

Tese de Doutorado

Uma abordagem da física estatística
a alguns problemas de biologia evolutiva

Edgardo Brigatti

Centro Brasileiro de Pesquisas Físicas

Rio de Janeiro, 2006

Conteúdo

Agradecimentos	7
Resumo	9
Abstract	11
1 Introdução	13
1.1 Biologia evolutiva: a teoria da evolução	15
1.2 Estudos de dinâmicas evolutivas	21
1.2.1 Auto-organização e emergência	22
1.2.2 Paisagens	23
1.2.3 Comportamento de escala e criticalidade	26
1.2.4 Sistemas dinâmicos: dinâmica de limiar e criticalidade	28
1.2.5 Computação booleana	29
2 Propriedades de escala em um modelo de envelhecimento	31
2.1 O fenômeno do envelhecimento e sua explicação em bases evolucionárias	31
2.2 Modelos teóricos e computacionais no estudo do envelhecimento	35
2.3 O modelo Penna para populações assexuadas e haplóides . . .	35
2.3.1 Comentários em relação ao fator de Verhulst, ao ta- manho da tira de bits e à idade mínima de reprodução	36
2.4 Procurando um comportamento de escala no modelo Penna sexuado	38
2.4.1 Resultados	40
2.4.2 O processo de renormalização	48

2.4.3	Resultados posteriores	50
3	Microevolução, transição descontínua e sustentabilidade em um sistema ecológico ameaçado por uma caça seletiva	53
3.1	Sustentabilidade ecológica e fenômenos de microevolução . . .	53
3.2	Modelos para a descrição dos aspectos fenotípicos de uma população	55
3.3	Um modelo para a descrição das populações de cabras das Ram Mountain	58
3.3.1	O tamanho da população	62
3.4	Resultados	63
3.4.1	Os dois estados extremos da dinâmica	63
3.4.2	A transição descontínua	66
3.4.3	Como a dinâmica representa uma microevolução	68
3.5	Quais informações o modelo fornece?	70
3.5.1	O interesse por uma dinâmica não trivial	70
3.5.2	A reconstrução de dados experimentais e a possibilidade de fazer algum tipo de previsão	70
4	Competição e dinâmica evolutiva: o fenômeno da especiação simpátrica	73
4.1	O problema biológico da especiação	73
4.1.1	Uma vasta literatura de modelos	74
4.1.2	Recentes críticas ao modelo de Dieckmann e Doebeli	76
4.1.3	Populações assexuadas	77
4.2	O modelo	78
4.2.1	Seleção natural e competição	78
4.2.2	Acasalamento preferencial positivo	80
4.3	Resultados	81
4.3.1	Populações assexuadas	81
4.3.2	Populações sexuadas	87
4.4	Considerações finais	92

5	Especiação e macroevolução: micro e macroevolução são desacopladas ?	95
5.1	Os modos microevolucionários são extrapoláveis a escalas geológicas?	95
5.2	Estudos precedentes	96
5.3	O modelo	99
5.3.1	Dependência dos parâmetros	101
5.4	Resultados	103
5.4.1	Análise estatística das distribuições de frequências . .	103
5.4.2	Ausência de extinções de massa	108
5.5	Conclusões	111
	Apêndice	113

Agradecimentos

Um agradecimento particular ao meu orientador, Itzhak Roditi, que permitiu a realização deste trabalho de tese.

Esta pesquisa nasceu por meio de uma constante colaboração com o Professor Jorge de Sá Martins, do grupo de Mecânica Estatística da Universidade Federal Fluminense. Agradeço a ele e a todo o grupo de pesquisa da UFF por ter-me oferecido inúmeras oportunidades de aprender. Sou grato também ao Professor Dietrich Stauffer e ao colega Veit Schwammle.

Por fim um agradecimento a todo o pessoal do instituto CBPF, ao seu secretariado, em particular na pessoa de Myriam S. Coutinho, e às agências CAPES e CNPq pelo apoio financeiro.

Resumo

Este trabalho de tese foi realizado com o fim de desenvolver uma pesquisa interdisciplinar, aplicando técnicas herdadas da física teórica e computacional a problemas provenientes da área da biologia. Foram abordados diferentes temas, todos relacionados ao estudo da emergência de fenômenos coletivos em modelos microscópicos, baseados na dinâmica de agentes individuais, que simulam sistemas biológicos evolutivos.

Desenvolvemos um programa de pesquisa, a partir do bem estabelecido modelo de envelhecimento biológico proposto por T.J.P. Penna, seguindo um caminho que foi ao mesmo tempo científico e pedagógico: abordamos, em primeiro lugar, alguns problemas relacionados com as propriedades de base do modelo para depois estudar temas relacionados com a evolução de uma única população e finalmente chegar à descrição de fenômenos macroevolutivos, caracterizados pela interação de numerosas espécies diferentes.

Por isso o nosso projeto iniciou-se com o estudo das propriedades de escala do modelo clássico de Penna. Em seguida, através de uma extensão natural deste modelo que permite a representação de um caráter fenotípico individual, foi possível descrever os efeitos de microevolução em um sistema ecológico ameaçado por uma caça seletiva, fornecendo dados quantitativos em relação à sustentabilidade ecológica desse biosistema. Em seguida, por meio da definição de um perfil adaptativo específico no espaço de fenótipos, foi estudado o fenômeno da especiação simpátrica, ou seja as condições que permitem a auto-organização de aglomerados fenotípicos estáveis e isolados reprodutivamente a partir de uma população única de indivíduos que se reproduzem sexualmente. Por fim, a partir de uma microdinâmica governada pelos mesmos mecanismos, foi desenvolvido um novo modelo com o objetivo

de descrever as características estatísticas de um sistema macro-evolutivo, ou seja a dialética dos processos de geração e extinção das espécies em um dado ecossistema.

Abstract

This thesis aims to develop an interdisciplinary research, studying some problems of biology with the use of techniques inherited from theoretical and computational physics. Different topics were addressed, all related to the study of emergence of collective behaviour in microscopics agent based models, simulating evolutive biological systems.

We unfolded our research program, starting from the well established model for biological ageing proposed by T.J.P. Penna, following a scientific and pedagogic path. At first we addressed some questions regarding the basic properties of the model, then we studied some problems connected with the evolution of a single population. At the end, we succeed in describing macroevolutive phenomena, characterised by the interaction of many different species.

We started our project studying the scale properties of the classic Penna model. In the following, throughout a natural extension of this model allowing for the representation of an individual phenotypic character, we described the effects of microevolution in an ecological system threaten by a selective hunting. We were able to state some quantitative predictions related with the problem of ecological sustentability of such biosystem. Furthermore, we studied the phenomenon of sympatric speciation determining the conditions that allows the self-organisation of stable, reproductively isolated, phenotypic clusters starting from a sexual interbreeding population, through the implementation of a specific adaptive landscape in the phenotypes space. Finally, in a system governed by a similar microdynamics, we developed a new model with the aim of describing the statistical properties of a macro-

evolutive system, that is the dialect of generation and extinctions of species in a given ecosystem.

Capítulo 1

Introdução

J'ai essayé plus d'une fois, comme tous mes amis, de m'enfermer dans un système pour y prêcher à mon aise. Mais un système est une espèce de damnation...

Je suis revenu chercher un asile dans l'impeccable naïveté.

C. Boudelaire, *Exposition universelle*

Hoje em dia, consultando o arquivo de artigos eletrônicos *arXiv*, no campo relativo a Física, será fácil encontrar, ao lado de estudos sobre os condensados de Bose-Einstein ou as junções Josephson, artigos que se ocupam do mercado de ações, da morfologia dos conglomerados urbanos ou da periodicidade nas extinções de massa - isso sem efetuar uma pesquisa mais específica na seção recém aparecida, que sob o nome de *quantitative biology*, dedica-se inteiramente a estudos de problemas biológicos - .

Sem que nada de evidente ou explícito acontecesse, parece que na cabeça de alguns físicos ocorreu uma pequena revolução, concretizada numa migração repentina dos argumentos de estudo em direção a novas áreas, bem longe dos fenômenos típicos relacionados com as ciências duras. Tal processo vem sendo acompanhado por um interesse crescente numa abordagem multidis-

ciplinar aos problemas de pesquisa, por meio da convergência de técnicas e conceitos provenientes de diferentes áreas da ciência. Trata-se de um desenvolvimento relacionável a uma nova compreensão da complementaridade entre diversas áreas de pesquisa, na esperança que tais esforços inovadores possam levar a novos resultados.

Este tipo de abordagem resulta fortemente motivado pelo fato que os estudos predominantemente reducionistas, extremamente especializados e setoriais, se por um lado obtiveram sucessos espetaculares, e ainda continuarão a ter, por outro geraram uma enorme massa de dados e informações de difícil interpretação e que necessitam de um novo aparato conceitual para poderem ser reelaboradas e integradas. Estes fatos são sintomas de uma dificuldade que a abordagem puramente reducionista e a aridez de uma especialização extrema vem encontrando. A habilidade em reduzir todos os fenômenos a poucas e elegantes leis fundamentais não garante a capacidade, partindo destas leis, de reconstruir a inteira panóplia do universo. Ademais, este projeto construcionista manifesta todos seus limites no confronto com a complexidade e as diferenças de escala [23, 24]. Na mesma direção, a prática científica vem mostrando a incompatibilidade dos seus métodos com um reducionismo eliminativo que considerava algumas ciências dispensáveis, redutíveis às ciências primárias [22].

Este imprevisto e formidável crescimento de estudos e pesquisas de caráter interdisciplinar efetuado por parte de físicos têm suas raízes também em razões bem mais pragmáticas: os sucessos que nos últimos trinta anos a física obteve na área da mecânica estatística, transições de fases e dinâmica não linear de sistemas desordenados geraram um novo arsenal teórico e conceitual, baseado nas idéias de escala, leis de potências, universalidade e auto organização que se revelou ideal para ser transposto a problemas de outras áreas do conhecimento.

Ademais este êxito foi acompanhado pela explosão da potência computacional das máquinas de cálculo, que não só permitem a solução de complicados problemas teóricos ou numéricos mas conseguiram definir um novo paradigma, sendo consideradas capazes de gerar experimentos automáticos e de ser fonte de dados virtuais. O clássico processo de simplificação e controle que está na base da constituição de experimentos encontra nos computado-

res um aliado ideal na definição de sistemas extremamente simplificados e totalmente controlados dos fenômenos que se quer indagar. Por fim, o uso destas máquinas com técnicas típicas da mecânica estatística tem mostrado toda a sua eficácia na modelagem de sistemas caracterizados por grande complexidade.

Se um crescimento maciço no número de físicos interessados em problemas interdisciplinares, relacionados com o estudo dos comportamentos coletivos seja em sistemas econômicos, sociais ou biológicos, pode ser situado no começo dos anos noventa, trabalhos marcados por uma visão, para aqueles tempos, heterodoxa aparecem, ainda que de forma mais esporádica, já nos anos '40. Nesse período foram editados dois trabalhos, escritos por Ettore Majorana [1] e Erwin Shrödinger [2], o primeiro apresentando uma analogia entre leis estatísticas na física e nas ciências sociais, o segundo discutindo o estudo de problemas biológicos do ponto de vista de um físico.

Por fim, temos também que admitir que o estabelecimento a pleno título deste tipo de investigações como uma nova área de pesquisa ainda não foi acompanhada por algum resultado extraordinário e nem por uma abertura conceitual marcante mas, sem dúvida, os efeitos da discussão e reelaboração destas diferentes abordagens e temas de pesquisa e a dialética que vem estabelecendo-se entre áreas científicas diferentes são conquistas consideráveis.

O trabalho desta tese é endereçado ao estudo de problemas de biologia evolutiva desenvolvido por meio de métodos e técnicas provenientes da física estatística. Por esta razão, logo em seguida, começaremos com a apresentação da moderna teoria da evolução, a moldura fundamental no interior da qual se situam todos os problemas de biologia evolutiva, e mais adiante mostraremos alguns conceitos e estudos que os físicos desenvolveram na área da dinâmica evolutiva.

1.1 **Biologia evolutiva: a teoria da evolução**

A moderna teoria da evolução é o resultado da integração de um rico e vivo debate, que perdura há mais de um século, sobre o primeiro trabalho de

Charles Darwin, *The origin of the species*, que, como afirmado pelo mesmo autor, definiu em um original “one long argument” a primeira estrutura e a base lógica dessa teoria. Com a sua obra, Darwin apresentou provas biológicas e argumentos teóricos para validar a hipótese que todos os organismos existentes são descendentes de um ou, no máximo, alguns poucos ancestrais comuns (ver figura 1.1).

Para chegar à demonstração do seu teorema, Darwin elaborou suas argumentações a partir de algumas considerações fundamentais baseadas em uma série de fatos inquestionáveis. Esses fatos podem ser resumidos em três pontos: todos os organismos se multiplicam, variam e são capazes de herdar algumas características próprias. São estas propriedades aquelas que definem os elementos essenciais dos sistemas biológicos: serem caracterizados por um processo reprodutivo que garante a hereditariedade dos caracteres junto ao aparecimento de novas variantes, através da produção de um número superabundante de recém nascidos.

