Speciation with gene flow in a heterogeneous virtual world: can physical obstacles accelerate speciation?

Abbas Golestani1,*, Robin Gras1,2 and Melania Cristescu2,3

1School of Computer Science, 2Department of Biology, and 3Great Lakes Institute, University of Windsor, Windsor, Canada

The origin of species remains one of the most controversial and least understood topics in evolution. While it is being widely accepted that complete cessation of gene-flow between populations owing to long-lasting geographical barriers results in a steady, irreversible increase of divergence and eventually speciation, the extent to which various degrees of habitat heterogeneity influences speciation rates is less well understood. Here, we investigate how small, randomly distributed physical obstacles influence the distribution of populations and species, the level of population connectivity (e.g. gene flow), as well as the mode and tempo of speciation in a virtual ecosystem composed of prey and predator species. We adapted an existing individual-based platform, EcoSim, to allow fine tuning of the gene flow's level between populations by adding various numbers of obstacles in the world. The platform implements a simple food chain consisting of primary producers, herbivores (prey) and predators. It allows complex intra- and inter-specific interactions, based on individual evolving behavioural models, as well as complex predator–prey dynamics and coevolution in spatially homogenous and heterogeneous worlds. We observed a direct and continuous increase in the speed of evolution (e.g. the rate of speciation) with the increasing number of obstacles in the world. The spatial distribution of species was also more compact in the world with obstacles than in the world without obstacles. Our results suggest that environmental heterogeneity and other factors affecting demographic stochasticity can directly influence speciation and extinction rates.

Keywords: speciation; gene flow; individual-based simulation; predator–prey system

1. INTRODUCTION
The relative contribution of geography and ecology to speciation remains one of the most controversial topics in evolutionary biology. Models of speciation that involve geographically unrestricted gene flow (sympatric speciation) or limited gene flow (parapatric speciation) are often considered unrealistic. The major theoretical problems with models that assume gene flow stem from the antagonism between selection and recombination and from the problem of coexistence [1]. It is generally assumed that while selection acts to maximize the fitness optimum of populations, generating genetic and phenotypic divergence, recombination continuously shuffles the co-adapted gene complexes and brings populations together. Moreover, sister species that are not sufficiently ecologically divergent are believed to experience competitive exclusion that leads to rapid extinction of emerging lineages [1]. The wide range of theoretical conditions that diminish these major conflicts [2,3] are considered by many critics to be biologically unrealistic, maintaining the long-lasting debate over the likelihood of speciation with gene flow.

While placing level of gene flow at the centre of speciation debates has extremely been successful in shaping research programmes and directions [1], the simple dichotomy of sympaty and allopatry along with the static spatial (biogeographical) context also has the potential to hinder progress in the field. It has been suggested that conclusions reached about the relative importance of various mechanisms of speciation can be drastically different if investigators use explicit geographical-pattern (biogeographical) concepts versus more demographic (population genetic) criteria that imply a strict condition of original panmixia outside the geographical context (e.g. sympatric, allopatric) [4]. Clearly, many species have strong schooling or homing behaviours or strong ecological preferences that result in a non-random distribution of genetic diversity at the onset of speciation. Moreover, habitat heterogeneity can often enhance the local structuring of genetic variation. Such strategies make the distinction between sympatric and micro-allopatric speciation scenarios hard to disentangle without very good knowledge of the early stages of speciation. Moreover, the intense debate over the geography of speciation has often left biologically relevant scenarios, such as the intermediate parapatric conditions, out of the research context.

There is no doubt that the complexity of natural systems poses a great challenge when one tries to assign speciation cases to discrete categories. Most species exhibit dynamic changes in distribution that involve population expansions and contractions, fragmentations and secondary contacts across evolutionary relevant time scales [4–6]. Moreover, macro-geographical barriers are also ephemeral on a larger geological scale. At a fine local scale, the effect of micro-geographical barriers depends largely on how important the structure of the habitat is for dispersal rates. It has recently been proposed that the complex context of speciation can be better understood outside the framework of a
classical geographical definition by focusing on the important evolutionary forces such as gene flow, selection and genetic drift [4].

