

# Understanding the dynamics underpinning the geographical distribution of the *Anastrepha fraterculus* complex in eastern Brazil

L. Sánchez<sup>1</sup>, A. L. P. Perondini<sup>2</sup> and D. Selivon<sup>2,\*</sup>

<sup>1</sup>Centro de Investigaciones Biológicas (C. S. I. C.), Ramiro de Maeztu 9, 28040 Madrid, Spain.

<sup>2</sup>Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo (USP), Rua do Matão, 277, 055080-090 São Paulo, SP, Brazil.

## ABSTRACT

Three species of the *Anastrepha fraterculus* complex (AF complex) occur in eastern Brazil, namely *Anastrepha* sp.1 aff. *fraterculus*, *Anastrepha* sp.2 aff. *fraterculus*, and *Anastrepha* sp.3 aff. *fraterculus*. Along their extensive distribution, which differs in relation to geographic topography, they present differences in the host fruit usage in some zones of sympatry. The dynamics underlying the extant distribution pattern of the three species of the AF complex in eastern Brazil was herein analysed by a novel application of the generalized logical formalism methodology. The present report illustrates that two factors, the uni-directional incompatibility between *sp1* and *sp2*, with the former overriding the second, besides a yet unknown environmental factor related to altitude, play an instructional role in the generation of the extant distribution pattern observed in nature. Simulations about temporal order of invasion of species in a given area showed that the outcome is independent of the order of arrival of the species. This kind of approach has also a potential to make predictions about the population dynamics after the introduction of species, or the eradication of species in a given area.

**KEYWORDS:** fruit flies, host fruit usage, cryptic species, generalized logical formalism, environment.

## INTRODUCTION

The fruticulture around the world suffer the attack of several pests that cause huge losses in the production of commercial varieties of fruits. Among these pests, fruit flies of the Tephritidae family are one of the most important, since they use fruits as substrate for larval development [1]. Species of *Anastrepha* causes losses estimated between 120 and 220 million dollars per year to the Brazilian fruticulture [2, 3]. Among these species, the nominal species *Anastrepha fraterculus* (Wiedemann) is widely distributed with virtually the same distribution of the genus, and is one of the most important pests from the economic point of view. This nominal species, however, comprises a complex of cryptic species, the *Anastrepha fraterculus* complex (AF complex). In Brazil, three species of the AF complex have been characterized through a pioneering analysis that integrates biological characteristics such as, genetic, karyotype, molecular, host usage, egg morphology, morphometry of adult structures and pre-zygotic and post-zygotic isolation mechanisms. Based on these data, three entities were characterized and provisionally named as distinct species, *A. sp.1* aff. *fraterculus*, *A. sp.2* aff. *fraterculus* and *A. sp.3* aff. *fraterculus* [4-9], and herein will be shortly named as *A. sp.1*, *A. sp.2*

---

\*Corresponding author: dselivon@ib.usp.br

and *A. sp.3*. Later on, morphometric analysis applied to more samples from extensive geographic areas of the Neotropical Region allowed the characterization of five more morphotypes in the AF complex [10, 11].

The three species are dispersed abundantly from north to south in eastern Brazil where they could be correctly identified [4, 5, 8]. The relief of this Brazilian region is characterized by an inland plateau with moderate to higher elevations. The plateau begins near the coast and advances toward the internal areas of the country. Such relief configuration defines a stripe of land at sea level, the coastal plains, running from south to northeast regions [12]. Extensive sampling in the last 20 years evidenced three sympatric zones of occurrence of the three species. One in the coastal plains from southeaster to south areas, and two in the inland plateau, the valley of the Paraíba do Sul River (States of São Paulo and Rio de Janeiro), and in the south (States of Santa Catarina and Rio Grande do Sul). Along the coastal plains, *A. sp.2* and *A. sp.3* are found sympatrically. In the valley of Paraíba do Sul River of the inland plateau, the three species coexist, *A. sp.2* and *A. sp.3* being less abundant than *A. sp.1* [7, 8, 13] while in the southern inland plateau, *A. sp.3* specimens are found occasionally with *A. sp.1* [7, 14].

The analysis of the relationships between the pest insects, their hosts and their distribution in a given area, has been mainly focused on determining which hosts a certain species of insect uses for its larval development. For this purpose, two descriptive methodologies have been exploited. One methodology is based on the idea of a bipartite graph, whose components are divided into two disjoint sets, one representing the insects and the second their host fruits. The arcs connecting the two sets represent the insect-fruit relationship, that is, an arc is drawn from a given insect to its host fruits. The second methodology represents the insect-fruit relationship by a matrix where cell  $if = 1$  means that insect  $i$  uses fruit  $f$  for its larval development and  $if = 0$  means that the insect does not use it. However, these two methodologies are merely descriptive and add nothing about the dynamics that created such a distribution pattern. With the purpose of obtaining information on the dynamics underlying the extant distribution

pattern of *A. sp.1*, *A. sp.2* and *A. sp.3* in eastern Brazilian regions, we used the generalized logical formalism [15, 16]. This analysis would throw light on the putative forces -insect and fruit interactions as well as environmental factors-involved in shaping the insect pattern distribution.

## MATERIALS AND METHODS

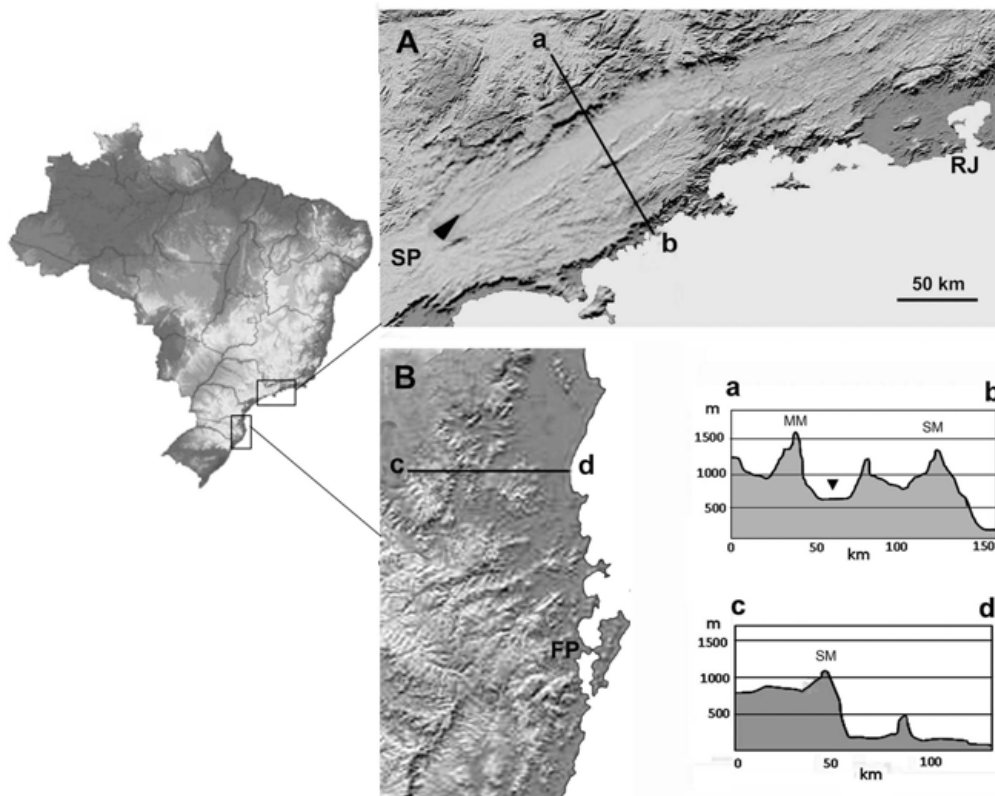
### Species distribution of the AF complex