Ademais Darwin identifica, como força responsável pelas mudanças evolutivas, a seleção natural, um mecanismo que determina diferentes sucessos reprodutivos às diferentes variantes, a vantagem das mais adaptadas ao ambiente no qual vivem. Tal força, premiando com um maior sucesso reprodutivo os indivíduos mais bem adaptados, permite a acumulação de variações biológicas com efeitos positivos. Com essa determinação dos mecanismos evolutivos, Darwin não só afirma a necessidade das mudanças evolutivas mas também a existência da adaptação dos indivíduos ao meio ambiente.

Por meio desta análise, Darwin definiu as propriedades dos sistemas biológicos em relação à evolução, especificando aquilo que é uma dinâmica evolutiva, diferenciando-a de qualquer outra evolução temporal simples. Os agentes de um sistema evolutivo tem que ser capazes de hereditariedade, variação e de ciclos reprodutivos caracterizados por um crescimento logístico. Também este último detalhe é importante: permite a existência de uma rica bacia de novos indivíduos onde a seleção natural pode produzir resultados altamente improváveis selecionando mudanças raras que podem rapidamente invadir a população graças aos efeitos que uma taxa de reprodução maior tem numa dinâmica de crescimento logístico.

Na versão original da teoria Darwiniana a força de seleção natural re-

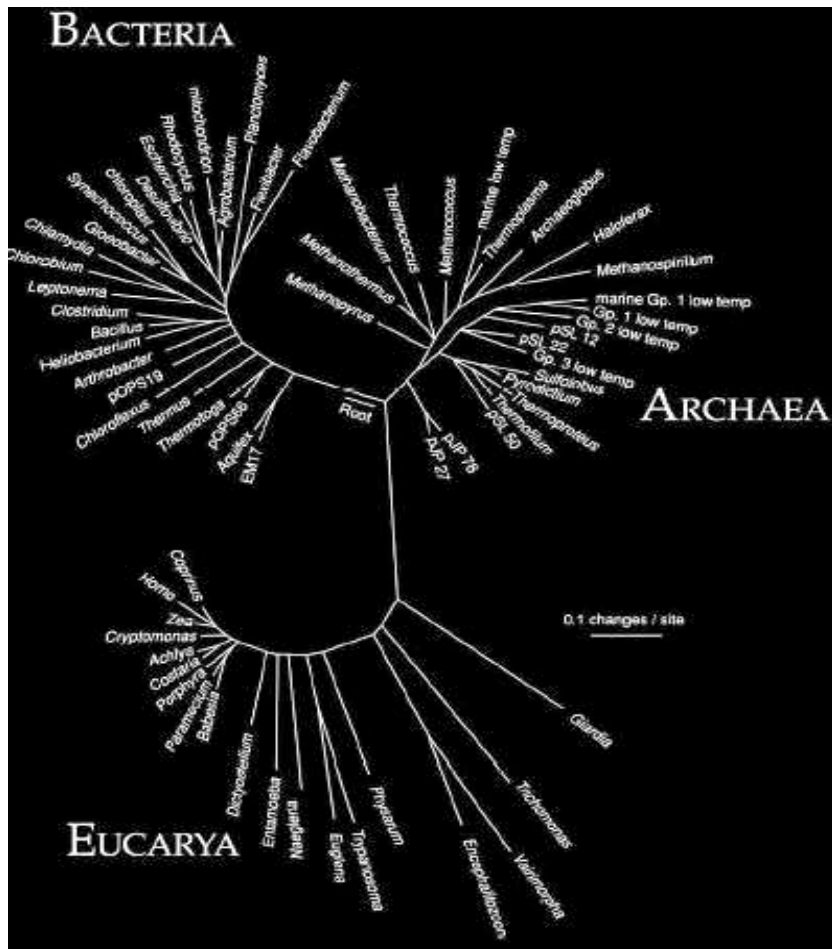


Figura 1.1: Árvore filogenética universal. São mostradas as relações entre Bacteria, Archaea e Eucarya.

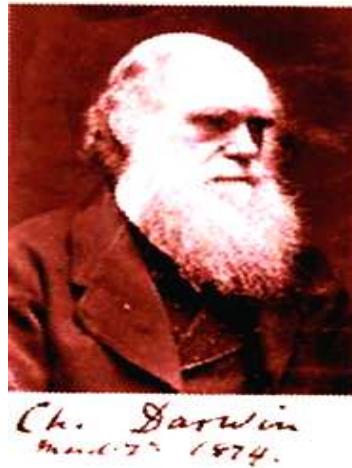


Figura 1.2: Charle Darwin, o pai da Teoria da Evolução

sulta caracterizada por três princípios [4]: ela opera a nível dos organismos como indivíduos, é o único mecanismo criativo gerador de adaptação por vias funcionalistas e sua atuação microscópica pode ser extrapolada a todas as escalas espaciais e temporais, definindo todas as propriedades macroscópicas da inteira riqueza de formas e diversidade da história da vida.

É interessante refletir um pouco sobre esta caracterização Darwiniana da seleção natural. Essa resulta fortemente ligada ao grande esforço reducionista e mecanicista com o qual Darwin quis criar uma teoria que, a partir de um nível microscópico, fosse capaz de explicar todas as variações macroscópicas. Uma grande influência neste percurso deve-se ao trabalho do economista Adam Smith, que viu as ordens macroscópicas da economia surgirem espontaneamente como consequência da dinâmica microscópica da luta individual para a vantagem pessoal. É neste quadro histórico que se torna possível compreender como foi natural para Darwin definir como único agente sobre o qual operava a seleção a entidade que naquele período era considerada como a mais básica, o organismo, e como de forma natural adaptação e funcionalismo foram considerados centrais na definição da força seletiva.

A obra de Darwin define a hereditariedade sem nada dizer dos seus mecanismos. Só no começo do século XX, com a redescoberta dos estudos

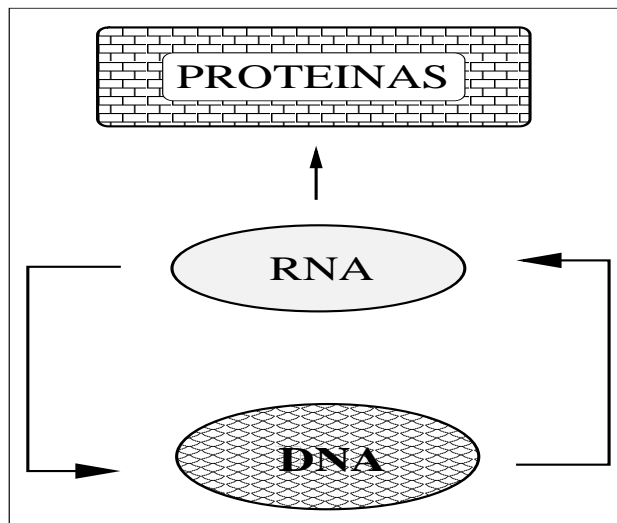


Figura 1.3: Uma simples representação do fluxo da informação genética segundo a visão do dogma central.

mendelianos e por meio dos primeiros passos da moderna teoria genética, foi possível identificar tais processos. Uma consequência destes novos estudos foi a definição de um novo paradigma pelo qual todas as mudanças evolutivas são recondutíveis às mudanças nos ácidos nucleicos dos organismos, os únicos meios de transmissão de informação de uma geração para outra. Estas novas idéias permitiram a constituição do quadro Neo Darwinista, baseado no dogma central [5], que afirma que a passagem de informações no sistema genético é permitida na direção $\text{DNA} \rightarrow \text{RNA} \rightarrow \text{proteínas}$ e, em alguns vírus, $\text{RNA} \rightarrow \text{DNA}$, mas nunca na direção das proteínas para o DNA, sepultando definitivamente qualquer idéia de caráter lamarkiano (ver figura 1.3).

O grande sucesso do casamento entre a teoria evolutiva de Darwin e a teoria genética, levará, nos anos 50, à definição de um novo paradigma na biologia evolutiva, chamado de Síntese Moderna, que tentará conjugar toda a fenomenologia das mudanças evolutivas de forma consistente com os mecanismos da moderna teoria genética. Como consequência, a teoria matemática da evolução se tornará dominada pela genética de populações [5], com explicações detalhadas da maioria dos mecanismos na escala molecular. Pela sua força preditiva, suas bases reducionistas e a sua coesão sedutora, esta teoria

se tornará um paradigma universal para a biologia.

De outro lado, são os resultados desses desenvolvimentos os primeiros a agitar uma nova discussão sobre os três princípios caracterizantes da seleção Darwiniana. Ao lado da idéia de uma força seletiva que opera só ao nível de organismo aparece a possibilidade de uma seleção que opera aos níveis de genes, células e até moléculas. O debate chega ao ponto de considerar possível que o gene seja a unidade de base sobre a qual opera a seleção.

Se com a Síntese Moderna a idéia que considera o organismo como a mônada fundamental da dinâmica evolutiva conhece uma crise profunda, é só a partir dos anos 70, com alguns trabalhos de Gould e Eldridge [80, 81], que vem sendo avançada a hipótese de um modelo hierárquico da seleção que passa a operar aos níveis, seja mais microscópicos dos organismos, como genes e células, seja a identificar entidades sujeitas a seleção em níveis mais macroscópicos, como demes (populações locais), espécies e clades (grupos de espécies que apresentam um ancestral comum), configurando desta forma uma crítica mais vasta a uma abordagem puramente reducionista.

Nessa ótica, em relação à segunda propriedade da evolução Darwiniana, vem configurando-se a hipótese que existe uma variedade de mecanismos operantes em nível evolutivo, e não unicamente a seleção. É esta uma tentativa de afastar-se das perspectivas unidirecionais que querem determinar a evolução em bases puramente adaptativas e segundo o modo funcionalista, ou seja, explicar cada órgão ou característica por meio da categoria da utilidade. Discute-se a necessidade de considerar os vínculos estruturais dos elementos evolutivos como origem de alguns fatos evolutivos. O surgimento de estruturas é produto não unicamente da adaptação causada pela seleção mas por algumas características dinâmicas ou estruturais que determinam a auto organização de novas estruturas, obtendo, segundo a expressão de Kauffman [7], “order for free”. Nas mesma direção, quer-se reavaliar a importância das interações geradas por mecanismos de realimentação entre organismos e ambiente ou entre as estruturas internas dos organismos e aquelas externas. Enfim, não pode ser esquecida a possibilidade de considerar evoluções não adaptativas, em particular depois da descoberta de como a variação da seqüência de alguns genes pode não ter efeitos sobre o valor adaptativo, descoberta que deu lugar ao surgimento da teoria neutral da evolução [6].

Concluindo, o estudo das grandes extinções de massa colocou em crise a idéia de poder-se extrapolar os resultados obtidos numa escala temporal a todas as outras, representando a inteira complexidade dos fenômenos evolutivos por meio da acumulação dos modos microscópicos.

O debate agora apresentado, em relação aos princípios da teoria, é profundamente relacionado com os problemas que uma abordagem puramente reducionista, não obstante todos os méritos e seus inegáveis sucessos, pode causar. A dificuldade em descrever uma seleção que opera contemporaneamente em diferentes níveis de uma estrutura hierarquizada, a presença de vínculos estruturais como causa de organização, a definição dos mecanismos que constituem estruturas complicadas em escalas diferente, é exatamente o problema que, em termos diferentes, é enfrentado na física dos sistemas complexos: dinâmicas fora do equilíbrio, não locais, não lineares, com mecanismos de realimentação, auto-organização, emergência e criticalidade.

1.2 Estudos de dinâmicas evolutivas

Tentaremos agora apresentar como novos métodos teóricos, conceitos interpretativos e algoritmos computacionais provenientes da física estatística, têm tido um interessante papel em abrir novos e significativos caminhos no estudo de problemas relacionados com a ciência da vida. Concentrando o nosso interesse no estudo da dinâmica dos processos evolutivos, a nossa atenção dirige-se aos trabalhos computacionais e ao uso de métodos estocásticos.

Um dos primeiros resultados fundamentais, derivado da aplicação de métodos da física estatística na descrição de fenômenos evolutivos, é o bem conhecido modelo de quasiespécies de Eigen [14]. Baseado nos empolgantes resultados do reducionismo da Síntese Moderna, modela a evolução como um sistema dinâmico estocástico e não linear, onde uma população de indivíduos, identificados pelo respectivo genótipo, difunde-se no espaço de todos os genótipos que, neste caso, representa seqüências de polinucleotídeos. A dinâmica do modelo segue as limitações impostas pela variação genética, pelo processo de desenvolvimento e pela seleção, com dependência do sucesso do fenótipo. Esta abordagem pioneira é baseada em modelos de cinética química. Não

obstante a sua importância, resulta incompleta, não conseguindo considerar as ações dos vínculos e a riqueza das escalas temporais. Se este trabalho pode ser considerado como paradigmático e fundamental, tanto que ainda muitos estudos de hoje se baseiam em idéias nele contidas, a direção que ultimamente estes estudos tomaram procura definir as condições para tentar realizar uma teoria integrativa das dinâmicas evolutivas, capaz de considerar os efeitos da seleção, os acidentes, a neutralidade, a relevância de estruturas e funções [10]. Com o fim de atingir este objetivo resulta relevante o uso não só de técnicas herdadas da mecânica estatística mas também de sistemas dinâmicos, como a teoria do caos, das bifurcações e da formação de padrões.