In the area of ecosystem simulation, individual-based modelling provides a bottom-up approach allowing for the consideration of the traits and behaviour of individual organisms. Instead of modelling an ecosystem as a whole, individual-based models treat individuals as ‘unique and discrete evolutionary entities’ [7]. By modelling organisms with varying characteristics (such as age, mating preferences and role in the ecosystem), the properties of the system can emerge from their complex interactions. This approach has advanced in various areas of application such as forest ecology, fish recruitment modelling and spatial heterogeneity depiction [8], and can effectively be used in speciation research. Yet, few attempts have been made to simulate a complex ecosystem. An example of such a system is the platform Echo [8], which includes an evolutionary mechanism allowing virtual agents to evolve. However, the organisms in Echo are relatively simple, and have no behavioural model. Another system often used to study long-term evolution is Avida [9] which allows studying emergence of complex behaviours in populations of evolving digital organisms. It nevertheless has limitations on the movement of individuals and is based on one or more external fitness function(s), which means that the system is more an optimization process rather than a complex virtual world. Other models, such as PolyWorld [10], BubbleWorld.Evo [11] or Framsticks [12], include more complex agents and behavioural models. They use artificial neural networks or systems of learned rules to evolve the agent’s behavioural model during their life or across generations. Unfortunately, these approaches are highly computationally expensive, allowing the implementation of populations composed of only few hundred agents. They are, therefore, more appropriate for studying the evolution of learning capacities than for studying large-scale evolutionary processes. More specialized predator–prey models have also been proposed [13], yet these particular models are dedicated to represent specific schooling behaviours, and the evolution is an offline mechanism using a genetic algorithm suggesting that the direction of the evolution is pre-determined.

In this study, we use an individual-based simulation approach to investigate how physical obstacles (the raggedness of the environment) influence population connectivity (e.g. gene flow), the distribution of species in space and time and ultimately the genomic cohesion of species and speciation rates. We explore how randomly placed physical obstacles that can be easily circumvented by agents, influence the pattern and rate of speciation. The major advantage of using an individual-based simulation approach in speciation research is that it allows for the modification and fine tuning of one property of the whole system at a time. The effects of these modifications can be determined within a reasonable time frame and over a large evolutionary scale.

2. MATERIAL AND METHODS

In this study, we use EcoSim (http://sites.google.com/site/ecosimgroup/research/ecosystem-simulation) [14], a versatile simulation platform that has been designed to investigate several broad ecological questions (e.g. community formation), as well as long-term evolutionary patterns and processes such as speciation and macroevolution. In this programme, two organism types, prey and predator, are simulated in a torus-like discrete world which is a 1000 × 1000 matrix of cells. Every cell can contain some amount of grass (varying in time) which is the primary producer. To observe the evolution of individual behaviour and ultimately ecosystems over thousands of generations, several conditions need to be fulfilled: (i) every individual should possess genomic information; (ii) this genetic material should affect the individual behaviour and consequently its fitness; (iii) the inheritance of the genetic material has to be done with the possibility of modification; (iv) a sufficiently high number of individuals should coexist at any time step and their behavioural model should allow for complex interactions and organizations to emerge; (v) a model for species identification, based on a measure of genomic similarity, has to be defined; and (vi) a large number of time steps need to be performed. These complex conditions pose computational challenges and require the use of a model which allows the compactness and easiness of computation with a high potential of complex representation.

