and remove or kill them, we will easily stop the spread of late blight disease.
- **(2):** Evolution algebra theory states that biologically stable races correspond to algebraic persistent elements. It predicts the periodicity of reproduction of stable races. This is helpful to understand the speed of spread of plant diseases.
- **(3):** Evolution algebra can re-recover progeny subpopulation. Furthermore, because these progeny subpopulations correspond to simple subalgebras, each race in the same subpopulation shares the same dynamics of reproduction and

spreading. Evolution algebras are therefore helpful to simplify the complexity of progeny population structure.

**(4):** Evolution algebra theory provides a way to compute the frequency of each race in progeny population given the reproduction rates, which are algebra structural constants. Practically, these frequencies can be measured, and so reproduction rates can be computed by formulae in evolution algebras. Therefore, evolution algebras will be a helpful tool to study many aspects of asexual reproduction process, like that of Oomycetes, Phytophthora.

**Acknowledgments.** I would like to thank Avner Friedman and Jean-Michel Morel for their suggestions. I would also like to thank C. William Birky Jr. for his explanation of uniparental inheritance by emails. The part of this work was done when the author was a postdoctoral fellow at MBI supported by the National Science Foundation by agreement No.0112050. The author also thank Reves Center at the College of William and Mary for supporting a conference travel to Shanghai, China.

## REFERENCES

[1] C. W. Jr. Birky, *The inheritance of genes in mitochondria and chloroplasts: Laws, mechanisms, and models*, Annu. Rev. Genet., **35** (2001), 125–148.
[2] C. W. Jr. Birky, "Inheritance of Mitochondrial Mutations," Mitochondrial DNA Mutations and Aging, Disease and Cancer, K.K. Singh, edit, Spring, 1998.
[3] Jianjun Paul Tian, "Evolution Algebras and their Applications," Lecture Note in Mathematics, **1921**, Springer, Berlin, 2008.
[4] G. Mendel, "Experiments in Plant-Hybridization," Classic Papers in Genetics, 1-20, J. A. Peter editor, Prentice-Hall Inc. 1959.
[5] Y. I. Lyubich, "Mathematical Structures in Population Genetics," Springer-Verlag, New York, 1992.
[6] A. Worz-Busekros, "Algebras in Genetics," Lecture Notes in Biomath. 36, Springer-Verlag, Berlin-New York, 1980.
[7] M. L. Reed, *Algebraic structure of genetic inheritance*, Bull. of AMS, **34** (1997), 107–130.
[8] N. W. Gillham, "Organelle Genes and Genomes," Oxford University Press, 1994.
[9] C. F. Emmerson, G. K. Brown and J. Poulton, *Synthesis of mitochondrial DNA in permeabilised human cultured cells*, Nucleic Acids Res., **29** (2001), e1–e1(1).
[10] F. Ling and T. Shibata, *Mhr1p-dependent concatemeric mitochondrial DNA formation for generating yeast mitochondrial homoplasmic cells*, Mol. Biol. Cell, **15** (2004), 310–322.
[11] Y. Tang, G. Manfredi, M. Hirano and E. A. Schon, *Maintenance of human rearranged mitochondrial DNAs in long-term transmitochondrial cell lines*, Mol. Biol. Cell, **11** (2000), 2349–2358.
[12] F. M. A. Samen, G. A. Secor and N. C. Gudmestad, *Variability in virulence among asexual progenies of Phytophthora infestans*, Phytopathology, **93** (2003), 293–304.
[13] W. E. Fry, and S. B. Goodwin, *Re-emergence of potato and tomato late blight in the United States*, Plant Disease, **81** (1997), 1349–1357.

Received April 2009; revised October 2009.

*E-mail address*: jptian@math.wm.edu