Do nosso lado, logo em seguida vamos apresentar alguns conceitos e métodos diretamente derivados da mecânica estatística, que resultam mais predominante e amplamente utilizados no panorama dos estudos de sistemas biológicos evolutivos.

1.2.1 Auto-organização e emergência

Os conceitos de auto-organização e emergência estão hoje em dia entre os mais usados e presentes na literatura das ciências não lineares sem, de outro lado, terem nunca sido definidos de forma sistemática e precisa. Estes conceitos são fortemente ligados ao problema da produção de entidades estruturadas e, por isso, conexos com o estudo da morfogênese. Obviamente o processo de adaptação, interpretável como criador de estruturas ordenadas, se encontra diretamente relacionado com estas idéias.

Com o termo auto-organização quer-se descrever uma propriedade bastante contra-intuitiva, ou seja a propriedade de sistemas que, deixados sozinhos, no lugar de aumentar o próprio estado de desordem e desorganização, autonomamente aumentam a própria estrutura. Em particular, o sistema aumenta o seu grau de organização no tempo, sem ser ordenado por agentes externos, ou seja não recebendo estímulos externos, como no caso de sistemas isolados, ou só estímulos não estruturados, como nos sistemas sujeitos a ruídos externos.

A dificuldade da definição da auto-organização está ligada ao conceito de

crescimento de organização ou crescimento de complexidade: a diminuição da entropia termodinâmica não é uma boa medida de organização em sistemas complexos, até o ponto que existem alguns sistemas biológicos onde a auto-organização é acompanhada por um aumento na entropia [15]. Por este motivo, uma das mais recentes definições afirma que um sistema se auto-organiza se, de forma autônoma, aumenta no tempo a sua complexidade estatística, como definido em [15]. Para uma introdução e uma tentativa de quantificação do conceito veja-se [15].

Podemos falar em propriedades emergentes quando alguma propriedade macroscópica de um sistema de muitos corpos não pode ser obtida através de uma operação estatística a partir de algum caráter puramente individual dos agentes microscópicos que definem o sistema. Por exemplo, se a magnetização pode ser obtida como uma simples média do valor de spin de cada elemento do sistema magnético em exame, a formação de aglomerados magnéticos não pode ser descrita e nem identificada através de uma operação deste tipo.

Um exemplo da relevância destes conceitos no trabalho aqui apresentado pode ser encontrado no segundo capítulo onde, no modelo Penna, a distribuição de genes mutados se auto-organiza concentrando-se na parte final dos genomas dos indivíduos, ou ainda no quarto e quinto capítulos onde o processo de especiação emerge através da auto-organização de distintos aglomerados fenotípicos.

1.2.2 Paisagens

A idéia de paisagem evolutiva (*adaptive landscape*) foi introduzida por S. Wright nos anos '30 como uma metáfora que representa os processos evolutivos. Esta imagem descreve estes processos como o movimento, através de passos sucessivos, dos indivíduos de uma população que escalam a superfície de um relevo à procura de um pico adaptativo, ou seja, como um processo de otimização na procura de um máximo local.

Foi possível enquadrar de forma mais quantitativa esta simples metáfora através da definição do conceito de paisagem adaptativa (paisagem adaptativa). O valor adaptativo (valor adaptativo) de um conjunto de indivíduos

de uma população, caracterizado por ter uma mesma característica fenotípica ou genotípica, é um número proporcional ao número médio da descendência gerada pelos indivíduos que possuem dita característica. Definido o valor adaptativo desta forma, é possível construir uma função adaptativa (*fitness function*) designando a cada ponto do espaço dos genótipos/fenótipos o respectivo valor adaptativo, ou seja da taxa de reprodução. Falaremos de paisagem adaptativa no caso que esta função seja também capaz de definir um gradiente que determina a dinâmica da população em exame.

Podemos então imaginar o processo evolutivo como a dinâmica de todos os indivíduos de uma população, que movem-se como pontos sobre a paisagem adaptativa.

É evidente a analogia que pode ser traçada entre a paisagem adaptativa e as funções de custo tanto utilizadas na física, como a energia livre na termodinâmica ou, de forma mais geral, as superfícies de energia. Agora, é em geral difícil poder identificar uma energia que determine um gradiente capaz de gerar toda a dinâmica do sistema. Uma exceção elegante e bem sucedida é o modelo apresentado no trabalho de revisão escrito por Luca Peliti [13]. Foi possível tornar muito forte a analogia entre paisagem adaptativa e superfície de energia relacionando este tipo de superfície com aquelas geradas por um Vidro de Spin, modelos caracterizados por um grande número de mínimos energéticos locais. Através desses estudos foi possível descrever a dinâmica de uma população assexuada que vive em um ambiente constante, definindo uma Hamiltoniana efetiva dada por:

$$\beta H = \beta \sum_{i,t} s_i(t) s'_i(t+1) + k \sum_t F(s'(t))$$

onde o vetor s_i representa o genótipo dos indivíduos, um vetor caracterizado pelos valores ± 1 , $\exp[F(s)]$ a função adaptativa e β é dependente da taxa de mutação dos indivíduos. Aparece evidente a analogia formal com um modelo de Vidro de Spin Quântico. Existe uma rica literatura onde foram estudadas várias paisagens adaptativas, ora caracterizados por superfícies lisas ou com poucos picos pontiagudos, ora por superfícies rugosas. O fato mais interessante dessa abordagem é a possibilidade de definir com clareza a transição do sistema de fases caracterizadas por estruturas de ordem bem definidas para outras diferentes (como a transição entre uma fase adaptativa e uma neutra

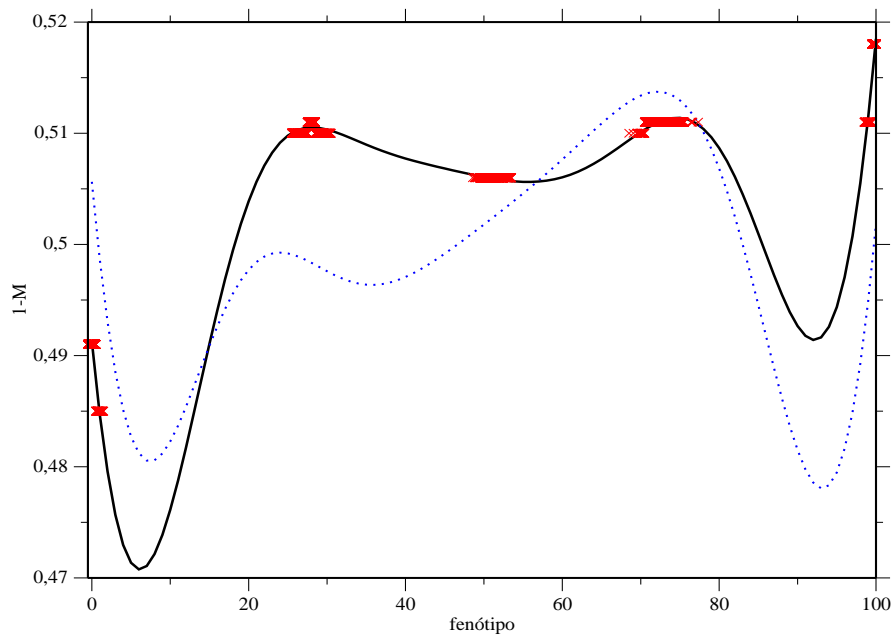


Figura 1.4: $1-M$, onde M é a probabilidade de morte do modelo apresentado no quarto capítulo. Esta função pode ser considerada proporcional a uma paisagem adaptativa (a taxa de reprodução é fixa, então o sucesso reprodutivo só depende do tempo de vida do indivíduo) ou a uma superfície de energia. As cruzes representam os indivíduos da população que movem-se em cima da paisagem adaptativa. A curva pontilhada indica a mesma função no intervalo temporal seguinte.

ao variar-se a taxa de mutação), em clara analogia com modelos da mecânica estatística.

É interessante notar como, sobretudo em sistemas mais complexos implementados com simulações de tipo Monte Carlo, a dinâmica do sistema, deixando de lado as complicações devidas à implementação dos detalhes do processo reprodutivo, é definida por uma seleção que é representada por um fator de morte. Neste caso, para uma taxa de reprodução fixa e igual para todos os indivíduos, a paisagem adaptativa é proporcional à função $1 - \mathcal{M}$, onde com \mathcal{M} indicamos a probabilidade de morte como implementada na simulação de Monte Carlo (ver figura 1.4).

Paisagens adaptativas dependentes do tempo, e que descrevem mecanismos de realimentação entre população e meio ambiente capazes de gerar uma dinâmica evolutiva fora do equilíbrio, são utilizados no estudo de fenômenos de ramificação e formação de aglomerados fenotípicos nos capítulos 3 e 4.

1.2.3 Comportamento de escala e criticalidade

A importância das idéias de escala é hoje em dia bem evidente. Esta abordagem resulta tanto mais promissora quanto mais é difícil lidar com a riqueza de detalhes do problema em exame. Situações típicas onde tais estudos têm tido grande sucesso são aquelas em sistemas onde os detalhes microscópicos não parecem ter uma grande importância nas escalas maiores e não resultam essenciais para obter o tipo de informação que está sendo pesquisada. Um grande número de diferentes relações de escalas é hoje em dia encontrado em fenômenos biológicos [11, 12, 16, 8]: entre a dimensão do corpo e a dos órgãos que o compõem, entre as taxas dos processos fisiológicos e os ciclos biológicos, nas árvores taxonômicas e na estatística dos dados relativos aos eventos de extinção. Esta freqüente presença de comportamentos de escala é de alguma forma relacionada com estruturas hierárquicas presentes e vem sendo usualmente descrita pela presença de leis de potência que representam diferentes distribuições estatísticas. Tais leis de potências assumem uma importância real se correspondem também a um genuíno comportamento de escala, ou seja se as leis de potências medidas estão relacionadas com uma invariância de escala na morfologia do espaço de definição do sistema.

Este tipo de abordagem finca suas raízes na moderna teoria dos fenômenos críticos, que permite relacionar a presença de leis de potência com propriedades geométricas do espaço relativo ao sistema físico em exame.

Tentando uma brevíssima apresentação (para uma introdução veja-se [17, 18]), podemos dizer que, para transições de fase contínuas, na vizinhança do ponto crítico se apresentam não analiticidades na energia do sistema, na forma de singularidades na derivada segunda ou superior. Em conseqüência, as funções de resposta mostram divergências que resultam descritas por leis de potência. As funções de resposta manifestam uma sensibilidade do sistema às perturbações externas e são ligadas com as flutuações que estas últimas provocam. Por este motivo, as divergências das funções de resposta podem ser relacionadas com uma correlação de longo alcance entre as flutuações. Tal correlação é descrita por um comprimento de correlação que, no ponto crítico, chega a ter tamanho infinito. Seguindo então uma hipótese dessa teoria, perto do ponto crítico a única escala de comprimento importante e natural do sistema é o comprimento de correlação. O fato de que este torna-se infinito no ponto crítico implica que, nas vizinhanças deste ponto o sistema é estatisticamente dominado por flutuações auto similares até a escala do comprimento de correlação.

Esta suposição pode ser formalizada afirmando que em tal ponto o sistema tem uma simetria de dilatação que gera um comportamento invariante na escala espacial, e em conseqüência em todas as funções dependentes de um comprimento, que decaem como potências da variável espacial. Dessa forma se define uma nova região, entre caos e ordem determinística, onde, graças à simetria de dilatação, é possível eliminar os graus de liberdade menores que o comprimento de correlação. É este o procedimento do grupo de renormalização que permite calcular os expoentes das leis de potência e justificar, através do seu sucesso preditivo, a suposição da hipótese de escala e a auto-similaridade. A partir desses resultados podemos concluir que a propriedade que define a criticalidade é o fato que a escala natural de comprimento da estrutura do sistema físico, ou seja o comprimento de correlação, torna-se infinita, gerando um comportamento espacial livre de escalas. É desta forma que são gerados comportamentos descritos por leis de potência.

Para conseguir encontrar uma relação entre esta teoria e problemas mais

propriamente ligados a fenômenos evolutivos é ainda necessário considerar um novo ingrediente, ou seja, o tempo, uma vez que temos que descrever fenômenos fora do equilíbrio.

Por fim, é interessante notar como existem alguns trabalhos [78, 19] que, embora num estágio preliminar e também não sem algumas dificuldade, tentam estudar o problema da especiação em analogia com uma transição de fase, como já o trabalho de Peliti fez na descrição de dinâmicas evolutivas e como será discutido no terceiro capítulo.