EcoSim uses a modified version of the fuzzy cognitive map (FCM) model [15] adapted for behavioural modelling. The FCM, coded in the genome of the individuals, is used not only as the behavioural model of the individuals but also as the vector of transmission of evolutionary information [14]. The genomes are coded with real numbers, and are therefore continuous. This approach offers compactness with a very low computational requirement, while having the capacity to represent complex notions. Although, each agent is represented by a unique map, which is an inherited modified combination of its parental genomes, the system can still manage several hundreds of thousands of agents simultaneously in the world. This FCM contains sensory inputs (e.g. predator_close, prey_close, food_close, mate_close, energy_low), internal concepts (e.g. fear, hunger, sexual need, curiosity, satisfaction) and motor concepts (e.g. escaping, searching for food, eating, breeding, socializing or trying to approach a potential sexual partner). Additionally, it includes links and weights representing the mutual influences (allowing feedback loops) of these concepts. Moreover, the activation level of each motor concept depends on a complex and a nonlinear combination of excitatory and inhibitory influences from both sensory inputs and internal concepts. Therefore, the action performed by an individual at a given time step is the one corresponding to the motor concept with the highest activation level (for details, see the electronic supplementary material, §1, figure S1). Links between concepts can appear or disappear, so that the structure and complexity of the maps can change during the evolutionary process.

Each agent also possesses several physical or life-history characteristics summarized in table 1. The energy is provided by the primary or secondary resources found in their environment. For example, prey individuals gain 250 units of energy by eating one unit of grass and predators gain 500 units of energy by eating one prey. At each time step, each agent spends energy depending on its action (e.g. breeding, eating, running) and on the complexity of its behaviour model (number of existing edges in its FCM). On an average, a movement action such as escape and exploration requires 50 units of energy, a reproduction action uses 110 units of energy and the choice of no action results in a small expenditure of 18 units of energy. The food chain implemented in
Table 1. Several physical and life-history characteristics of individuals.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Predator</th>
<th>Prey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum age</td>
<td>42 time steps ± 0–6</td>
<td>46 time steps ± 0–18</td>
</tr>
<tr>
<td>Minimum age of reproduction</td>
<td>8 time steps</td>
<td>6 time steps</td>
</tr>
<tr>
<td>Maximum speed</td>
<td>11 cells per time step</td>
<td>6 cells per time step</td>
</tr>
<tr>
<td>Vision distance</td>
<td>25 cells maximum</td>
<td>20 cells maximum</td>
</tr>
<tr>
<td>Level of energy at initialization of the system</td>
<td>1000 units</td>
<td>650 units</td>
</tr>
<tr>
<td>Average speed</td>
<td>1.4 cells per time step</td>
<td>1.2 cells per time step</td>
</tr>
<tr>
<td>Average level of energy</td>
<td>415 units</td>
<td>350 units</td>
</tr>
<tr>
<td>Maximum level of energy</td>
<td>1000 units</td>
<td>650 units</td>
</tr>
<tr>
<td>Average number of reproduction action during life</td>
<td>1.14</td>
<td>1.49</td>
</tr>
<tr>
<td>Average length of life</td>
<td>16 time steps</td>
<td>12 time steps</td>
</tr>
</tbody>
</table>

Our system consists of three levels; primary producers, predators and prey, allowing complex interactions between agents with and between trophical levels.

One of the actions performed by the individuals is reproduction. Several factors play roles in reproduction. For reproduction to be successful, the two parents need to be in the same cell, have enough energy, choose the reproduction action and be genetically similar. The organisms cannot determine their genetic similarity with their potential partner. They try to mate and if the partner is too dissimilar, the reproduction fails. The result of the reproduction action is a unique offspring with a genome which represents a combination of the parental genomes. The newborn receives an initial amount of energy equivalent to the energy that the two parents spend in reproduction.