In southern areas of eastern Brazil the structural (topographic) relief comprises the coastal plains at sea level and the inland plateau with higher altitudes (Fig. 1) [17]. Beginning in the most southern areas of the inland plateau and running parallel to the coast up to Rio de Janeiro, exists a scarp, the Serra do Mar (Sea Mountains). Slightly north of São Paulo, at west and running in a NNE direction parallel to the Serra do Mar, exists another scarp in the inland plateau, Serra da Mantiqueira. Between these two mountain ranges exist a valley, the valley of the Paraíba do Sul River that extend up to the Atlantic Ocean north of Rio de Janeiro (Fig. 1A). In the inland plateau, including the valley, altitude varies from 500-800 m, while in the mountain ridges it may reach 1000 to >1500 m [18]. Hence, one may consider that three basic altitudinal landscapes are found in the southeastern region: lands at sea level, at an intermediate and at higher altitudes in the inland plateau. Similarly, the three basic landscapes are also present in south Brazil (Fig. 1B).

For the present analyses, data on the collection of the three species of the AF complex, *A. sp.1*, *A. sp.2* and *A. sp.3*, in the coastal areas, and in the inland plateau of southeastern and southern areas of Brazil were used. Although the three species, *A. sp.1*, *A. sp.2* and *A. sp.3*, may use various host fruits for larval development, in the present study the analyses will be restricted to the three main host ones, a native fruit *Psidium guajava* (guavas), and exotic species *Citrus sp.* (oranges) and tropical almond (*Terminalia cattapa*) [5, 7, 8, 13]. These fruits used by the species were found in the three sympatric zones as shown in Table 1.

### The model of dynamic analysis

The model was defined using the generalized logical formalism [15, 16] and the software tool



**Fig. 1.** Details of the relief in small areas in southeastern (A) and southern (B) Brazilian regions. The coastal land is shown in dark gray and the inland plateau in different shades of light gray (insets based on [17]). The graphs on the right show the approximate topographic profiles of transects a-b and c-d in the maps: MM, Mantiqueira Mountains; SM, Sea Mountains (based on [18]). The arrowheads in A and graphic a-b, indicate the Paraíba do Sul valley, and RJ, SP and FP indicate the location of the cities of Rio de Janeiro, São Paulo and Florianópolis.

**Table 1.** Host fruit\* usage (A) in the three zones of sympatry (B, C) by *Anastrepha fraterculus* species in eastern Brazil.

Species	sp1	sp2	sp3
<b>A. Host fruit usage</b>	GUA	GUA, ORA, TAL	GUA, TAL
<b>B. Presence of species in sympatry</b>			
Coastal region	Not present	GUA, ORA, TAL	GUA, TAL
<b>C. Presence of species in the inland zones of sympatry</b>			
Paraíba do Sul River	GUA	ORA	GUA, TAL
South Inland Plateau	GUA	Not present	GUA, TAL

\* GUA, guava; ORA, orange; TAL, tropical almond.

GINsim [19]. Further details of the equations and logical parameters for the networks are given in Table 2. Briefly, the relationship between the insects and their host fruits is represented as a

network. The nodes stand for the species, the fruits and the colonization of the fruits by the species. Each node is assigned as a discrete variable that describes the node state, with a maximal level

**Table 2.** Equations and logical parameters for the Networks.

Network element	Functional level	Equation
2A. Equations for Figure 1.		
SP2-GUA	1	SP2 & GUA
SP2-ORA	1	SP2 & ORA
SP2-TAL	1	SP2 & TAL
SP3-GUA	1	SP3 & GUA
SP3-TAL	1	SP3 & TAL
2B. Logical parameters for the Network shown in Figure 1.		
SP2	1	Constant to initial value, input = {0, 1}
SP3	1	Constant to initial value, input = {0, 1}
GUA	1	Constant to initial value, input = {0, 1}
ORA	1	Constant to initial value, input = {0, 1}
TAL	1	Constant to initial value, input = {0, 1}
SP2-GUA	1	$K_{SP2-GUA}(\{SP2,1\}, \{GUA,1\}) = 1$
SP2-ORA	1	$K_{SP2-ORA}(\{SP2,1\}, \{ORA,1\}) = 1$
SP2-TAL	1	$K_{SP2-TAL}(\{SP2,1\}, \{TAL,1\}) = 1$
SP3-GUA	1	$K_{SP3-GUA}(\{SP3,1\}, \{GUA,1\}) = 1$
SP3-TAL	1	$K_{SP3-TAL}(\{SP3,1\}, \{TAL,1\}) = 1$
Interactions are denoted by the name of their source component, together with the level(s) for which they are operative. Logical parameters define the effect of a given combination of interactions (in parenthesis) operating on a given component. For example, $K_{SP2-GUA}(\{SP2,1\}, \{GUA,1\}) = 1$ indicates that the value of SP2-GUA is 1 in the presence of SP2 and GUA, both at their functional level 1. Only non-zero logical parameters are listed.		
2C. Equations for the Network shown in Figure 2.		
SP1-GUA	1	SP1 & GUA & !SP2-GUA
SP2-GUA	1	SP2 & GUA & !SP1-GUA
SP2-ORA	1	SP2 & ORA
SP3-GUA	1	SP3 & GUA
2D. Logical parameters for the Network shown in Figure 2.		
SP2	1	Constant to initial value, input = {0, 1}
SP3	1	Constant to initial value, input = {0, 1}
GUA	1	Constant to initial value, input = {0, 1}
ORA	1	Constant to initial value, input = {0, 1}
SP1-GUA	1	$K_{SP1-GUA}(\{SP1,1\}, \{GUA,1\}) = 1$
SP2-GUA	1	$K_{SP2-GUA}(\{SP2,1\}, \{GUA,1\}) = 1$
SP2-ORA	1	$K_{SP2-ORA}(\{SP2,1\}, \{ORA,1\}) = 1$

Table 2 continued ...