1.2.4 Sistemas dinâmicos: dinâmica de limiar e criticidade

Como exposto acima, o estado crítico de sistemas termodinâmicos é bem entendido e descrito em seus detalhes a partir da energia livre do sistema e com o uso do Grupo de Renormalização de Wilson. Ao contrário, não se sabe como descrever com um formalismo equivalente sistemas dinâmicos. Não obstante esta dificuldade, o ponto crítico, que em sistemas no equilíbrio é alcançado através de uma variação precisa no valor do parâmetro de controle, em alguns sistemas fora do equilíbrio é um atrator da dinâmica, robusto em relação aos parâmetros do sistema e independente das condições iniciais do mesmo. Este fenômeno é conhecido como Criticalidade Auto Organizada (SOC) [20]: o sistema se organiza em um estado com as propriedades características dos sistemas de equilíbrio que estão em um ponto crítico, apresentando auto-similaridade e leis de potências nas distribuições estatísticas, seja no espaço seja no tempo, ambas caracterizados por ausência de escalas.

Uma propriedade comum, característica destes sistemas é o fato de serem submetidos a um estresse externo contínuo, ao qual respondem com eventos discretos. Ou seja, a força aplicada cresce até superar um limiar e só neste ponto a energia conservada é relaxada repentinamente. Esta dinâmica de limiar resulta então caracterizada por uma separação espacial nas escalas temporais entre o estresse externo, lento e contínuo, e a relaxação interna, rápida e pontuada, ou seja, caracterizada por longos intervalos de *stasis* (sem atividade) e imprevistas eclosões de atividade. São estas as propriedades

que definem uma dinâmica pontuada, onde são gerados numerosos eventos caracterizados por distribuições sem escala característica, onde um pequeno aumento na intensidade do estresse pode causar respostas de qualquer tamanho.

O interesse por esta teoria no estudo de fenômenos biológicos é conexo com os famosos resultados obtidos por meio do modelo de Bak-Sneppen [83], o primeiro modelo capaz de sugerir como os fenômenos evolutivos podem ser caracterizados por um comportamento de equilíbrio pontuado e que lançou a hipótese de que fossem organizados em um estado crítico [12, 90]. No último capítulo do nosso trabalho apresentaremos um novo modelo para o estudo da dinâmica de eventos de originação e extinção, que pretende continuar uma análise das dinâmicas evolutivas no quadro traçado por este tipo de interpretação.

1.2.5 Computação booleana

Uma palavra enfim sobre umas técnicas de programação amplamente usadas na implementação de programas para a descrição de sistemas nos quais os agentes podem ser representados por um vetor de variáveis booleanas [21, 9], como acontece na maioria dos modelos por nós estudados.

No caso da escolha dos valores 0/1, estes vetores são simplesmente associáveis à representação binária das palavras do computador. Por esta razão, podem ser armazenados de forma muito simples e compacta, não necessitando muito espaço de memória, e podem ser manipulados bit a bit através de operações lógicas (AND, OR e XOR), muito mais rápidas e eficientes que qualquer operação algébrica. Graças ao uso dessas técnicas baseadas em tiras de bits é possível estudar simulações de grande tamanho e evoluções temporais longas. Isso permite obter dados comparáveis com aqueles encontrados na Natureza, como mostrado no capítulo 2, ou resolver controvérsias teóricas que, usando outras técnicas, a princípio não parecem abordáveis por meios computacionais, como fica evidente no capítulo 3.

Capítulo 2

Propriedades de escala em um modelo de envelhecimento

mundus senescit

Gregorio di Tours

2.1 O fenômeno do envelhecimento e sua explicação em bases evolucionárias

Com o termo envelhecimento consideramos todas as mudanças que ocorrem na última parte da vida de um organismo e causam uma redução na probabilidade de sobrevivência e/ou um declínio nas capacidades reprodutivas [26]. Na busca das causas que podem ser consideradas responsáveis por esse processo é natural dividi-las entre fatores puramente extrínsecos, determinados pelas influências do meio ambiente que agem danificando em vários níveis os indivíduos (do DNA até órgãos inteiros), e fatores intrínsecos ligados a elementos puramente genéticos. Se os primeiros fatores são inevitáveis e puramente casuais, tanto que não necessitam nenhuma explicação evolucionária, a existência e a natureza das causas genéticas precisam de uma abordagem diferente. O fato que mais sugere e desafia uma explicação deste fenômeno em bases evolutivas é a grande variedade na longevidade de orga-

nismos sujeitos aos mesmos riscos de natureza ambiental, ou seja, sujeitos à mesma exposição às influências negativas do meio ambiente.

Na tentativa de definir o cenário que determina este processo evolutivo é possível identificar duas forças principais.

Pela simples consideração que um gene que adia o período de senescência e/ou aumenta a capacidade reprodutiva gera uma vantagem evolutiva, podemos concluir que a seleção natural premia indivíduos portadores de tais genes. Então, evolutivamente, esta seleção teria que causar uma tendência uniforme a aumentar o período fértil e, mais em geral, a longevidade dos vários organismos. Ao contrário, para explicar as diferenças nas longevidades das várias espécies é necessário achar uma força que se oponha a este processo. Essa força tem que ser identificada com todos os fatores aleatórios de morte que agem na vida de um indivíduo, tais como acidentes ou os mesmos fatores extrínsecos de envelhecimento, e que então diminuem a força de seleção natural antes apresentada [28]. Ou seja, o envelhecimento é gerado por um equilíbrio entre uma força seletiva que tende de forma monótona a aumentar a longevidade dos organismos e uma força estocástica, com efeitos cumulativos no tempo, que, matando os indivíduos antes deles experimentarem a força seletiva, diminui definitivamente os efeitos de tal seleção. Também é crucial notar como essa força, embora representando fatores extrínsecos e embora sendo de natureza puramente estocástica, tem em realidade um efeito direto sobre os genomas, sendo a responsável pela diminuição dos efeitos de seleção sobre genes defeituosos ou pela fixação de genes pleiotrópicos (com efeitos favoráveis em um período da vida e negativos em outros). É possível encontrar uma confirmação destas teorias evolutivas por meio de dados biológicos. O fato de que as hipóteses evolutivas fazem uma ligação entre as expectativas de vida e a fertilidade em idades maiores está presente em vários organismos para os quais a fertilidade aumenta com a idade, como, por exemplo, em vários peixes e plantas. Mais, é possível encontrar provas empíricas da tese que a longevidade aumenta em populações sujeitas a um alto grau de seleção natural, e então, em correspondência, um baixo risco de mortalidade de natureza aleatória. Esse fato é comprovado por dados provenientes de animais criados em cativeiro e em condições ambientais ótimas. As espécies com maior expectativa de vida são aquelas que na natureza têm um

baixo risco de mortalidade. Por exemplo, os pássaros têm uma longevidade maior que os mamíferos de tamanho comparável e, entre estes últimos, os morcegos têm uma expectativa de vida maior. Isso porque a capacidade de voar diminui as chances de ser predado. Da mesma forma, também cágados e tartarugas vivem mais que outros répteis e, entre os pássaros, os que não voam são conhecidos por ter uma vida mais breve [27].

Hoje em dia, as teorias evolutivas do fenômeno do envelhecimento que mais conseguiram sucesso em explicar os dados biológicos são duas. Uma primeira, nascida com alguns trabalhos de Medawar [29, 30], é baseada na idéia que os genes relacionados com a fertilidade e que condicionam as chances de sobrevivência são dependentes da idade do indivíduo, ou seja afetam o indivíduo só a partir de uma idade específica. Isso significa que os genes que mostram os seus efeitos na parte final da vida estão sujeito a uma menor pressão seletiva por causa do fato que boa parte dos indivíduos que os carregam já morreram ou se tornaram inférteis. Nessa hipótese o envelhecimento é então gerado por uma acumulação de mutações desfavoráveis e ativas na última parte do ciclo vital (explicação mutacional [31]). Tais mutações são geradas com uma frequência suficientemente alta para balancear o equilíbrio entre seleção e mutações em favor dessas últimas. Essa causa de envelhecimento é devida a uma impossibilidade de se alcançar um estado ótimo, causa uma seleção não suficientemente forte que permita uma evolução puramente adaptativa.

Uma outra explicação é baseada na hipótese [32, 33] de que as possíveis combinações das possibilidades reprodutivas e de sobrevivência em diferentes idades não são totalmente livres mas têm que obedecer a alguns vínculos. Partindo dessas condições de contorno definidas pelas estruturas do material biológico (genético, fisiológico e ambiental), o processo evolutivo tende a maximizar o valor adaptativo das diferentes espécies e este resultado é obtido aumentando as chances no primeiro período de vida, fato que implica uma diminuição no período tardio (explicação baseada na optimalidade). Por essa explicação o envelhecimento é gerado por efeitos deletérios que aparecem tarde na vida como consequência indirecta de processos favoráveis na primeira parte (por exemplo, o custo que os indivíduos que têm a chance de se reproduzir pagam). É interessante notar que, embora essa teoria pode ser

estruturada de um ponto de vista puramente genético em termos de genes pleiotrópicos (que afetam o fenótipo em mais que um aspecto - nesse caso na sobrevivência e fertilidade em diferentes idades -), estes argumentos são independentes dos detalhes genéticos [34].

É muito difícil determinar por meio de experimentos biológicos qual das duas explicações é a mais plausível. Uma prática de estudo consolidada faz uso de exemplares de *Drosophila* criados artificialmente. Um experimento baseado em uma idéia de Edney e Gill [28], restringe a possibilidade de se reproduzir só a adultos jovens. Por esse motivo, a parte final do tempo de vida desses animais não resulta exposta a seleção natural. Como isso seja causa do envelhecimento pode ser estabelecido medindo a fecundidade (no lugar da longevidade): efetivamente aparece uma diminuição apenas na fecundidade tardia dos insetos, ou seja aquela relativa ao período onde a ausência de seleção causa a acumulação de mutações deletérias. Estes fatos sugerem que o envelhecimento evolui pela acumulação de alelos deletérios [35]. Não obstante este resultado seja favorável à teoria da acumulação de mutações deletérias, considerando outros dados, sempre provenientes de experimentos que fazem uso de *Drosophilas*, é difícil estabelecer uma primazia entre as duas teorias e parece que o envelhecimento seja em parte gerado pela seleção para obter um ciclo vital ótimo, em parte resultado da acumulação de mutações deletérias [34].

Embora não seja ainda possível definir com exatidão e precisão uma teoria do envelhecimento, através das hipóteses precedentemente apresentadas, é possível compreender como não se pode responsabilizar pelo envelhecimento só poucos genes ou uma única causa geral, mas, ao contrário, o envelhecimento apresenta uma natureza poligenética e fortemente dependente de várias razões fisiológicas e ecológicas. Por este motivo, também a possibilidade de se aumentar intrinsecamente a longevidade humana não pode ser pensada em termos puramente individuais. Ao contrário, um aumento na idade média de reprodução e uma diminuição dos fatores extrínsecos de mortes e infertilidade na população poderão causar, só depois de muitas gerações, uma diminuição, de natureza evolucionária, das taxas de envelhecimento nos seres humanos [34].

2.2 Modelos teóricos e computacionais no estudo do envelhecimento

O número de teorias relacionadas com o fenômeno do envelhecimento é tão grande que um trabalho de revisão do gerontólogo Medved [36] chegou a classificar até 300, incluindo como primeiras teorias, datadas nas cercanias dos anos 50, os trabalhos do Medawar [29] e Williams [27]. Em particular, entre os primeiros trabalhos baseados em um estudo matemático podemos citar o modelo de Partridge-Barton publicado em 1993 [34], que assume só dois intervalos de idade: juventude e maturidade. Só em seguida começaram a se desenvolver também simulações baseadas em técnicas de Monte Carlo [26, 31, 9], dentre as quais o modelo Penna [25], que é hoje em dia a técnica mais comum de simulação para este gênero de problemas. O modelo é baseado na explicação mutacional, e então na acumulação de mutações deletérias. O problema e os resultados relativos à relação entre o modelo Penna e as teorias evolutivas acima apresentadas e o sucesso do modelo na explicação e previsão de resultados experimentais podem ser encontrados em vários trabalhos de revisão (ver as referências [9, 37, 38]).

Segue uma rápida apresentação do modelo com um enfoque dirigido à introdução do problema específico objeto do nosso estudo: as suas propriedades de escala.