EcoSim implements a species concept directly related to the genotypic cluster definition [16] in which a species is a set of individuals associated with the average of the genetic characteristics of its members. The species map is computed based on the average FCM matrices of all individuals. It is considered that a species splits if the distance between the genomes of the two most dissimilar agents is greater than a predefined threshold [17,18]. Our species recognition method involves the use of a 2-means clustering algorithm in which an initial species is split into two new species, each one of them containing the agents that are mutually the most similar. Since the birth and death of individuals influences the general composition of species [19] and speciation and extinction can occur at any time step, species membership is evaluated at each time step. Although speciation events are discrete, the complete speciation process is spread over multiple time steps. With time, a species will progressively contain individuals that are genetically more and more dissimilar up to an arbitrary threshold where it is considered that the species split. After splitting, the two sister species are still very similar and hybridization events can occur. Two individuals can interbreed if their genomic distance is smaller than an arbitrary threshold (half of the speciation threshold) even if they are designated as members of two sister species by our clustering algorithm (for details, see the electronic supplementary material, §2, figure S2). The frequency of hybridization events ranges between 2 and 5 per cent of the overall number of reproductive events. On average, prey and predator individuals live for 12 and 16 time steps, respectively, with some individuals living for up to 60 time steps. If we consider that a generation is the time needed for a new born individual to have its first reproduction, a generation is about eight time steps. This indicates that a single run with 15 000 time steps corresponds to about 2000 generations.

In terms of computational time, the speed of simulation per generation is related to the number of individuals. Recent executions of the simulation with an average of 250 000 individuals produced approximately 15 000 time steps in 35 days. Each time step involves the time needed for each agent to perceive its environment, make a decision, perform its action, as well as the time required to update the species membership, including speciation events and record relevant parameters (e.g. the quantity of available food). Although a run can involve the birth of more than one billion individuals and thousands of species, all evolutionary events, the mental state and action of every agent, are saved for every time step of every run. This thorough tracking system enables us to extract, to measure and to correlate parameters that are useful to understand the underlying properties of such a complex system.

Several studies evaluated the capacity of the EcoSim platform to model real ecosystems and make realistic predictions regarding species abundance patterns [19] and the complexity levels of the simulation [18]. These studies show that the communities of species generated by the simulation follow the same lognormal law as natural communities and that EcoSim can help evaluate the overall level of diversity of a given community. Moreover, it was shown that the simulation has a deterministic chaotic behaviour [20], similar to many natural systems [21], and that the spatial distribution of the individuals has multi-fractal characteristics [22].

In order to measure the effect of the raggedness of the environment on population fragmentation and the speciation process, we included small physical obstacles that obstruct the movement (dispersal) of agents. The presence of obstacle cells in the world is expected to impede the movement of our agents, change their spatial distribution, and in turn influence the gene flow between populations [for details, see the electronic supplementary material, §3]. The reduction of gene flow between populations is proportional to the raggedness of the environment. We control the level of gene flow by changing the percentage of obstacle cells. We investigated how this impediment in the movement of organisms, without any complete extrinsic barrier separating two subpopulations of an initial species, affects the speciation frequency and the number of coexisting species. We compared three different situations. We considered a neutral configuration with no obstacles (the ‘density of obstacles (0%)’ experiment). We also considered two virtual worlds with various numbers of obstacles: 1 and 10 per cent.
For example, in the experiment ‘density of obstacle (10%)’, 10 per cent of cells in the world are obstacles. For each experiment, we conducted 10 independent runs using the same parameters and averaged the results. To ensure that our results are not dependent on special parameter values, several speciation thresholds for prey (0.65, 1.3, 2.6) and predator (0.75, 1.5, 3) were used. As the results for the three speciation thresholds were very similar, we averaged them. All the results represent the average of 30 experiments (10 runs × three speciation thresholds). To avoid any bias owing to a variation in the number of free cells available after the addition of obstacles, we maintained the number of free cells constant by increasing the world size accordingly. We analysed the variation of the number of species for both prey and predator during the simulation process for several numbers of obstacles. We also observed other properties of the whole system, such as individual behaviours, spatial distribution of species and the assembly and dynamics of ecological communities.