SP3-GUA	1	$K_{SP3-GUA}(\{SP3,1\}, \{GUA,1\}) = 1$
Interactions are denoted by the name of their source component, together with the level(s) for which they are operative. Logical parameters define the effect of a given combination of interactions (in parenthesis) operating on a given component. For example, $K_{SP1-GUA}(\{SP1,1\}, \{GUA,1\}) = 1$ indicates that the value of SP1-GUA is 1 in the presence of SP2 and GUA , both at their functional level 1. Only non-zero logical parameters are listed.		
2E. Equations for the Network in Figure 3.		
SP1-GUA	1	SP1 & GUA & EF:2 & !SP2-GUA
SP2-GUA	1	SP2 & GUA & (!EF:1   !EF:2) & EF:3 & !SP1-GUA
	2	SP2 & GUA & (EF:1   EF:2) & !EF:3 & !SP1-GUA
SP3-GUA	1	SP3 & GUA & !EF:1 & (EF:2   EF:3)
	2	SP3 & GUA & EF:1 & !EF:2 & !EF:3
2F. Logical parameters for the Network shown in Figure 3.		
EF		Constant to initial value, input={0, 1, 2, 3}
SP1	1	Constant to initial value, input={0, 1}
SP2	1	Constant to initial value, input={0, 1}
SP3	1	Constant to initial value, input={0, 1}
GUA	1	Constant to initial value, input={0, 1}
SP1-GUA	1	$K_{SP1-GUA}(\{EF,2\}, \{SP1,1\}, \{GUA,1\}) = 1$
SP2-GUA	2	$K_{SP2-GUA}(\{EF,3\}, \{SP2,1\}, \{GUA,1\}) = 1$
		$K_{SP2-GUA}(\{EF,1\}, \{SP2,1\}, \{GUA,1\}) = 2$
		$K_{SP2-GUA}(\{EF,2\}, \{SP2,1\}, \{GUA,1\}) = 2$
SP3-GUA	2	$K_{SP3-GUA}(\{EF,2\}, \{SP3,1\}, \{GUA,1\}) = 1$
		$K_{SP3-GUA}(\{EF,3\}, \{SP3,1\}, \{GUA,1\}) = 1$
		$K_{SP3-GUA}(\{EF,1\}, \{SP3,1\}, \{GUA,1\}) = 2$

Interactions are denoted by the name of their source component, together with the level(s) for which they are operative. Logical parameters define the effect of a given combination of interactions (in parenthesis) operating on a given component. For example,  $K_{SP1-GUA}(\{SP1,1\}, \{GUA,1\}) = 1$  indicates that the value of SP1-GUA is 1 in the presence of SP2 and GUA , both at their functional level 1. Only non-zero logical parameters are listed.

defining the highest qualitative functional level of the regulatory node (this maximal level equals 1 in the simplest Boolean case). The arcs represent either the colonization of fruits by species or the interactions between the species for colonizing the fruits. Each arc represents an interaction and is assigned a threshold, which defines the smallest level of the interaction source for which the interaction is operative. Logical parameters qualitatively describe the effects of the interactions controlling the states of the network nodes.

A state of the network is represented by a vector, which includes the current (discrete) node levels. Given a state, one can determine which interactions

are operative and the values of the logical parameters indicate the nodes called to change their levels (i.e. those nodes whose current levels differ from the values of the logical parameters). In general, for a given state, all possible elementary transitions (i.e. switching of a single node level, to a neighbouring integer level) are considered, thus leading to as many outgoing transitions as updating calls (asynchronous updating). Depending on the structure of the regulatory graph and on the values of the logical parameters, a model leads to a large but finite number of dynamical pathways. In turn, these are represented in the form of a state transition graph, where nodes represent states and directed arcs represent transitions between states.

Model dynamics are defined by the specification of initial state(s) and of an updating scheme, which refers to how network component levels are updated. The asynchronous update is generally chosen as it amounts to generate all possible trajectories from the specified initial conditions: e.g. when two component levels are changed, these updates are done asynchronously, defining two transitions towards alternative successor states. The resulting state transition graph may then include bifurcation states from which, depending on the choice of the successor state, the final stable state may differ; in other words, in the absence of any further constraint, two alternative outcomes are reachable.

In the particular case that is been analysed, *sp1*, *sp2* and *sp3* as well as the fruits, guavas (GUA), oranges (ORA) and tropical almonds (TAL) are inputs of the network. The use of a fruit by an insect for its larval development is represented by a variable whose value depends on the insect species and the fruits, and on putative positive or negative interactions between the insects for colonizing the same fruit. Thus, for example, the node (variable) SP1-GUA means that *sp1* uses guavas for its larval development. From a theoretical point of view, there are nine nodes in the present study: SP1-GUA, SP1-ORA, SP1-TAL, SP2-GUA, SP2-ORA, SP2-TAL, SP3-GUA, SP3-ORA and SP3-TAL. However, note that not all these cases were found in nature since *sp1*, *sp2* and *sp3* do not colonize every species of fruits. The stable states that resulted from the dynamic analysis represent the insect-fruit association, whose biological meaning is that a given insect is adapted in exploring a given resource for the larval development. Equations and logical parameters for the networks are given in Table 2.

The analyses that follow were developed in three steps involving distinct elements of the relationship between the insects and their host fruits in the three zones of sympatry. In the first, the species distribution and their association with the host fruits were evaluated. In a second step, and based on the topography of the sympatric zones, the influence of a putative environmental factor in the distribution of the three species was added. In a third step, simulations about the order of colonization by the three species of a given area will be presented.

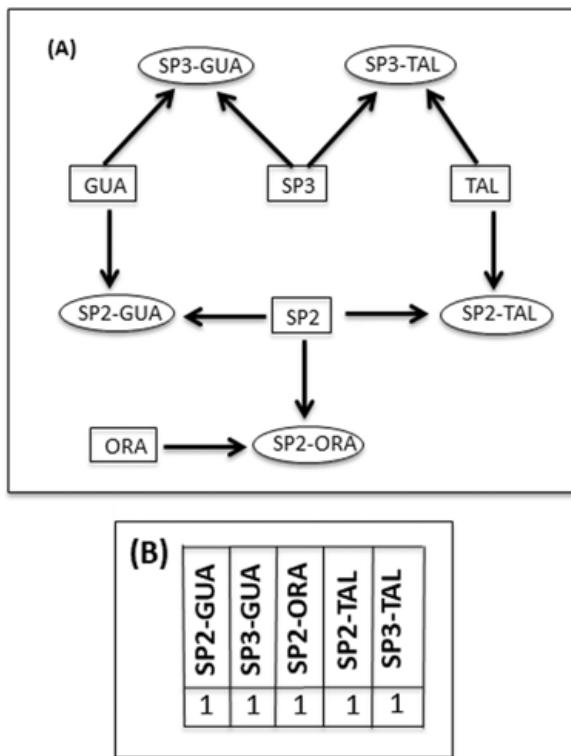
## RESULTS

### Host fruits and species distributions

The main host fruit usage by the three species of the AF complex in southern Brazil is given in Table 1A. Tables 1B and 1C give details of the species distribution in the three sympatric regions, the coastal plains and the two areas in the inland plateau, the valley of the Paraíba do Sul River and the southern most region. Note that *sp2* and *sp3* can simultaneously explore in the same fruit, and that *sp1* and *sp2* do not occur simultaneously in guavas [8], suggesting the existence of an incompatibility between these two species, regardless which is(are) the factor(s).