2.3 O modelo Penna para populações assexuadas e haplóides

Na sua forma mais simples, aquela que descreve populações assexuadas e haplóides (com cromossomas que se apresentam sozinhos), o Modelo Penna representa cada agente por uma tira de 32 bits, ou seja uma palavra de computador [25]. Tal tira codifica apenas genes dependentes da idade do indivíduo. A prática de codificar o material genético por meio de uma sucessão de valores booleanos (0/1, ± 1) é muito comum em vários modelos presentes na literatura [37]. No nosso caso, particularmente bem sucedida é a técnica de representar estes valores diretamente através da base binária e a inter-

pretação de tais valores em relação a uma leitura do genoma em sucessão temporal, como ficará claro mais adiante. Para uma dada tira o valor 1 codifica um alelo deletério, o valor 0 um alelo não deletério (alelo *wild-type*). A cada passo temporal um loco do genoma do indivíduo é lido e quando o número de alelos deletérios contados nesse processo supera um valor limite T o indivíduo morre. Depois dessa seleção e no mesmo passo temporal, os indivíduos (que são todos considerados fêmeas) que já atingiram uma idade maior que a idade mínima de reprodução (IMR) geram F filhos. Tal geração é obtida por intermédio de uma cópia da tira da mãe onde é introduzido um número M de mutações em diferentes locos, escolhidos aleatoriamente. Todas as mutações deixam o novo loco com um alelo do tipo 1, ou seja, são efetuadas só mutações deletérias e não é considerada a possibilidade de recuperar um gene mutado. Tudo isso é comparável com os processos biológicos conhecidos, onde as mutações desfavoráveis são duas ordens de grandeza mais prováveis que as favoráveis. Outras causas de morte também são incluídas no modelo. Em primeiro lugar, cada indivíduo tem uma idade máxima que coincide com o tamanho da tira que o representa - 32 nesse caso -. Além disso é também necessário introduzir um outro fator de regulação do tamanho da população para evitar a sua explosão. Este fator, chamado de fator de Verhulst, é uma probabilidade de morte do tipo campo médio dependente do tempo e da densidade da população e representa os limites de espaço e de nutrientes ao qual uma população real é sujeita. Esta probabilidade é dada por $V(t) = N(t)/K$, onde $N(t)$ é o tamanho da população no tempo t e K é o parâmetro que quantifica estas limitações ambientais. Um código Fortran desse modelo, que será usado como base para todas as simulações seguintes, pode ser encontrado na referencia [9] ou no Apêndice.

2.3.1 Comentários em relação ao fator de Verhulst, ao tamanho da tira de bits e à idade mínima de reprodução

O uso do fator de Verhulst é muito comum, a causa da necessidade de lidar com populações finitas. Se do lado biológico o fato de considerar-se a presença de mortes por causas não genéticas pode parecer realista, por outro

lado, sobretudo no caso dos mais complexos organismos diplóides, na natureza as mortes puramente casuais são dificilmente justificáveis. Mais, pelo fato que diferentes estratégias de regulação têm um impacto inesperado sobre a distribuição genética final das populações, a escolha da mais oportuna técnica de implementação computacional põe uma questão de difícil solução. A sua forma mais comum, como descrita acima, que depende da densidade da população e opera a cada passo temporal, é herdada de modelos precedentes para crescimento logístico onde gera dinâmicas bem estabelecidas, que são comparáveis com dados encontrados em populações reais [26, 44]. Uma alternativa na implementação desse fator de mortes aleatórias prevê um fator de Verhulst do mesmo tipo mas que age só no primeiro passo da vida de um indivíduo [45]. Esse uso do fator de Verhulst é equivalente a uma probabilidade de reprodução dependente do tamanho da população, ou seja a deixar uma fêmea reproduzir com uma probabilidade igual a $1 - V(t)$. Essa escolha é de particular interesse porque permite controlar a população de uma forma independente do tamanho do genoma, pelo fato que um indivíduo nunca é sujeito aos efeitos de um fator de morte externo. Ao contrário, na versão do modelo com a estratégia de Verhulst usual, cada indivíduo tem uma probabilidade de morte a cada passo temporal e a efetiva probabilidade de morte não genética é trivialmente crescente com o tamanho do genoma.

O comprimento padrão da tira de bits usado no modelo Penna é de 32 bits, pela facilidade de implementação em processadores com palavras de tal tamanho. Uma comparação desse valor com o tamanho natural da maioria dos seres vivos (que varia entre 10^4 e 10^5 unidades realmente funcionais) faz parecer o modelo profundamente pouco realista. Na realidade, como primeiro fato tem que se observar que só um pequeno número dessas unidades pode ser relacionada com enfermidades graves dependentes da idade. Mais, apesar de indicar com o termo genoma a tira de 32 bits, essa representa diretamente as conseqüências de mutações deletérias ou de erros contidos na informação genética, ou seja as doenças.

Deixando de lado essa última consideração e continuando a referir-nos à tira como genoma, em um estudo sobre os dados de mortalidade da população alemã, o genoma utilizado continha 128 bits [41]. Como resultado complementar, esse estudo sugeriu que não existia uma grande diferença qua-

litativa entre 32 e 128 bits e que era possível comparar dados para genomas de diferentes tamanhos escalando alguns parâmetros. Estes resultados, junto ao efetivo interesse em se afirmar que o modelo não é dependente do tamanho do genoma, são as principais motivações que nos levaram a abordar, em primeiro lugar, as propriedades de escala do modelo.

É interessante notar como o parâmetro de idade mínima de reprodução é de grande importância, sendo o principal regulador da distribuição da idade na população. Isso ocorre porque o começo do envelhecimento normalmente coincide com o valor de IMR , em total acordo com as teorias evolutivas descritas no começo do capítulo e comprovado de forma espetacular nos fenômenos da senescência catastrófica observada nas simulações de espécies semélparas (que reproduzem uma única vez, todos os indivíduos na mesma idade) [40]. Por esse motivo, em geral, também a expectativa de vida é fortemente dependente desse parâmetro. Estes resultados, considerados usualmente uma consequência trivial da teoria de acumulação das mutações, como será mostrado na frente, não são totalmente transponíveis para o caso de uma probabilidade de reprodução dependente do tamanho da população.

Por fim, é importante lembrar que o valor da idade mínima de reprodução tem que ser maior que o valor de T (valor limite de alelos deletérios ativados); em caso contrário, os indivíduos não estariam sujeitos à seleção natural responsável por premiar os genomas mais livres de alelos defeituosos.

2.4 Procurando um comportamento de escala no modelo Penna sexuado

Os interesses em esclarecer um possível comportamento de escala do modelo Penna são múltiplos. Do ponto de vista biológico, como já afirmado, seria importante achar um modelo independente do tamanho do genoma. Isso, entre outros fatos, tornaria redundante a questão relativa ao efetivo número de genes relacionados com o fenômeno de envelhecimento. Mais, considerando a estrutura de idade do genoma, de um ponto de vista físico, possíveis propriedades de escala no genoma significariam propriedade de escala temporal do modelo, que resultaria independente da escolha do passo temporal de Monte

Carlo.

Um primeiro estudo em relação a esse problema [42] não obteve resultados conclusivos. Nesse caso a simulação realizada usava um modelo assexuado e com um fator de Verhulst clássico, como descrito na seção 2.3.1, e tentava comparar diretamente resultados com diferentes parâmetros reescalados. Uma análise similar [43], mas para um modelo contínuo no tempo e com um genótipo representado por valores reais, obteve um resultado não facilmente mapeável na versão discreta do modelo. Considerando estes resultados precedentes, decidimos concentrar nossa atenção sobre o modelo Penna para indivíduos diplóides (cujos cromossomas se organizam em pares de cromossomas semelhantes) e que usam sexo para se reproduzir. Isso se justifica por ser esse o modelo mais utilizado e completo e por estarmos mais interessados em obter uma análise qualitativa mas que possa sugerir as propriedades de escala de um modelo considerado bastante realista, do que em obter resultados exatos com modelos altamente simplificados e pouco aptos para obter resultados confrontáveis com dados biológicos. Por esse motivo temos que introduzir as diferenças fundamentais entre o modelo assexuado e o modelo diplóide e sexuado.

A representação de uma população diplóide é feita introduzindo-se uma segunda tira de bits no genoma, que então fica composto por duas tiras de bits acopladas e lidas simultaneamente: em cada passo temporal são considerados dois alelos, provenientes um de uma tira e o outro da segunda. Dessa forma aparece no modelo o novo conceito biológico da dominância entre alelos. No nosso caso o gene considerado é deletério se os dois alelos têm o caráter deletério (loco homocigoto). No caso heterocigoto, normalmente, o gene expresso é o *wild type*, exceto no caso em que naquele loco o alelo 1 seja definido como dominante. Ou seja, em geral o alelo mutado é recessivo, com exceção de H locos, fixados no começo da simulação, onde tal alelo é considerado dominante. A implementação da reprodução tenta representar a principal fenomenologia do processo meiótico. O primeiro passo compreende a realização do cruzamento (*crossing-over*) do genoma da mãe. Por isso as duas tiras são cortadas em uma posição aleatória e os quatro pedaços resultantes são recombinados de forma que o pedaço esquerdo da primeira tira é adicionado ao pedaço direito da segunda e analogamente com os outros dois,

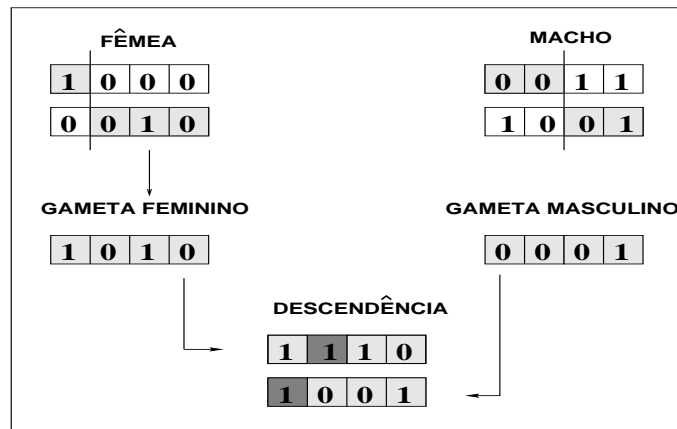


Figura 2.1: Representação do ciclo reprodutivo. O genoma diplóide atravessa o processo de cruzamento e recombinação e na segunda passagem é escolhido o gameta haplóide. O ciclo termina com a introdução das mutações (alelos com uma sombra mais escura) e a combinação dos gametas que leva a formação do novo indivíduo.

como ilustrado na figura 2.1. Das duas novas tiras é selecionada uma que torna-se o gameta feminino. Da mesma forma é produzido o gameta masculino de um macho aleatoriamente escolhido no interior da população. Da união dos dois gametas, depois que os mesmos tenham sofrido M mutações em cada uma das duas tiras, é gerado o novo indivíduo, que tem seu gênero escolhido também de forma aleatória. Com isso termina o processo reprodutivo, que então compreende cruzamento, recombinação e mutação. É possível encontrar algumas reflexões e críticas do ponto de vista de um biólogo, em relação ao modelo aqui apresentado, na referência [39].

2.4.1 Resultados

A análise aqui apresentada será focada nas propriedades de escala da distribuição de idade da população. As distribuições estudadas são normalizadas, não considerando como um dado essencial o fato que o tamanho da população cresce com o aumento da dimensão do genoma. Antes de apresentar estes resultados é interessante mostrar como efetivamente a escolha do fator de Verhulst é decisiva na nossa análise. Na figura 2.2 é possível apreciar como,

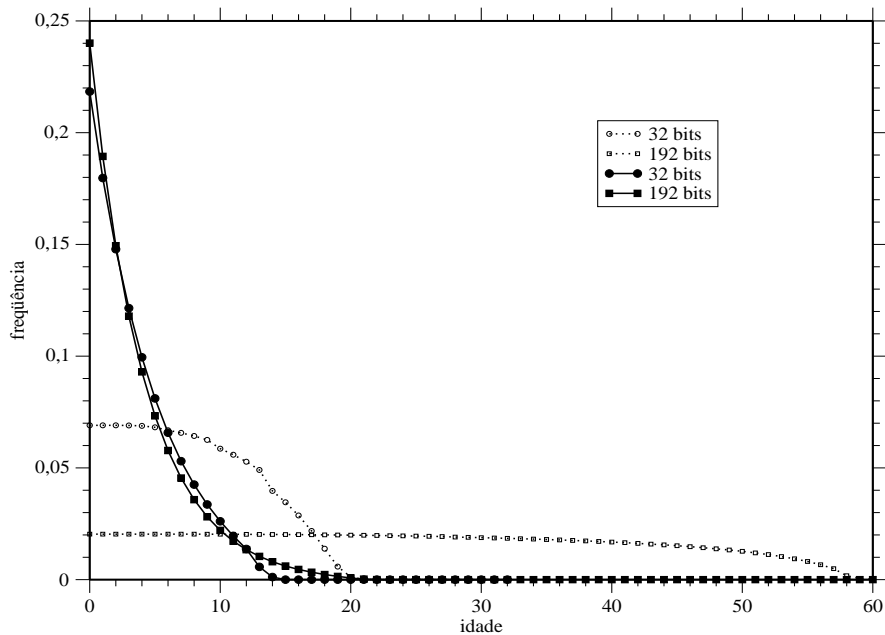


Figura 2.2: O traço contínuo representa as distribuições de idades geradas com um modelo que adota um fator de Verhulst ativo a cada passo temporal. O traço descontínuo representa as distribuições de idade geradas com um modelo que adota uma probabilidade de reprodução dependente do tamanho da população. A idade do começo da reprodução é fixada em 8 em todas essas simulações.

no caso em que se aplica um fator de morte aleatória a cada passo temporal, é difícil estabelecer uma lei de escala. As distribuições são fortemente afetadas no tempo por esse fator e tendem a se sobrepôr. Com uma probabilidade de reprodução fixa, o começo do envelhecimento coincide com a idade mínima de reprodução, e isso é um outro fator decisivo na impossibilidade de se achar uma forma de lei de escala simples e clara.