3. RESULTS AND DISCUSSIONS

(a) Global patterns
We investigated the global behaviour of our system in different situations, by varying the number of obstacles. We measured and monitored several representative characteristics of the system, such as the number of species, individual behaviour and spatial distribution of the individuals that can give insight into the evolutionary processes that shape the biodiversity of our virtual world. An overview of the distribution of species reveals that individuals show a strong clustering distribution with circular or spiral shapes.

Individual's distribution forming spiral waves is a common property of predator–prey models (see figure 1). The prey near the wave break has the capacity to escape from the predators sideways. A subpopulation of prey then finds itself in a region relatively free from predators. In this predator-free zone, prey individuals start dispersing rapidly forming a circular expanding region. The predation pressure creates successive interactions between prey and predators over time, the same pattern repeats over and over again, leading to the formation of spirals. Strong and robust spiral waves have been commonly observed in complex and dynamic biological systems [23]. Self-organized spiral patterns have been seen not only within chemical reactions [24] but also among host–parasitoid or predator–prey systems [23,25–27] even when the world is uniform in terms of environment’s raggedness [25,27]. Our system is closer to a predator–prey system than a host–parasite system. Both the behaviour of the species (dispersal ability), population numbers of prey and predators, life-history characteristics (life span, energy expenditure), reflect more a predator–prey system than a host–parasite system. The global pattern of population distribution in the world with obstacles is very

![Figure 1. An overview of the distribution of species and populations in the world with density of obstacles (10%) and the density of obstacles (0%) experiment.](image)

- (a) View of the whole world in the density of obstacles (0%) experiment.
- (b) Magnified part of the world in density of obstacles (0%) experiment.
- (c) View of the entire world with obstacles.
- (d) Magnified part of the world with obstacles. The blue squares are obstacle cells and dots are individuals. Different coloured dots represent different prey species and white dots represent predator species.

![Image]
similar to the one observed in the world of the experiment with density of obstacles (0%) (figure 1) and these oscillatory dynamics are probably owing to the predator–prey interaction [28]. It has been observed that the size and number of spirals in all the experiments are almost the same (table 2). Therefore, the existence of such patterns is unlikely to explain the differences in speciation rates between different experiments.

(b) Species richness and relative species abundance

The most important quantities of the system that we monitored through time were species richness (the total number of species), as well as species abundance (the number of individuals per species) in the world during the entire simulation process and across multiple simulations. Our results indicate that the number of species for both prey and predators increases directly with the number of obstacles (figure 2). Our results reveal that the total number of prey and predator species in the world is higher in the two configurations with obstacles compared with the no obstacles configuration. Moreover, the number of species in the configuration with 10 per cent obstacles is much higher than the number of species for another configuration with obstacles.

The computed average and standard deviation of the number of species during the whole process for prey population (table 3) reveal clear differences between the three experiments. This suggests that the speciation rate is directly proportional to the restriction of movement and therefore to gene flow between populations. As the total numbers of individuals in the three configurations are almost the same, it follows that the number of individuals per species decreases when obstacles are added in the world.

The computed ANOVA statistical tests [29] revealed that the observed differences in species richness between the experiments were significantly different (table 4). Even if the values obtained are higher than the 5 per cent confidence threshold, the difference observed between the inter-class values and the between-classes values strongly confirms that the variations observed for the number of species are highly significant for prey species. The average number of species with the density

---

**Figure 2.** Comparison between numbers of (a) prey and (b) predator species in the whole world during 16 000 time steps. Each curve represents an average value obtained from 30 independent runs with three different speciation thresholds (blue line, density of obstacles (0%); green line, density of obstacles (1%); red line, density of obstacles (10%)).

**Table 2.** Average and standard deviation of the number and size of spirals in the 30 independent runs of every configuration.