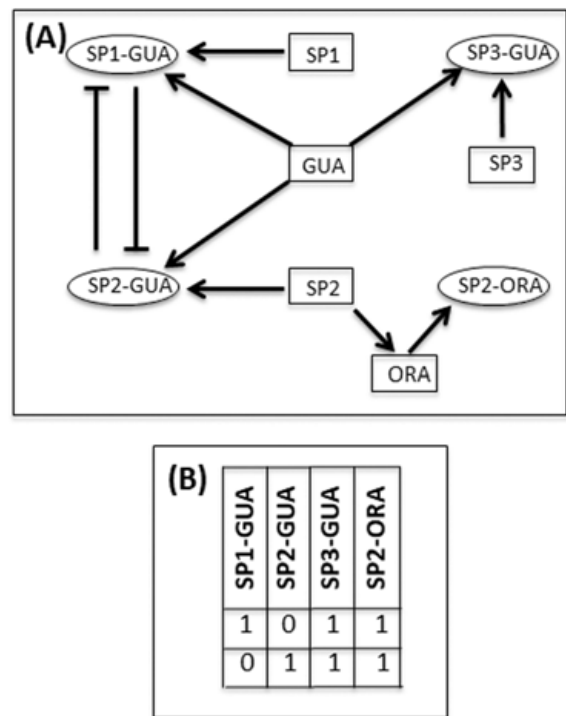
The network corresponding to the species distribution in the sympatric coastal region is shown in Fig. 2A. First, *sp1* does not occur at sea level, and only *sp2* and *sp3* were found in this region. Since *sp2* and *sp3* can be found together in guavas and in tropical almond, it was considered that there is no incompatibility between these two species. The dynamics of this network gives rise to a single stable state (Fig. 2B), in agreement with experimental observations, corresponding to the formation of the complexes, SP2-GUA, SP3-GUA, SP2-ORA, SP2-TAL and SP3-TAL.

The distribution of *sp1*, *sp2* and *sp3* and the fruits they use in the two sympatric regions, valley of Paraíba do Sul River and southern areas of the inland plateau, is very similar and is given in Table 1C. Two remarkable facts characterize the host fruit usage. Firstly, *sp1* or *sp3* has never been found colonizing oranges. Secondly, although *sp2* can colonize guavas in the coastal areas, it has not yet been found colonizing guavas with *sp1* in the inland plateau, either in distinct fruits or in the same fruit at the same time. This fact suggests the existence of a bi-directional incompatibility -whatever reason- between *sp1* and *sp2*, as indicated by the negative blunt arrows connecting SP1-GUA and SP2-GUA in Fig. 3A. Such incompatibility does not exist between *sp2* and *sp3* as both can simultaneously colonize guavas and tropical almond in coastal areas. The network corresponding to the species distribution in the sympatric regions of valley of Paraíba do Sul River and southern areas of the inland plateau is shown in Fig. 3A, while the detailed dynamics of this network is shown in Fig. 4. The analysis gives rise to two



**Fig. 2.** Colonization of the coastal region by *sp2* and *sp3*. (A) Network determining the distribution pattern of these species. Rectangles and ellipsoids stand for inputs and variables, respectively. Normal arrows represent positive interactions. (B) Final stable state attained after the invasion of the region by the species. The number 1 indicates that the complexes SP2-GUA, SP3-GUA, SP2-ORA, SP2-TAL and SP3-TAL are formed.

theoretical stable states (Fig. 3B). Both of them are characterized by the formation of complexes SP3-GUA and SP2-ORA, differing in the third complex, which is either SP1-GUA or SP2-GUA. This difference would be a consequence of the assumed bi-directional incompatibility between *sp1* and *sp2*. One of the state-transitions correspond to the formation of states SP1-GUA=1 and SP2-GUA=0 (Fig. 3B). The other state-transitions correspond to the formation of states SP1-GUA=0 and SP2-GUA=1 (Fig. 3B). It was discovered, however, that in nature *sp1* but not *sp2* was always found in guavas, and since *sp2* can use guavas as a substrate for larval development, as it happens in the coastal region (see above), it is concluded that, in reality, there must exist a uni-directional incompatibility of *sp1* upon *sp2*,

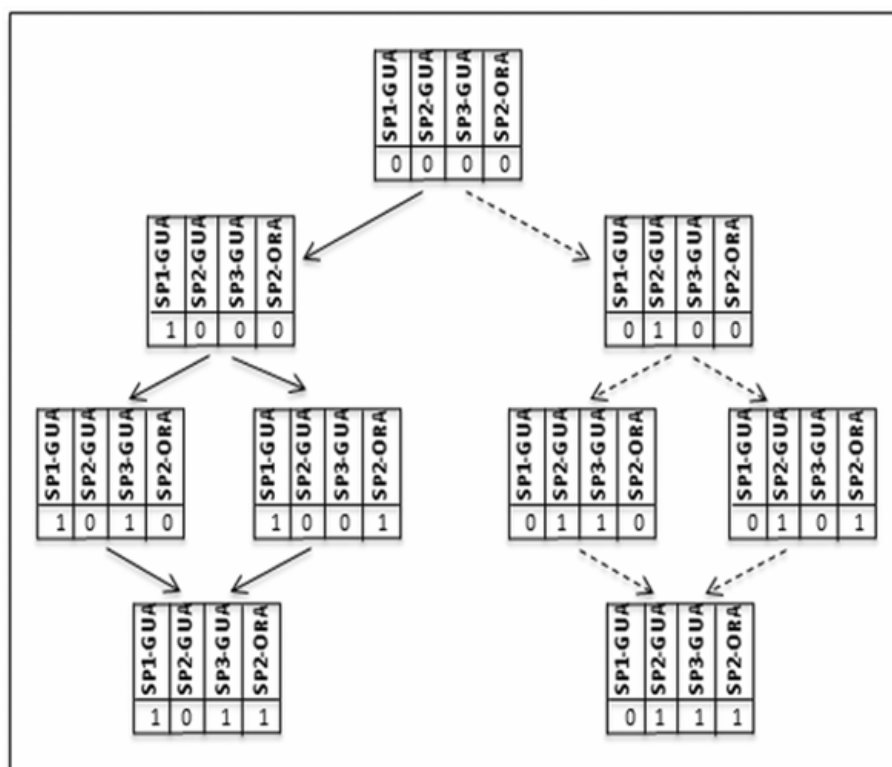


**Fig. 3.** Colonization of the Paraíba do Sul valley and southern regions by *sp1*, *sp2* and *sp3*. (A) Network determining the distribution pattern of these species. Rectangles and ellipsoids stand for inputs and variables, respectively. Normal and blunt arrows represent positive and negative interactions, respectively. (B) Final stable states attained after the invasion of the region by the species. The number 1 indicates the formation of the complexes insect-fruit, whereas the number 0 indicates that they are not formed.

with *sp1* overriding *sp2*. Moreover, the observed differences in the host fruit usage by *sp1*, *sp2* and *sp3* in the coastal and in the inland sympatric regions, mentioned above, are suggestive that putative influences of the environment are also acting on their distribution in these areas.

**Environment and distribution of *A. sp.1*, *A. sp.2* and *A. sp.3* in guavas**

Since the three species are capable of developing in guavas, a native host fruit, data of only this host was used in the analysis concerning the effect of the environment on the geographical distribution of AF complex. To study such distribution in the present analysis, the existence of an environmental factor (EF) related to altitude, whatever its nature, influencing the distribution pattern was assumed.



**Fig. 4.** Detailed dynamics of the colonization of the Paraíba do Sul valley and Southern regions by *sp1*, *sp2* and *sp3*. Network determining the distribution pattern of these species. Rectangles and ellipsoids stand for inputs and variables, respectively. Normal and blunt arrows represent positive and negative interactions, respectively. Final stable states attained after the invasion of the region by the species. The formation of the complex insect-fruit is indicated by the number 1, whereas the number 0 indicates that they are not formed.