Os resultados relativos a nossa análise, com a probabilidade de reprodução dependente do tamanho da população, são apresentados na Fig. 2.3. Ao contrário da figura precedente, o começo do envelhecimento é retardado nos genomas maiores. A distribuição de idade, mesmo mostrando uma tendência a decrescer exatamente no começo da idade reprodutiva, o faz com derivadas muito pequenas, fato que pode ser bem apreciado na análise da função de mortalidade (ver figura 2.5) que apresenta um platô para idade pequenas. Em

geral, nesse caso, a expectativa de vida cresce linearmente com o tamanho do genoma, como ficará evidente adiante, e não é fortemente dependente do parâmetro que controla o começo da idade reprodutiva.

As distribuições de idade mostradas na figura são relativas a várias simulações do modelo, cada uma com o número de bits do genoma diferente. O primeiro usa uma tira de 32 bits. Sendo esse valor uma unidade natural nesse modelo computacional rodado com processadores com palavras de 32 bits, os genomas das outras simulações são escolhido com um tamanho múltiplo desse valor: 64, 96, 192 e 224 bits. Para todas essas simulações, os parâmetros que controlam o número de locos dominantes, a idade do começo da reprodução, o número de filhos e mutações em cada geração e o valor do limite de mutações deletérias são exatamente os mesmos. Também nesse estudo não consideramos os números de cruzamentos no processo de meiose como um parâmetro variável, mas ao contrário é fixo e igual a um. A variação desse parâmetro tem um interessante impacto sobre o comportamento das simulações, como estudado em um recente trabalho do qual voltaremos a falar [46]. O término da idade reprodutiva difere para genomas diferentes e corresponde sempre à idade máxima dos indivíduos (32, 64, 96, 192 ou 224). Essa escolha é sugerida pelo fato de que, com a nossa escolha de parâmetros, as simulações geram populações que se auto-sustentam e que, no equilíbrio, não são muito sensíveis à mudança desse parâmetro. Desta forma, a nossa escolha não introduz uma assimetria indesejável. Da figura é possível observar como as distribuições, ainda que qualitativamente parecidas, mostram diferenças bem apreciáveis. Esse fato nos sugeriu procurar uma lei de escala que permitisse achar uma relação entre duas idades diferentes (t_1 e t_2) pelas quais a integral das duas distribuições correspondentes ao diferentes tamanho do genoma ($\rho_1(t)$ e $\rho_2(t)$) chegasse a ter o mesmo valor. Formalmente podemos expressar esse conceito afirmando procurar uma reescala temporal $t_2 = F(t_1)$ que resolva a equação:

$$\int_0^{t_1} dx \rho_1(x) = \int_0^{t_2} dx \rho_2(x) \quad (2.1)$$

A solução resulta ser uma relação linear simples. Em todos os casos apresentados a integral da distribuição de idades, como função do seu limite superior, começa apresentando um crescimento linear, nas idades onde

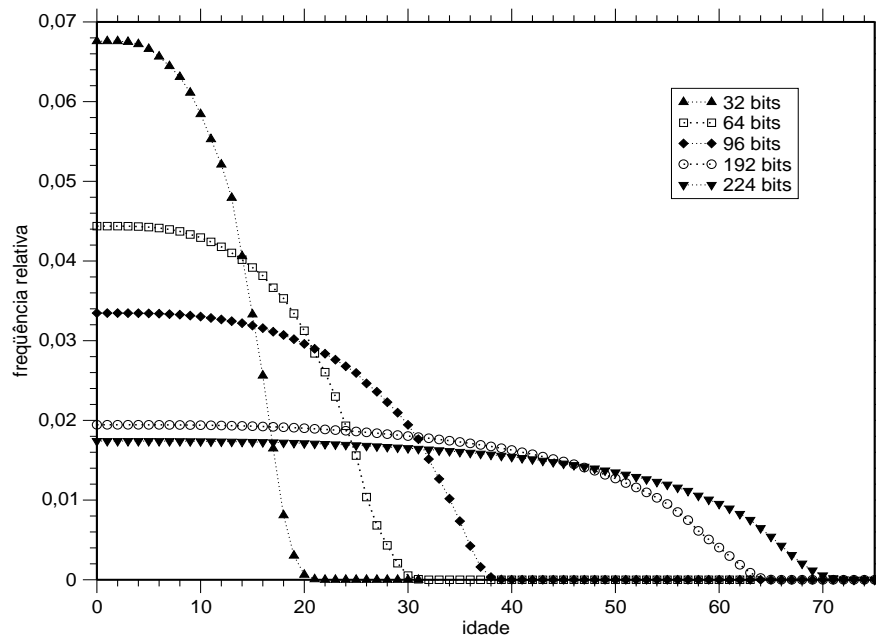


Figura 2.3: Distribuição de idades da população no modelo com 32, 64, 96, 192 e 224 bits. Este modelo tem probabilidade de reprodução dependente do tamanho da população e controlada pelo parâmetro de Verhulst. Os parâmetros usados nas simulações são: parâmetro de Verhulst (400000), população inicial (1000), idade mínima de reprodução (8), número de filhos a cada estação reprodutiva (4), máximo valor permitido de doenças ativas (3), número de mutações adquiridas no nascimento em cada tira de bits (1) e número de locos dominantes (6). Em cada medida apresentada efetuamos uma média sobre os últimos 1000 passos temporais de 10 diferentes realizações, a partir do momento em que todas as distribuições podem ser consideradas estacionárias (as simulações terminaram entre o passo 50000 e 200000, dependendo do tamanho do genoma).

j	2	3	6	7
c	1.53	2.02	3.45	3.83
d	0.51	1.06	2.58	3.12

Tabela 2.1: Os coeficientes c e d por $i = 1$ e alguns valores de j

a distribuição é essencialmente constante, e termina saturando. Este comportamento sugere que uma relação linear entre os valores temporais pode satisfazer à igualdade integral. Se $y_i = a(i) + b(i)t_i$ é a regressão da parte linear da função integral da distribuição $\rho_i(t)$, $y_i = y_j$ leva à relação que estamos procurando:

$$t_j = (b(i)/b(j)) \cdot t_i + \left(\frac{a(i) - a(j)}{b(j)} \right) = c(i, j) \cdot t_i + d(i, j) \quad (2.2)$$

Cada índice, i ou j , está definido como o tamanho da tira de bits dividido por 32. É possível obter os coeficientes dessa relação de escala fazendo a regressão de cada função integral e utilizando a fórmula derivada acima para calcular $c(i, j)$ e $d(i, j)$. Por simplicidade na notação utilizada omitimos o primeiro índice caso seja igual a 1. Na tabela 2.1 são apresentados os resultados relativos a estes coeficientes para alguns valores de j e para $i = 1$.

Essas simples relações de transformação permitem reescalar a função integral em sua totalidade, e não unicamente na parte linear. Também as distribuições de idade resultam reescaláveis: tomando o inverso das transformações realizadas com os coeficientes da Tabela 2.1 e em seguida renormalizando as funções reescaladas é possível aproximar-se da distribuição do modelo para 32 bits (ver Fig. 2.4).

Olhando os coeficientes da Tabela 2.1 é possível sugerir uma aproximação simples para o valor do coeficiente angular da transformação da escala temporal:

$$c(j) \simeq [1 + 0.5 \cdot (j - 1)]. \quad (2.3)$$

Uma relação parecida também existe para os termos $d(j)$ que são proporcionais à diferença entre os termos constantes na regressão da função integral e, por isso, dependem do valor da distribuição de idade em zero.

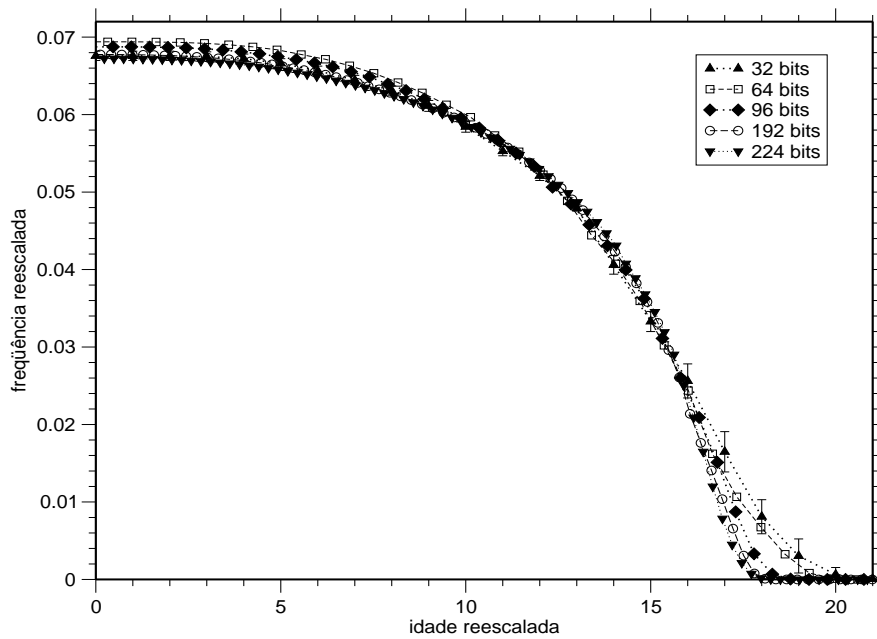


Figura 2.4: Transformando a escala temporal das distribuições obtidas com genomas múltiplos de 32 usando o inverso das transformações com coeficientes dados na Tabela 2.1, e depois renormalizando-as, é possível aproximar-se da distribuição para 32 bits.

Concluindo, os resultados das nossas simulações sugerem a existência de transformações de escala resumidas pelas seguintes leis:

$$\begin{aligned} 32 &\longrightarrow N = 32 \cdot j \\ t &\longrightarrow [1 + 0.5 \cdot (j - 1)]t + [0.5 \cdot (j - 1)] \\ \rho &\longrightarrow [1 + 0.5 \cdot (j - 1)]^{-1}\rho \end{aligned} \tag{2.4}$$

onde N é o número de bits no genoma do indivíduo e j um inteiro.

Dos resultados numéricos parece possível efetuar uma reescala das distribuições de idade ($\rho = \rho(t)$) com o uso dessas leis. Apesar disso, resulta bastante difícil poder afirmar o quanto essas leis de transformações permitem uma reescala puramente qualitativa ou se as funções reescaladas são estatisticamente indistinguíveis entre si. Por isso, se faz interessante uma comparação entre as funções de mortalidade, que podem ser diretamente derivadas das distribuições de idade através da relação seguinte:

$$f(a) = \log(\rho(a)/\rho(a + 1)), \tag{2.5}$$

onde $\rho(a)$ é o valor da distribuição na idade a . Estes novos dados são interessantes para esclarecer a qualidade das nossas leis pelo fato que os dados experimentais disponíveis são expressos por funções deste tipo.

Na figura 2.5 são representadas as funções de mortalidade depois de efetuada a reescala das distribuições de idade. Em uma escala linear, parece que as funções colapsam para idades pequenas e divergem claramente para idades maiores. Na figura menor, em uma escala log-linear, é possível ver como, no intervalo considerado, que apresenta só idades pequenas, as funções de mortalidade têm o mesmo comportamento, mas a diferença entre elas cresce com o incremento do tamanho do genoma. Desses dados, e com a ajuda das barras de erros apresentadas, é possível concluir que o colapso entre as funções não é totalmente alcançado.

É então possível concluir essa análise afirmando que as distribuições de idade reescaladas, embora tenham comportamentos similares, não correspondem quantitativamente para todos os tamanhos dos genomas.

É também interessante dar atenção ao efeito que a mudança no tamanho do genoma causa na frequência de bits deletérios no genoma. Tais distribuições mostram um aspecto típico do modelo. É possível reconhecer duas

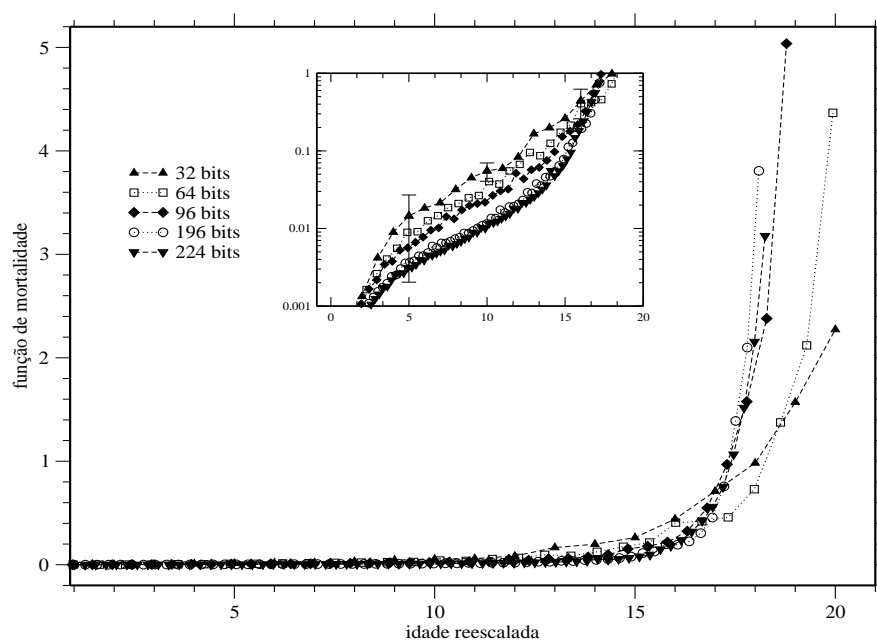


Figura 2.5: As funções de mortalidade computadas das distribuições de idade reescaladas (Fig. 2.4). A figura menor mostra as mesmas funções em uma escala semi-logarítmica para idades menores que 15. São também representadas em três pontos as correspondentes barras de erro.