<table>
<thead>
<tr>
<th>density of obstacle (%)</th>
<th>no. spirals mean</th>
<th>size of spirals in cells mean</th>
<th>size of spirals in cells s.d.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>41</td>
<td>186</td>
<td>67</td>
</tr>
<tr>
<td>1%</td>
<td>38</td>
<td>188</td>
<td>58</td>
</tr>
<tr>
<td>10%</td>
<td>43</td>
<td>181</td>
<td>73</td>
</tr>
</tbody>
</table>

**Table 3.** Average and standard deviation of the number of species in the 30 independent runs for every configuration.

<table>
<thead>
<tr>
<th>density of obstacle (%)</th>
<th>Nr prey species mean</th>
<th>s.d.</th>
<th>median</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>27</td>
<td>9</td>
<td>25</td>
</tr>
<tr>
<td>1%</td>
<td>68</td>
<td>16</td>
<td>61</td>
</tr>
<tr>
<td>10%</td>
<td>94</td>
<td>19</td>
<td>82</td>
</tr>
</tbody>
</table>

**Table 4.** Measure of distribution dissimilarity by ANOVA test for the time series representing the number of prey species. (Each comparison includes either two different experiments or a single experiment compared with itself. For two distributions, the higher the value is, the more dissimilar the distributions are. When a single distribution is considered, the higher the value is, the higher the variance between the 30 runs of the same configuration is.)

<table>
<thead>
<tr>
<th>density of obstacle (%)</th>
<th>density of obstacle (1%)</th>
<th>density of obstacle (10%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>4.3016 × 10^3</td>
<td>2.5184 × 10^4</td>
</tr>
<tr>
<td>1%</td>
<td>116.629</td>
<td>517.2141</td>
</tr>
<tr>
<td>10%</td>
<td>517.2141</td>
<td>93.1408</td>
</tr>
</tbody>
</table>
of obstacles (10%) configuration is more than five times that of the density of obstacles (0%) configuration.

Our results clearly show that the addition of obstacles in the world is associated with an increase in the number of species. Population genetic theory predicts that natural selection and genetic drift cause populations to diverge from each other while migration resulting in gene flow acts in an opposite direction creating genetic homogeneity. We suggest that obstacles lead to an impediment in dispersal, more geographical isolation, less migration and gene flow. This overall lower level of population connectivity leads to rapid differentiation. Eventually, populations will contain individuals with genome dissimilarities higher than the speciation threshold, leading to speciation.

(c) Variation in individual behaviours
The results on species richness confirm that the restriction in the movement of species owing to scattered physical obstacles is strongly correlated with the frequency of speciation events. In order to verify that the increased speciation rate cannot be explained by other factors such as the change in the behaviour of the agents, we monitored the actions performed by prey individuals in the three configurations. Most of the actions chosen by the individuals (e.g. feeding, predator avoidance, prey chasing) were the same in the three configurations although a slight difference was identified on reproduction and socialization (figure 3). The decision for the individual to conduct these actions is done first using their behavioural models. Different situations can then lead to a failure of individuals to execute these actions. The reproduction action fails either because the agent cannot find a partner for reproduction, or has insufficient energy, or the two partners are genetically too different. The action of socialization fails because the agent cannot reach the place where the chosen partners are located. Our results suggest that the number of failed socialization events is much higher and much more variable in the density of obstacles (0%) configuration than in all obstacle experiments (see figure 3b,d). This is probably a direct consequence of the fact that the species have much smaller spatial distribution in obstacle configurations, making the socialization action easier to perform (the genetic similarity between agents is not considered for this action). However, the number of failed reproduction actions is significantly higher when the ruggedness of the world increases (figure 3c). This is probably owing to the higher genetic distance between individuals often found in close proximity (data not shown). This result enforces the hypothesis that the presence of obstacle cells in the world increases the genetic distance and therefore the speciation rate even in situations where the heterospecific individuals are more spatially compact.