For the sake of the present analyses, and based on Gontijo-Pascutti *et al.* [18] three altitudinal categories were recognized: EF = 1 at sea level, and two levels in the inland plateau, EF = 2 (intermediate level) and EF = 3 (high level), as shown in Fig. 1. In relation to altitude, *sp1*, *sp2* and *sp3*, showed distinct distributions, as shown in Table 3: *sp1* is not present at sea level but abundant at intermediate and high levels; *sp2* and *sp3* are abundant at sea level and although present, they are less abundant at intermediate and high levels. Distribution of *sp2* deserves special mention. The putative effect of the environment on the colonization pattern of guavas by *sp2* may be hidden by the inhibitory effect that *sp1* exerts upon *sp2* for colonizing the same fruit (see above). To overcome this inconvenience, we used the orange-colonization distribution of *sp2* to analyse the effect of the environment on its distribution pattern since *sp1* does not interfere because this

species is incapable of exploring orange. Table 3 presents the pattern distribution of *sp2* on colonizing orange depending on the environment.

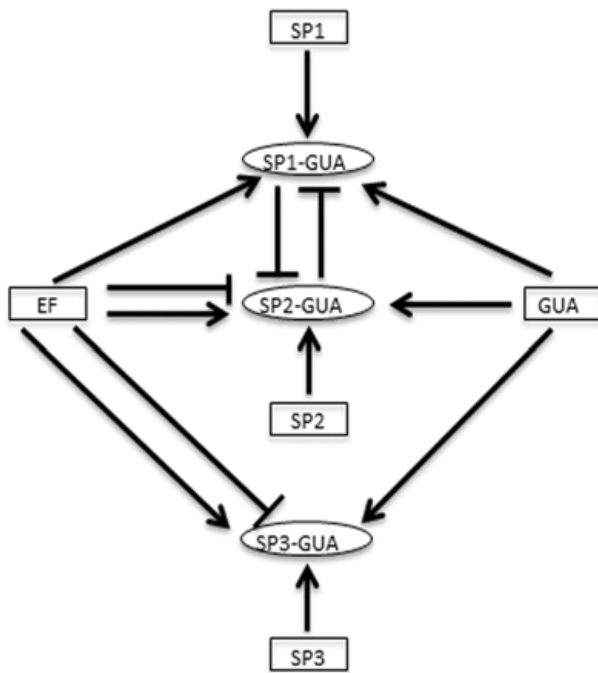
Fig. 5 models this ecological scenario for the environment-dependent distribution of *sp1*, *sp2* and *sp3*. The EF factor is an input of the network, and as before, SP1, SP2 and SP3 plus GUA (guava) constituted inputs, whereas SP1-GUA, SP2-GUA and SP3-GUA represent the variables corresponding to the species populating guavas. Since *sp1* is either absent or present, SP1-GUA is represented by a Boolean variable, that is SP1-GUA = {0, 1}, where 0 means absence and 1 means presence. Since *sp2* colonizes orange in the three ecological niches (Table 3), the absence of *sp2* in guavas at intermediate and high levels cannot be due to an effect of the environmental factor. Rather, the absence of *sp2* in guavas would be related to its putative incompatibility with *sp1*, as explained above. Indeed, the geographical



**Table 3.** Geographical distribution of *sp1*, *sp2* and *sp3* colonizing guava depending on the environment.

Environmental factor (EF)	SP1/Guava	SP2		SP3/Guava
		SP2/Guava	SP2/Orange*	
Sea level (EF = 1)	Not present	Abundant	Abundant	Abundant
Intermediate level (EF = 2)	Abundant	Not present	Abundant	Less abundant
High level (EF = 3)	Abundant	Not present	Less abundant	Less abundant

\*See text for details of sp2/orange column.



**Fig. 5.** Network corresponding to the effect of the environment (EF) on the distribution pattern of *sp1*, *sp2* and *sp3* species. Rectangles and ellipsoids stand for inputs and variables, respectively. Normal and blunt arrows represent positive and negative interactions, respectively. Note the dual interaction (negative and positive) of EF on SP2-GUA and SP3-GUA. The positive interaction of EF upon SP2-GUA is operative at functional threshold levels  $EF = \{1, 2\}$ , whereas the negative interaction is operative at  $EF = \{3\}$ . The positive interaction of EF upon SP3-GUA is operative at functional threshold level  $EF = \{1\}$ , whereas the negative interaction is operative at  $EF = \{2, 3\}$ .

distribution of *sp1* and *sp2* is complementary. There is, however, a certain effect of the environmental factor on *sp2*, because *sp2* is abundant in oranges at sea and intermediate level, but less abundant at high level. It is considered here that the effect of the environment on the distribution pattern of *sp2* is independent of the fruit, and that the distribution pattern of *sp2* in oranges and guavas would be the same. Hence, SP2-GUA is represented by a multivalued logical variable:  $SP2-GUA = \{0, 1, 2\}$ , where 0 means absence, 1 means low abundance and 2 means abundant. Note in Table 4 that when SP1-GUA has the value 1, SP2-GUA has the value 0; whereas when SP1-GUA has the value 0, SP2-GUA can have a value 1 or 2 depending on the environmental signal -this latter distribution is shown in Table 3. With respect to *sp3*, it is abundant at sea level and low abundant at intermediate and high levels. Hence, SP3-GUA is also represented by a multivalued logical variable:  $SP3-GUA = \{0, 1, 2\}$ .

The dynamical analysis was done for each ecological niche defined by the value of the environmental factor. Moreover, three different scenarios regarding the temporal sequence of the arrival of each species in an environmental area were simulated: the three species invade simultaneously, there is invasion of *sp2* and *sp3* but not *sp1*, and there is invasion of *sp1* and *sp3* but not *sp2*, as is shown below.

**Simulations about the order of invasion in sympatric zones**

(1) The first analysis considers that *sp1*, *sp2* and *sp3* simultaneously invade an area. This scenario

**Table 4.** Presence\* of *sp1*, *sp2* and *sp3* after the dynamical simulation of the species invading and colonizing guavas depending on the environment.