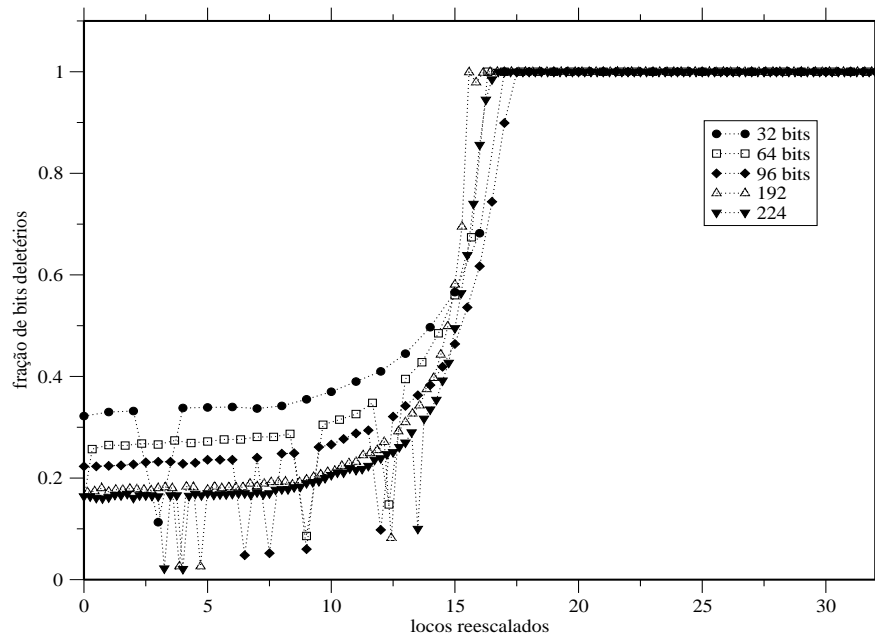


Figura 2.6: Efeito que a mudança no tamanho do genoma causa na frequência de bits deletérios do genoma. São apresentadas as funções resescaladas.

partes no genoma: uma primeira onde a frequência de bits mutados é inferior a um, responsável pela estrutura de idade da população. Uma segunda parte, bem distinguível a partir do loco 18, onde todos os bits têm um valor 1 e que não têm muita importância na dinâmica do modelo.

Em relação ao comportamento de escala, operando uma transformação na escala dos tempos do tipo daquela representada na equação 2.4, é possível obter distribuições similares em relação à posição na qual o genoma fica saturado por bits deletérios (figura 2.6). Ao contrário, não foi possível achar uma lei de escala simples em relação ao valor da primeira parte do genoma, onde resulta que a tamanhos maiores da tira corresponde uma menor frequência de bits deletérios.

2.4.2 O processo de renormalização

O coeficiente angular das transformações de escala 2.4 é obviamente fortemente dependente dos valores dos parâmetros da simulação. Particularmente

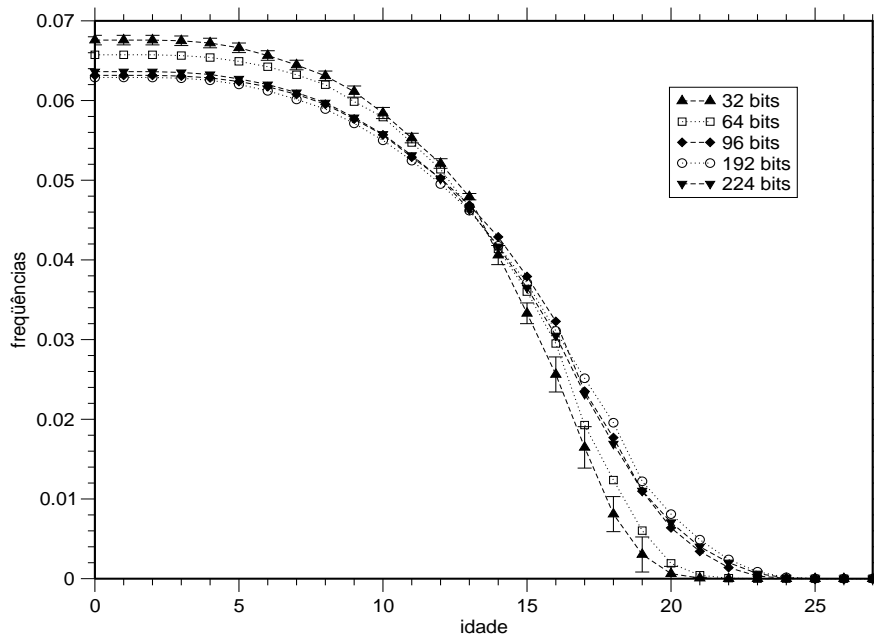


Figura 2.7: Diferentes simulações no caso onde o número de mutações e de locos dominantes são renormalizados em relação ao tamanho da tira de bits com o objetivo de manter constantes as respectivas densidades no genoma. Todas as simulações têm uma duração de 50000 passos de Monte Carlo.

interessante seria uma escolha dos parâmetros capaz de fornecer leis de transformação com coeficiente angular unitário. Nesse caso, seria possível obter a solução do modelo para 32-bits a partir de qualquer outro, só mudando-se adequadamente os valores dos parâmetros. Ou seja, ao mudar-se o tamanho da tira de bits, efetuar uma renormalização dos parâmetros deixaria os resultados indistinguíveis.

Concentramos a nossa atenção em apenas dois parâmetros, o número de locos dominantes para os alelos deletérios e o número de mutações geradas em cada geração. A idéia é de manter constante a densidade de mutações e locos dominantes no genoma, de forma independente do tamanho da tira de bits. Resulta então suficiente multiplicar os valores originais destes dois parâmetros no modelo a 32 bits por i , o tamanho do genoma dividido por 32. Os resultados desse processo de renormalização dos parâmetros é mostrado na figura 2.7.

É possível notar como as funções de distribuição das idades têm um comportamento parecido mas não corresponde quantitativamente ao mesmo valor para todos os tamanhos do genoma.

2.4.3 Resultados posteriores

A publicação de nosso trabalho sobre as propriedades de escala do modelo Penna reavivou na comunidade a discussão a respeito do tema. Num interessante estudo posterior ao nosso trabalho, os autores desenvolvem e ampliam a análise acima apresentada. Este artigo [46] tenta encontrar relações mais exatas na reescala das diferentes distribuições de idade. Com essa finalidade, foi estudado um modelo sem locos dominantes e onde a probabilidade de cruzamento é mantida proporcional ao tamanho da tira de bits: ao aumentar-se a dimensão do genoma, aumenta-se o número de processos de cruzamento. Ao contrário, no nosso modelo há sempre um único processo de cruzamento, independentemente do tamanho da tira de bits. Mudando também o valor de IMR com o fim de mantê-lo proporcional ao tamanho do genoma, são obtidas distribuições de idade similares, com uma concordância com uma possível lei de escala melhor que nos resultados precedentemente examinados. Mais, trata-se de uma lei de escala ainda mais simples do que aquelas apresentadas na equação 2.4, e correspondem a reescalar o tempo com o tamanho do genoma:

$$\begin{cases} 32 \longrightarrow N = 32 \cdot j \\ t \longrightarrow j \cdot t \end{cases}$$

Ademais, em concordância com a nossa análise, ao lado dessa grande similaridade nas distribuições de idades foram também observados diferentes estados genéticos e no potencial reprodutivo: ao crescer o genoma, a frequência de bits deletérios diminui e o tamanho da população cresce.

Enfim, renormalizando a taxa de mutação proporcionalmente ao tamanho do genoma e mantendo IMR constante, são obtidas distribuições de idades todas parecidas. De fato, só a primeira porção do genoma é responsável pela estrutura de idades da população e, com esta estratégia, esta porção relevante

é pouco afetada pela mudança de tamanho.

Foi também publicado um artigo com um estudo analítico das propriedades de escala de um modelo similar [47]. Trata-se de um modelo bem simplificado: assexuado, com um valor de $T = 1$ e sem idade mínima de reprodução. Baseia-se em uma equação que descreve o limite termodinâmico, para uma grande população, através da qual é possível modelar o sistema através de uma dinâmica determinista.

O estudo mostra que esta implementação do modelo Penna não exibe uma invariância de escala exata, embora tal invariância seja obtida no limite de uma taxa de mutação por bits que tende a zero. Este limite corresponde a um modelo Penna contínuo, que efetivamente resulta invariante por escala.

Estes resultados são coerentes com aqueles mostrados na figura 2.4, mesmo quando os estudos analisam implementações do modelo bastantes diferentes. Efetivamente, também no nosso caso, é possível notar como, ao crescer o tamanho da tira de bits - o que causa uma diminuição do valor da taxa de mutação de bits (passando de um modelo a 32 bits para um a 224 muda de 0.03 para 0.004) - as curvas se aproximam sempre mais uma da outra, sugerindo a existência de uma curva limite, correspondente a tamanhos infinitos (e então a uma taxa de mutação por bits perto de 0), onde o colapso seria mais realístico.

Capítulo 3

Microevolução, transição descontínua e sustentabilidade em um sistema ecológico ameaçado por uma caça seletiva

...
Quand on le laisse seul
Le monde mental
Ment
Monumentalment.

J. Prévert, *Paroles*

3.1 Sustentabilidade ecológica e fenômenos de microevolução

O impacto das atividades antrópicas nos equilíbrios ecológicos é, hoje em dia, um problema ubíquo e urgente, que desafia a comunidade científica na procura de novas idéias e soluções mais efetivas para enfrentar o desafio de uma gestão sustentável dos recursos naturais. Recentemente, nessa vasta e complexa área de estudos, caracterizada tradicionalmente por uma abordagem focada em estudos demográficos e fatores ecológicos, novos dados e

estudos têm mostrado como é também necessário considerar a existência de problemas derivados de mudanças de natureza puramente evolucionária [48]. Estes novos dados, que documentam mudanças evolucionárias causadas pela atividade do homem, sugerem que, para enfrentar problemas de biologia conservacionista, sejam efetuados também estudos evolutivos; ignorar os efeitos evolutivos pode causar conseqüências imprevisíveis e decisões inapropriadas na gerência dos recursos naturais. Um exemplo claro e evidente desse tipo de problema é apreciável através dos efeitos que uma atividade de pesca excessiva pode ter sobre as populações destes animais [50]. Pela sua natureza paradigmática, é possível olhar os efeitos dessa atividade como se foram os dados provenientes de um gênero de experimento de grande escala sobre a evolução. Dos resultados desse experimento é possível obter claras evidências do fato que algumas atividades humanas podem ser causa de uma aceleração dos processos evolucionários naturais. Isso porque a pesca causa nas populações de peixes uma mortalidade altamente seletiva dependente do tamanho dos indivíduos, ou seja, de um particular traço da presa. No entanto, o fato desse traço ser variável e que, dependendo do seu valor, confere uma vantagem ou desvantagem na sobrevivência do indivíduo, e conseqüentemente no sucesso reprodutivo, somado ao fato dessa variação no traço ser herdada, causa uma evolução na população em direção oposta à polarização causada pela coleta [51]. Ou seja, especificamente, regras de pesca que proíbem a captura de indivíduos menores que um determinado tamanho causam, ao longo das gerações, populações de peixes caracterizados, diretamente a nível genômico, por um tamanho menor e uma taxa de crescimento mais lenta.

É possível achar com facilidade outros exemplos deste fenômeno em diferentes áreas da biologia. Por exemplo, é bem estabelecido que as bactérias são capazes de desenvolver rapidamente uma resistência a vários antibióticos. Na natureza, sem a intervenção humana, essa imunidade só pode aparecer com uma probabilidade muito baixa. Ao contrário, a partir do uso sistemático desse tipo de drogas, a evolução da imunidade pode ser alcançada em poucos anos depois da introdução no comércio de um novo antibiótico [49].

Um outro exemplo bem documentado descreve a rápida resposta evolutiva à virulência de um parasita. Na Austrália em 1950, a tentativa de controlar a superpopulação de coelhos europeus que, na ausência de predadores,

tinham chegado a ser uma verdadeira praga, sugeriu a introdução do vírus Myxoma. No primeiro verão o vírus se difundiu dramaticamente matando a maioria dos coelhos infectados, mas rapidamente os coelhos desenvolveram uma resistência ao parasita. De alguns estudos de laboratório foi possível estabelecer como em sete anos a probabilidade de morte tinha caído de 90% para 30% [52].