(d) Spatial distribution of populations and species
Spatial distribution of species was shown to have a strong influence on their extinction rates [30,31] as well as the relationship between physical distance and genetic distance [32]. For example, it has often been documented that the risk for species to become extinct increases when its distribution area is reduced and when its demography is low [30]. Moreover, a regular increase of genetic distance with increasing geographical distance has been predicted [32] and documented often in broad phylogeographical surveys or organisms with various life-history attributes and dispersal abilities [33].

To evaluate the spatial distribution of the species, we used a measure, based on an average distance of all the
members of a species to its physical centre. This measure is expressed in a number of cells and gives an accurate evaluation of the distribution area of a particular species in the world. The average and standard deviation of individuals’ average distances around the centre of each species taken from 10 independent runs show that the species have a more compact distribution in the obstacle versions (table 5).

Knowing that in a torus world of size 1000 × 1000 cells the largest possible distance between two points is about 700 cells, the average values observed for the density of obstacles (0%) configuration, which can be more than 180 cells, are quite large. This suggests that, in the density of obstacles (0%) configuration, many species have a widespread spatial distribution covering a large part of the world. By contrast, in the world with obstacles, species show a much more restricted geographical distribution which means that the species’ spatial distribution decreases proportionally with the increase in the number of obstacles. These results are also confirmed by the strong negative correlation between number of obstacles and the maximum observed spatial distribution of a species (table 6). The high value of the standard deviation for all configurations can easily be explained by the high variability of the number of individuals by species.

Our communities of species display a log-normal distribution pattern commonly found in nature [19]. This property leads to an important diversity in terms of number of individuals per species, which in turn explains the observed high variance in spatial distribution. Obstacle cells strongly affect the spatial distribution of the individuals by reducing the total number of subdivided populations that constitute a species (for details, see the electronic supplementary material, §3, figure S4). For most of the species, several spatially separated populations are observed in the density of obstacles (0%) configuration, whereas only one compact population is observed in the obstacle configurations.

(e) Fuzzy cognitive map evolution
The composition of species in our virtual world depends on the fine balance between speciation and extinction. The high species richness in the obstacle worlds could be owing to an accelerated speciation rate, a decelerated extinction rate, or a combination of both. In order to investigate the factors driving the biodiversity of the three virtual worlds, we analysed the level of genetic divergences between the initial genome and the genome of all individuals at every time step. In the initial populations, every prey or predator individual has the same genome. However, after roughly 50 time steps (about six generations), when all the founding individuals are dead, the genome of each agent is unique, and is the outcome of the evolutionary process. To evaluate the speed of evolution in our simulation, we compared the average distance [14] between all existing prey or predator genomes at any time step with the two initial genomes of prey and predators.

This average distance computed for a total of 5000 time steps (figure 4) indicates that the overall genetic divergence of the community of prey and predator species is greater in obstacle trials than in the density of obstacles (0%) experiment. The more obstacles in the world, the steeper the slope of the curve. This suggests that evolution accelerates with the number of obstacles.

Given that this result shows only global information (at the community level) about speciation patterns, changes at the intraspecific level or between closely related species are also very informative. We measured the speed of divergence between two sister species after a speciation event. In EcoSim, a species is associated with a genome which corresponds to the average genome of all its individuals allowing us to compute a distance between the ‘genome’ (called centre) of two species. We considered 20 independent speciation events for each of the three configurations. We then computed the distance between the centres of the two new species after the speciation event.
occurred across 250 time steps (figure 5). It can be noticed that, as expected, sister species diverge quite quickly after speciation. After speciation, hybridization events are quite rare because in each of the newly emerged species the individuals are highly similar, but dissimilar from the ones of the sister species. As a result, gene flow between the two species is probably very low and leads to fast divergence. More noteworthy, the speed of divergence is much higher when there are obstacles in the world. Once again, it is clear that this phenomenon is continuous, as the speed of divergence increases proportionally with the number of obstacles.