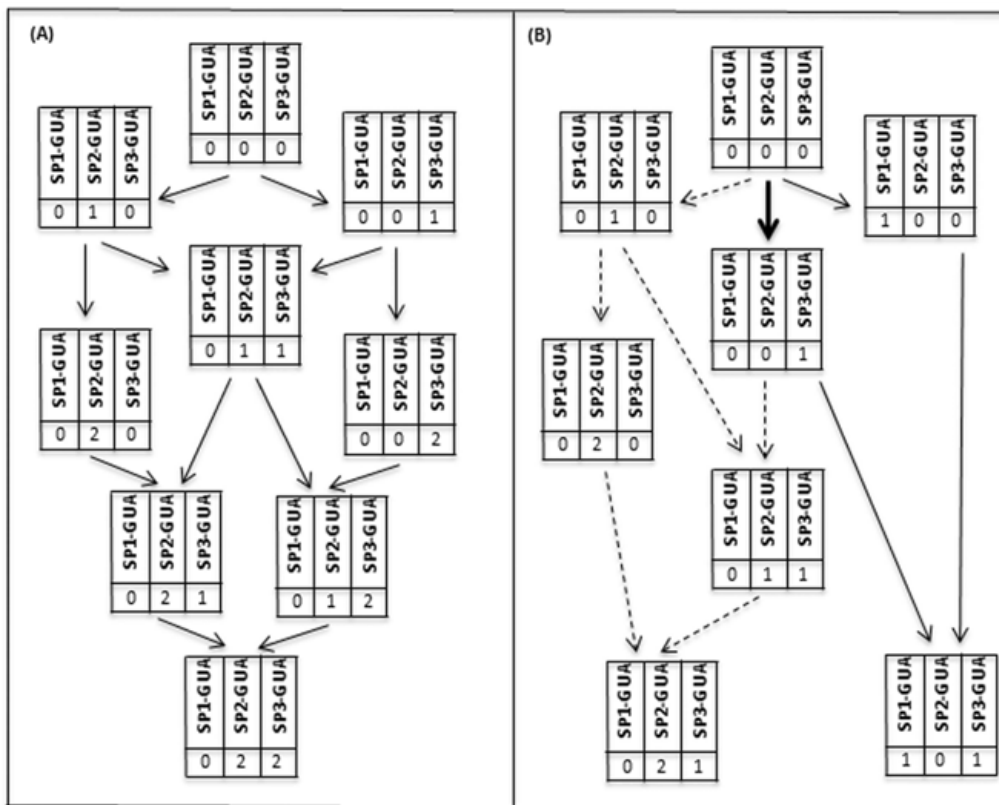
Invasion	EF Level	Colonization						
		Initial status				Stable states		
		GUA	SP1	SP2	SP3	SP1-GUA	SP2-GUA	SP3-GUA
<i>sp1, sp2, sp3</i>	Sea level	1	1	1	1	0	2	2
	Intermediate level	1	1	1	1	1	0	1
		1	1	1	1	0	2	1
	High level	1	1	1	1	1	0	1
1		1	1	1	0	1	1	
<i>sp2, sp3, not sp1</i>	Sea level	1	0	1	1	0	2	2
	Intermediate level	1	0	1	1	0	2	1
	High level	1	0	1	1	0	1	1
<i>sp1, sp3 not sp2</i>	Sea level	1	1	0	1	0	0	2
	Intermediate level	1	1	0	1	1	0	1
	High level	1	1	0	1	1	0	1

\*(0) Absence. (1) Abundance in the case of Boolean variable SP1-GUA and low presence in the case of multivalued variables SP2-GUA and SP3-GUA. (2) Abundance in the case of multivalued variables SP2-GUA and SP3-GUA.

corresponds to the niche with the environmental factor  $EF = 1$ . The detailed colonization dynamics of the sea level niche is shown in Fig. 6A. A single stable state was attained (Table 4), which agrees with the observed experimental distribution of species of the AF complex at the sea level, that is, absence of *sp1* ( $SP1-GUA = 0$ ) and abundance of *sp2* ( $SP2-GUA = 2$ ) and *sp3* ( $SP3-GUA = 2$ ). Notice that *sp1* does not occur in that geographical region so that species *sp2* can infest guavas.

The next scenario corresponds to the intermediate level niche ( $EF = 2$ ), whose detailed colonization dynamics is shown in Fig. 6B. Two stable states are reached (Table 4). The first state corresponds to the *sp1*-populated guavas ( $SP1-GUA = 1$ ) and absence of *sp2* in guavas ( $SP2-GUA = 0$ ). The second state corresponds to the absence of *sp1* ( $SP1-GUA = 0$ ) and abundance of *sp2* ( $SP2-GUA = 2$ ) in guavas.

The transition state that is common to both is indicated in Fig. 6B. Of note, the common state is characterized by both variables  $SP1-GUA$  and  $SP2-GUA$  having the value zero. If *sp1* is present, the next transition state is [ $SP1-GUA = 1$  &  $SP2-GUA = 0$ ] that ends into the final stable state characterized by the presence of *sp1* and absence of *sp2*. On the contrary, if *sp1* is absent, the next transition state is [ $SP1-GUA = 0$  &  $SP2-GUA = 1$ ] that ends into the final state characterized by the absence of *sp1* and the abundant presence of *sp2*. The stable state found in nature was the one corresponding to the presence of *sp1* and absence of *sp2*. The next scenario corresponds to the colonization of the high level niche ( $EF = 3$ ). Two stable states are reached (Table 4). The first state corresponds to the *sp1*-populated guavas ( $SP1-GUA = 1$ ) and the absence of *sp2* ( $SP2-GUA = 0$ ). The second state corresponds to the absence of *sp1* ( $SP1-GUA = 0$ ) and less abundance of *sp2* ( $SP2-GUA = 1$ ). This latter stable state was

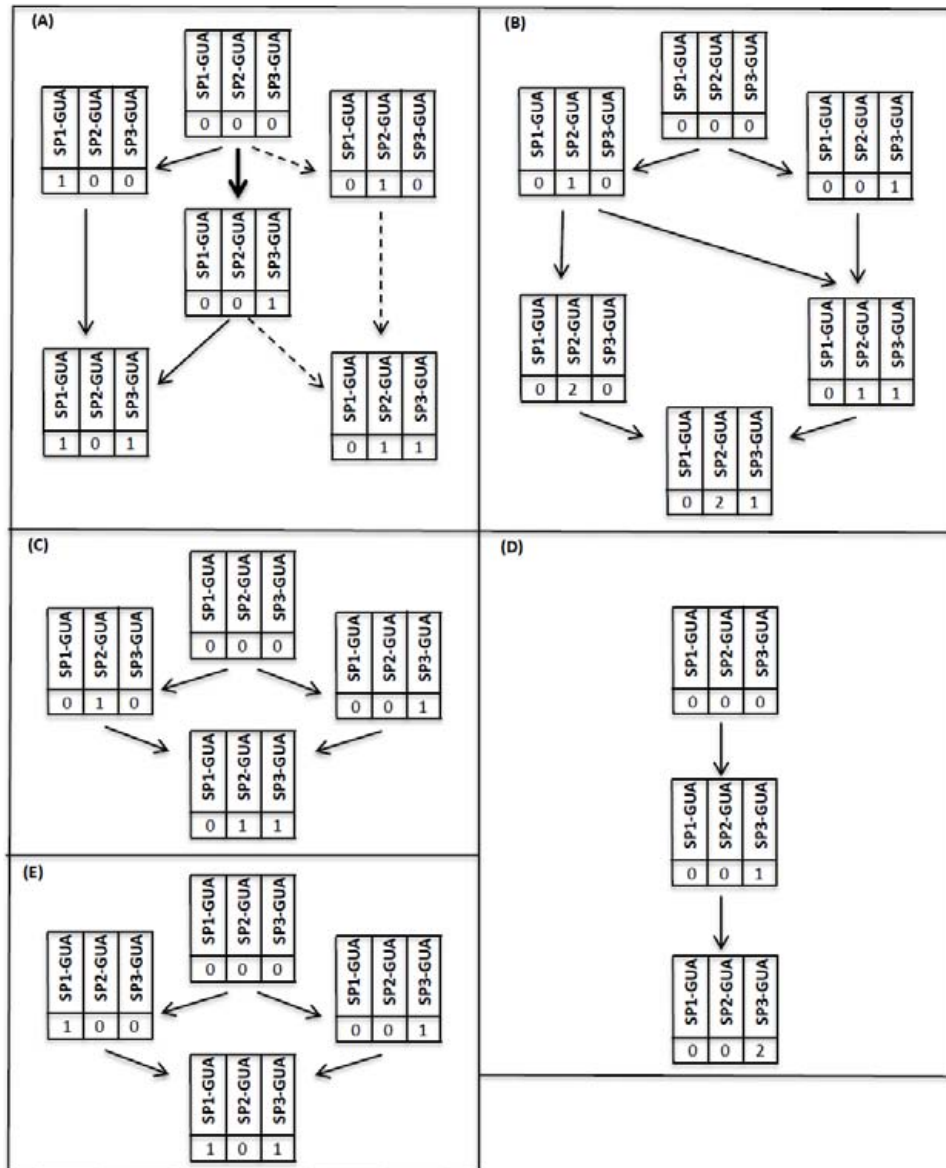


**Fig. 6.** Effect of the environment on the distribution of *sp1*, *sp2* and *sp3* colonizing guavas. State transition graphs correspond to the dynamic analysis of the network shown in Figure 3 in different scenarios. 0 means absence, 1 means abundance in the case of Boolean variable SP1-GUA and low presence in the case of multivalued variables SP2-GUA and SP3-GUA, 2 means abundance in case of multivalued variables SP2-GUA and SP3-GUA. **(A)** The three species *sp1*, *sp2* and *sp3* simultaneously invade the sea level region with the environmental factor EF = 1. The same dynamics occurs when the species *sp2* and *sp3* but not *sp1* simultaneously invade the sea level region. **(B)** The three species *sp1*, *sp2* and *sp3* simultaneously invade the intermediate region with the environmental factor EF = 2.

the one found in nature. The detailed dynamics presented in Fig. 7A show that the pathway that will be followed depends on the presence/absence of *sp1*. Again, the observed state found in nature is that in which *sp1* is present and *sp2* is absent in agreement with the idea that *sp1* overrides *sp2* and/or *sp1* arrived first to the region.

(2) The next simulation considers that *sp2* and *sp3* but not *sp1* invade an environmental area. The colonization dynamics of the sea level niche (EF = 1) is the same as in the scenario where the three species simultaneously invaded this region, reaching a single stable state [SP1-GUA = 0 & SP2-GUA = 2] (Fig. 6A and Table 4), corresponding to the absence of *sp1* and abundance of *sp2* and

*sp3*. This is so because the environmental factor (EF = 1) negatively affects *sp1*. The detailed colonization dynamics of the intermediate level niche (EF = 2) is shown in Fig. 7B and Table 4. A single final stable state is reached where *sp2* is abundant, *sp3* is less abundant and *sp1* is absent since it does not invade the region under this simulation. The detailed colonization dynamics of the high level niche (EF = 3) is shown in Fig. 7C and Table 4. A single final stable state is reached where *sp2* and *sp3* are less abundant, and *sp1* is absent since it does not invade the region under this simulation. Thus, under a hypothesis that *sp1* would be absent at the inland plateau, *sp3*, and also *sp2* would populate guavas in the two environmental areas.



**Fig. 7.** Effects of the environment on the distribution of *sp1*, *sp2* and *sp3* colonizing guavas. State transition graphs are similar to Fig. 6. **(A)** The three species simultaneously invade the high level region with the environmental factor  $EF = 3$ . **(B)** The species *sp2* and *sp3* but not *sp1* simultaneously invade the intermediate region ( $EF = 2$ ). **(C)** The species *sp2* and *sp3* simultaneously invade the high region ( $EF = 3$ ). **(D)** The species *sp1* and *sp3* but not *sp2* simultaneously invade the sea level region ( $EF = 1$ ). **(E)** The species *sp1* and *sp3* but not *sp2* simultaneously invade the intermediate region ( $EF = 2$ ). The same dynamics is predicted for colonization of the high region with  $EF = 3$ .

(3) The third simulation considers *sp1* and *sp3* but not *sp2* invading an environmental area. The model predicts a single stable state characterized by the abundant presence of only *sp3*, whereas *sp1* is absent, because it cannot live under  $EF = 1$  conditions, and *sp2* is absent since it does not

invade the region under this simulation. The detailed colonization dynamics of the sea level niche ( $EF = 1$ ) is shown in Fig. 7D and Table 4. The detailed colonization dynamics of the intermediate level niche ( $EF = 2$ ) is shown in Fig. 7E and Table 4. The model predicts a single

stable state characterized by the presence of *sp1* and less abundant presence of *sp3*, and *sp2* is absent because it does not invade the region under this simulation. The same dynamics is predicted for colonization of the high level niche,  $EF = 3$  (Table 3).

## DISCUSSION

The three species of the *Anastrepha fraterculus* complex inhabit a large extension in eastern Brazil where they are found in allopatry but are sympatric and syntopic in some areas, and use diverse fruits for larvae development [5, 7, 8]. The present analysis was aimed to understand the putative forces that shaped the present distribution pattern of these species. For this purpose, a methodology that has been successfully used to analyse the dynamics of regulatory genetic networks generating morphogenetic patterns during development was applied. [20-28]. In the majority of cases the knowledge we have about an ensemble, that is, its components and their interactions, is largely qualitative, which is formulated in verbal models, whose nature precludes testing their adequacy from a dynamical point of view. This is the situation of the ecological ensemble -the insects and their fruit hosts- facing here. The methodology applied, however, supports an integrative understanding of the inter-dependent behaviours of the ecological components involved, and it further allows to exploring the necessity of additional players and/or new emergent interactions to ensure a correct functioning of the ecological system at stake. Indeed, only through a formal analysis can we understand properties that emerge from the global consideration of the ecological ensemble.

Based on reported results concerning the distribution of the species of AF complex in Brazil [5, 7, 8, 13, 29, 30], a “species network” with the three species and their fruit hosts was constructed, thus resembling regulatory genetic networks. The dynamical analysis of this “species network” provided insights on the relationships between the species themselves and of them with their host fruits, as well as with the environment. It was found that two factors, the uni-directional incompatibility between species *sp1* and *sp2*, with the former overriding the second, and a yet unknown environmental factor, played an instrumental role

in the generation of the distribution pattern observed in nature.

Following the results above described, two scenarios can account for the history of the present distribution of *sp1* and *sp2* in the valley of Paraíba do Sul River. Considering that guava is a native fruit and orange an exotic one, it was previously proposed that *sp1* invaded first those regions colonizing guavas and that *sp2* arrived later colonizing the exotic oranges [8]. The present analyses employed these facts but the simulations considered distinct alternatives. Recall that *sp1* cannot explore oranges. When *sp2* arrived later it did not populate guavas -though it has the capacity to do so- because *sp1* prevents it. Consequently, *sp2* colonizes oranges. An alternative scenario could be that *sp2* arrived first and colonized both guavas and oranges. The later invasion of *sp1* could displace *sp2* from guavas, due to the incompatibility of *sp1* upon *sp2*, and the latter species will remain in oranges where it is protected from *sp1* as this cannot live in oranges. This is compatible with the idea that *sp1* overrides *sp2* and/or *sp1* arrives into the region first [8]. This new insight indicates that the order in the influx of both species would be irrelevant with regard to their distribution, being the incompatibility of *sp1* upon *sp2* the crucial, key factor shaping the observed distribution. Although possible from a theoretical point of view, we believe that the second scenario seems less plausible because it would be expected *a priori* to find guavas inhabited by *sp1* and others populated by *sp2* in a given area -contrary to experimental observation- unless enough time has passed for *sp1* to be able to displace *sp2* from all guavas.