To understand if the reduction in gene flow is enough to explain the speed of divergence, we also considered the effect of obstacles on spatial distribution of sister species. We computed the average geographical distance between the physical centres of the emerging species after speciation events occurred and across 250 time steps for 20 speciation events (figure 6). We observed that the physical distance between species is smaller in the world with obstacles than that without. This result can be correlated to the fact that the number of individuals per species is smaller and the spatial distribution of individuals is more compact for an obstacle than density of obstacles (0%) configurations. It is interesting to note that even if there are high variations in these spatial distances, there are no visible trends to an increase of distance between species after speciation in the three configurations.

4. CONCLUSIONS
Using a modified EcoSim individual-based platform to implement various degrees of physical obstacles that restrict the movement of individuals and probably reduce gene flow, we compared three different configurations with different densities of obstacles. It is clear
that the speciation rate and species diversity is directly proportional to the roughness of the physical environment. Our study also reveals that species are more spatially compact in the configurations with obstacles than in the world of the experiment with density of obstacles (0%). Moreover, the continuous reduction in spatial distribution as the number of obstacles increases results in low levels of gene flow between sister species. Therefore, the rapid genomic divergence between species should be directly linked to the reduction of movement owing to obstacles that result in low gene flow and rapid divergence between subdivided populations. We investigated several factors that could be involved in the increase of speciation rate, such as the individual's behaviours, the spatial distribution of the species, or the overall speed of evolution (increase in genetic divergence between sister species). We show that the faster divergence between populations and accelerated speciation cannot be explained by an increase of spatial separation during the initial stage of speciation or different behaviours of the individuals. We suggest that this is probably owing to the significantly lower population sizes in obstacle configurations. This reduced size, results in more pronounced genetic drift and rapid differentiation between populations that experience relatively low levels of gene flow.

It is well accepted that the effect of micro-geographical barriers (e.g. the raggedness of the environment) to maintain population cohesion and the genetic homogeneity of a species depends heavily on the intrinsic properties of the species (e.g. dispersal ability, intra- and inter-specific interactions). We suggest that the complex context of speciation can be better understood outside the framework of a classical geographical definition of speciation (e.g. sympatric, allopatric) by focusing on the complex interactions at the community level.

Speciation and extinctions are very important processes that influence the species composition of an ecosystem at a particular time and the long-term dynamics of ecological communities. Our approach allows testing of the unified neutral theory of biodiversity and biogeography proposed by Hubbell [34] which suggests that the persistence of ecologically equivalent species in sympatry across relevant time scales might not depend strictly on complex niche differences. Our results suggest that factors affecting demographic stochasticity (e.g. factors shaping the density of individuals in a local area, the extent of the distribution of species in space) can influence speciation and extinction rates and ultimately the distribution of relative species abundance. Our approach has demonstrated its use in modelling several important biological problems and it seems possible to modify it to represent many new ones. However, the FCM model has some limitations because it cannot evolve new sensory inputs or new actions. The complexity of the model also grows with the square of the number of such concepts, limiting this application to relatively simple behavioural models. Since our simulation takes into account spatial information and individual behaviour while allowing the creation of new species and, more importantly, the growth rate of species is not fixed, it is difficult to determine if varying the growth rates of species or predator pressure would lead to different distributional patterns without testing it. A more in depth analysis of the effect of reproduction rates and predator pressure on the spiral formation is needed.

We thank D. Bock, J. Vaillant and B. MacPherson for comments on the manuscript. This work was supported by the NSERC grant ORGPIN 341854, the CRC grant 950–2-3617 and the CFI grant 203617 and is made possible by the facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET: www.sharcnet.ca).

REFERENCES
22 Golestani, A. & Gras, R. 2011 Multifractal phenomena in EcoSim, a large scale individual-based ecosystem simulation. Int. Conf. on Artificial Intelligence (ICAIF). Las Vegas: CSREA Press, pp. 991–999.