This approach suggests that future investigations should be addressed to get insights on the factors involved in the incompatibility between species and in the effect of environment. Moreover, it pointed out that the ability to explore a host fruit for larval development might be more important than the order of arrival of a species in a given area. The formal generalized logical formalism applied to an extensive knowledge about species geographic distribution and host usage, can not only provide insights into the dynamics involved in the extant distribution of fruit flies and/or other insect species but also contribute to predictions about expected results after the introduction of species, or the eradication of species, in a given area.

## CONCLUSION

The dynamics underlying the distribution pattern of *Anastrepha fraterculus* complex in a Brazilian region was analysed using the generalized logical methodology. The results of such an analysis showed that the distribution pattern found in nature is a consequence of at least two factors. One is the uni-directional incompatibility between species *sp1* and *sp2*, with the former overriding the second. The second is a yet unknown environmental factor. Finally, this study has shown that the methodology used is suitable for analysing the forces involved in the creation of different insect distribution patterns in nature.

## ACKNOWLEDGMENTS

This study was supported by a grant from Fundação de Amparo à Pesquisa do Estado de São Paulo, Brazil (FAPESP 2016/00782-7) to DS. LS received a financial support from the Coordenação de Aperfeiçoamento de Pessoal de Nivel Superior (CAPES/PROEX-2016) for a short stay at the Instituto de Biociências, Universidade de São Paulo, Brazil.

## CONFLICT OF INTEREST STATEMENT

The authors declare no competing interest.

## REFERENCES

- White, I. M. and Elson-Harris, M. M. 1992, Fruit Flies of Economic Significance: Their Identification and Bionomics, CAB International, Wallingford, UK.
- Oliveira, C. M., Auad, A., Mendes, S. M. and Frizzar, M. R. 2012, J. Appl. Entomol, 137, 1-15.
- Taira, T. L., Abot, A. R., Nicácio, J., Uchôa, M. A., Rodrigues, S. R. and Guimarães, J. A. 2013, Rev. Brasil. Entomol., 57, 300-308.
- Selivon, D. and Perondini, A. L. P. 1998, Ann. Entomol. Soc. Am. USA, 91, 473-478.
- Selivon, D. and Perondini, A. L. P. 2007, Moscas de la Fruta en Latinoamérica (Diptera: Tephritidae): Diversidad, Biología y Manejo. V. Hernández-Ortiz (Ed.), SyG Editores, Distrito Federal, México, 101.
- Selivon, D., Perondini, A. L. P. and Morgante, J. S. 1999, Gen. Mol. Biol., 22, 507-510.
- Selivon, D., Vretos, C., Fontes, L. and Perondini, A. L. P. 2004, Proc. 6th International Fruit Flies Symposium. B. Barnes (Ed.), Isteg Scientific Publications, Irene, South Africa, 253.
- Selivon, D., Perondini, A. L. P. and Morgante, J. S. 2005, Ann. Entomol. Soc. Am. USA, 98, 367-381.
- Goday, C., Selivon, D., Perondini, A. L. P., Greciano, P. G. and Ruiz, M. F. 2006, Cytogen. Gen. Res., 114, 70-76.
- Hernández-Ortiz, V., Bartolucci, A. F., Morales-Valles, P., Frías, D. and Selivon, D. 2012, Ann. Entomol. Soc. Am. USA, 105, 305-318.
- Hernández-Ortiz, V., Canal, N. A., Tigrero, J. O., Ruíz-Hurtado, F. M. and Dzúl-Cauich, J. F. 2015, Resolution of Cryptic Species Complexes of Tephritidae Pests to Enhance SIT Application and Facilitate International Trade. M. De Meyer, A. R. Clarke and J. Hendrichs (Eds.), ZooKeys, 540, 75-124.
- Hasui, Y., Carneiro, C. O. R., Almeida, F. F. M. and Bartorelli, A. 2012, Geologia do Brasil. Beca-BALL Edições, São Paulo.
- Prezotto, L. F., Perondini, A. L. P., Hernández-Ortiz, V., Marino, C. L. and Selivon, D. 2017, Syst. Appl. Microbiol., 40, 59-67.
- Kovaleski, A., Uramoto, K., Sugayama, R. I., Canal, N. A. and Malavasi, A. 1999, Rev. Brasil. Entomol., 43, 229-234.
- Thomas, R. and D'Ari, R. 1990, Biological Feedback, CRC Press, Boca Ratón, Florida.
- Chaouiya, C., Rem, E., Mossé, B. and Thieffry, D. 2003, Lecture Notes, Contr. Information Sci., 294, 119-126.
- Miranda, E. E. 2005, (Coord.) Brasil em Rel evo. Embrapa Monitoramento por Satélite, Campinas. <http://www.relevobr.cnpm.embrapa.br> [last accessed May 20, 2017].
- Gontijo-Pascutti, A. H. F., Hasui, Y., Santos, M., Soares Jr., A. V. and Souza, I. A. 2012, Geologia do Brasil. Y. Hasui, C. O. R. Carneiro, F. F. M. Almeida and A. Bartorelli (Eds.), Beca-BALL Edições, São Paulo, 549.
- Chaouiya, C., Naldi, A. and Thieffry, D. 2012, Methods in Molecular Biology, 1, Bacterial Molecular Networks, Part 3, 84, 463-479.

- 
20. Sánchez, L., van Helden, J. and Thieffry, D. 1997, *J. Theor. Biol.*, 189, 377-389.
  21. Sánchez, L. and Thieffry, D. 2001, *J. Theor. Biol.*, 211, 115-141.
  22. Sánchez, L. and Thieffry, D. 2003, *J. Theor. Biol.*, 224, 517-537.
  23. Thieffry, D. and Sánchez, L. 2003, *Curr. Opi. Gen. Develop.*, 13, 326-330.
  24. González, A., Chaouiya, C. and Thieffry, D. 2006, *Genet.*, 174, 1625-1634.
  25. González, A., Chaouiya, C. and Thieffry, D. 2008, *Bioinform.*, 15, 234-240.
  26. Sánchez, L., Chaouiya, C. and Thieffry, D. 2008, *Int. J. Develop. Biol.*, 52, 1059-1075.
  27. Fauré, A., Vreede, B. M., Sucena, E. and Chaouiya, C. 2014, *PLoS Comput. Biol.*, 10, 415, e1003527.
  28. Sánchez, L. and Chaouiya, C. 2016, *BMC Syst. Biol.*, 10, 37.
  29. Rull, J., Abraham, S., Kovaleski, A., Segura, D. F., Islam, A., Wornoyaporn, V., Dammlage, T., Santo Tomas, U. and Vera, M. T. 2012, *Bull. Entomol. Res.*, 102, 435-443.
  30. Roriz, A. K. P., Japyassú, H. F. and Joachin-Bravo, I. S. 2016, *Entomol. Exp. Appl.*, 162, 346-357.