Do exame desses primeiros estudos, embora seja muito difícil obter uma clara medida quantitativa da velocidade da resposta evolutiva, percebe-se como a variação evolucionária pode ser muito rápida, comparável com o tempo de vida de um ser humano. Nos referimos a esse tipo de resposta evolutiva, com uma escala temporal rápida, com o termo microevolução [48]. Uma série de dados recentemente publicada em relação a um estudo de larga duração sobre os efeitos da caça de troféus em uma população de cabras de grandes chifres nas Ram Mountain (Canadá) [53], deu uma sustentação válida a esse gênero de considerações permitindo uma medida direta e em nível genômico deste tipo de efeito. Esse estudo revela como, no intervalo temporal de só 30 anos, uma caça seletiva baseada em um caráter do fenótipo - nesse caso particular o tamanho dos chifres - pode ter causado o desaparecimento dos genes que conferiam um crescimento rápido do corpo e dos chifres. No caso, foi observado um declínio no valor médio das características herdáveis (*breeding value*) relativas ao peso e ao tamanho dos chifres, fato que indica uma resposta microevolucionária à seleção causada pela caça.

3.2 Modelos para a descrição dos aspectos fenotípicos de uma população

Para enfrentar este gênero de estudos é necessário, como primeiro passo, definir uma caracterização precisa dos atributos físicos ou bioquímicos de cada indivíduo (fenótipo), diretamente relacionado com a sua constituição genética (genótipo). Isso porque o processo evolutivo opera com uma força seletiva no espaço dos fenótipos, os quais são diretamente associados a valores adaptativos diferentes. A dinâmica na qual estamos interessados é fruto da interação entre um mapeamento não injetivo (muitos a um) dos genótipos nos fenóti-

pos e dos fenótipos nos valores adaptativos. Os dados biológicos confirmaram que esta dupla dinâmica é caracterizada por comportamentos muito diferentes: em uma escala macroscópica o fenótipo não muda por longos períodos, com pulos repentinos interrompidos por longos períodos de stasis [54], apesar de que, durante esses intervalos, a evolução em termos de seqüências genéticas não pára [55], na escala microscópica, difundindo ao longo de caminhos seletivamente neutros.

Em nosso processo de modelização vamos procurar manter este mapeamento não injetivo entre genótipos e fenótipos. Ao fazer isso, endereçamos a nossa atenção unicamente às diversidades geradas por diferenças no genótipo dos indivíduos. Ou seja, não consideramos o fato que as distinções entre genótipo e fenótipo são afetadas também pelo meio ambiente, até o ponto em que indivíduos com o mesmo genótipo possam desenvolver diferentes fenótipos. Endereçamos, pois, a nossa atenção unicamente à diversidade genética, a qual pode gerar uma eventual variação fenotípica.

Por isso, como primeiro elemento, nos permitimos associar diferentes alelos a cada loco, implementando, em particular, locos dialélicos. Ademais, vamos considerar um modelo capaz de representar indivíduos diplóides, distinguindo entre locos homocigotos e heterocigotos e introduzindo também o fenômeno das dominâncias. Por fim, queremos descrever traços poligênicos, ou seja que são influenciados por numerosos genes. Podemos concluir dizendo que vamos implementar um fenótipo consistente com a representação de um traço quantitativo (*quantitative trait*) mas em uma situação simplificada: o traço é poligênico mas o meio ambiente têm um efeito não relevante sobre o fenótipo. Mais, consideramos traços quantitativos ideais, onde os genes são associados aleatoriamente (*linkage equilibrium*) e os seus efeitos são puramente aditivos [56].

Com a finalidade de realizar a modelização de um traço fenotípico com estas propriedades, é suficiente representá-lo por meio de uma tira de bits com as mesmas características daquela que representa a estrutura de idade do indivíduo: 32 bits, diplóide e com os alelos que assumem os valores 0 ou 1. O valor 0 é associado a um alelo igual ao presente na população original - alelo *wild type* -, o valor 1 aos alelos que sofreram uma mudança e tornaram-se alelos mutantes.

Uma vez definida a parte do material genético que codifica o traço fenotípico temos que associar um valor de fenótipo com este pedaço de genoma. Isso é feito somando-se o conjunto de mutações (bit com valor 1) ativas presente na tira de bits relativa ao traço. Com mutação ativa entendemos um loco homozigoto, onde os bits dos dois alelos homólogos têm o valor 1, ou um heterozigoto no qual o alelo mutante é dominante. Como no caso da tira que representa a estrutura de idade, o número de locos onde os alelos mutantes são dominantes é fixo e a posição deles é escolhida aleatoriamente no começo da simulação. Seguindo este processo, o valor do fenótipo é um inteiro entre 0 e 32 e é diferente da simples soma de todos os uns presentes nas tiras de bits. Adiante nos referiremos a este segundo valor pelo termo “valor da tira de bits”. A diferença essencial está no fato de que o valor do fenótipo considera a natureza diplóide do genoma, por meio dos efeitos da dominância e da homozigose. Ao contrário, o valor da tira de bits é simplesmente relacionado à frequência do alelo 1 no genoma.

A variabilidade genética e, em correspondência, a fenotípica, é assegurada pelas ações das mutações. As mutações são obtidas trocando-se aleatoriamente o valor de um bit da tira. Os bits podem mudar de 0 para 1 ou de 1 para 0, ao contrário dos alelos da tira de bits que descreve a idade, onde são implementadas só mutações de 0 para 1. Esta simplificação não considera o fato que é realmente difícil de que um alelo mutado recupere a sua atividade precedente.

Resumindo, o genoma inteiro de cada indivíduo contém, além da tira estruturada por idade e que controla o relógio biológico dos indivíduos, outras duplas de tiras de bits que codificam os traços fenotípicos e são responsáveis pelas interações intraespecíficas e/ou com o meio ambiente. Com esta técnica, já bem estabelecida, é possível implementar vários tipos de seleções no interior de uma população [58, 59, 60].

3.3 Um modelo para a descrição das populações de cabras das Ram Mountain

Numa publicação prévia [57], uma implementação simples do modelo Penna conseguiu obter algumas predições em relação às mudanças causadas por uma pesca intensiva de uma população com uma forte correlação entre idade e tamanho dos animais. Este último fato permitiu o uso de uma única tira de bits, representando a estrutura de idade do indivíduo, para descrever inteiramente as interações mais relevantes do sistema. No caso, não foi então necessário implementar nenhum traço fenotípico. Esta situação difere do nosso caso, no qual somos obrigados a definir para cada indivíduo um correspondente fenótipo.

Com essa finalidade consideramos um modelo simplificado, inspirado na dinâmica do repertório genético da população das Ram Mountain, onde o modelo computacional de referência é o modelo Penna [25], adequadamente adaptado para obter soluções estáveis para populações sujeitas a fenômenos de microevolução e para descrever com detalhe a dinâmica relativa. Consideramos então uma população sexual e diplóide, onde cada indivíduo é representado por um genoma simulado por duas tiras de bits acopladas. O comportamento coletivo destes agentes é definido pelas regras bem estabelecidas do modelo Penna, assim como descrito no capítulo precedente. A única diferença é o fato que, desta vez, cada indivíduo é representado por dois pedaços de genoma: um primeiro relativo à estrutura de idade, enquanto o segundo determina o traço biológico que caracteriza o indivíduo. As regras do ciclo reprodutivo são esquematizadas na Figura 3.1. O fenótipo do indivíduo, como descrito no parágrafo precedente, é obtido somando-se todas as mutações ativas presentes na tira de bits que representa esse traço. Por fim, a população é sujeita a um fator de Verhulst clássico [58], onde a probabilidade de morte depende da densidade da população.

Estas condições geram uma dinâmica de população simples e bem conhecida: a parte do genoma estruturada por idade causa o envelhecimento da população respeitando a lei de Gompertz [9]. De outro lado, o valor da tira de bits que determina o traço fenotípico, não sendo sujeita a nenhuma interação e sendo governada unicamente pela tendência à equiprobabilidade dos

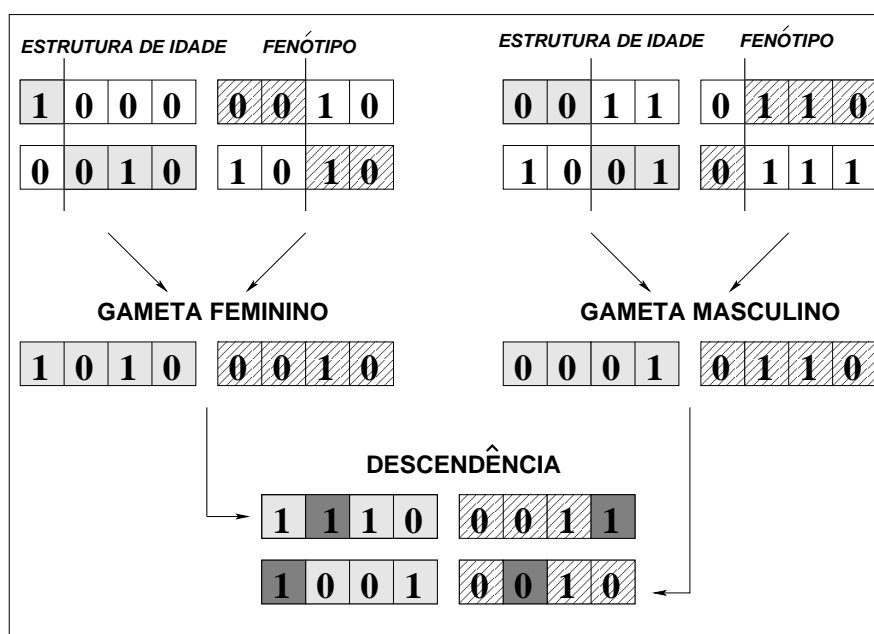


Figura 3.1: Representação do ciclo reprodutivo. O genoma diplóide é apresentado com a sua parte estruturada por idade e a tira de bits que codifica o fenótipo (fundo com linhas diagonais). Na primeira linha é realizado o cruzamento. Na segunda passagem vem escolhido o gameta haplóide. Por fim, são juntadas algumas novas mutações (quadrados mais escuros) e os gametas combinam formando um novo indivíduo.

dois valores alélicos, gera uma distribuição Gaussiana com um valor médio em 16.

A esse modelo básico é necessário juntar as duas forças principais que governam a dinâmica da distribuição dos fenótipos: a seleção sexual correlacionada com o aspecto fenotípico e a caça seletiva baseada no valor do fenótipo.

As regras implementadas para definir essas forças são inspiradas nos detalhes relativos aos dados coletados na população das Ram Mountain [62]. No exame desse estudo emerge que a seleção sexual aumenta a probabilidades de sucesso no acasalamento com o aumentar da idade, do tamanho dos chifres e do corpo das cabras. Ou seja, é observado um fenômeno que tende a concentrar as paternidades nos animais mais favorecidos em algumas características fenotípicas. Na implementação do nosso modelo essa dependência do sucesso reprodutivo em dois fatores é representada, por razões de simplicidade, em um único traço. Ademais, pela mesma razão, não consideramos que o sucesso no acasalamento tenha também uma dependência na idade do indivíduo. Dessas considerações parece evidente que o primeiro ingrediente essencial na descrição de uma dinâmica realista seja permitir a concentração da paternidade. Por isso é importante permitir que um macho possa acasalar mais que uma vez em cada ciclo reprodutivo. A seleção é implementada nessa forma: a cada passo temporal, cada fêmea que seja reprodutivamente ativa, faz uma escolha não casual do seu parceiro. Esta seleção é feita escolhendo-se entre 20 machos aquele que apresenta o máximo valor do fenótipo. Indicamos com o termo de dinâmica extrema este tipo de escolha. A introdução desta seleção sexual causa uma deriva na distribuição da tira de bits que representa o traço fenotípico em direção a valores maiores (na figura 3.4 é possível observar como a distribuição, depois de 100000 passos temporais, é centrada próxima de 26).

A pressão de caça que simulamos é inspirada nas condições e nas regulamentações vigentes na região das Ram Mountain [53]. Estas prescrições só permitem a caça de animais mais velhos de que quatro anos e com chifres que chegam a completar a curva (ver figura 3.2), isso até um número máximo estabelecido de animais caçados.

No modelo permitimos a caça de indivíduos machos mais velhos que 4



Figura 3.2: Um exemplar de macho com chifres que chegam a completar a curva inteira. Em consequência das atuais leis de preservação ambiental, estes exemplares caracterizados por grandes troféus são os mais caçados.

anos e com um valor do fenótipo maior que 16. A coleta dos animais é realizada seguindo esses passos: são feitas um número de tentativas ao acaso (aproximadamente igual à dimensão da população) de encontrar indivíduos que satisfazem as características acima apresentadas. No caso sejam encontrados, são eliminados da simulação. A rotina de caça termina quando o número de animais mortos chega ao número máximo de animais caçáveis, valor fixado no começo da simulação.

Com as regras agora definidas deixamos a dinâmica populacional evoluir por 50000 passos temporais. Depois de ter alcançado esse primeiro equilíbrio, a seleção de caça é ativada. Estimamos por meio de algumas simulações demoradas que, com esse modelo, seja suficiente esperar outros 50000 passos de Monte Carlo para alcançar o equilíbrio final.

3.3.1 O tamanho da população

Uma atenção particular tem que ser prestada em relação ao tamanho da população. Estima-se a população real das Ram Mountain, durante os anos de observação, próxima de 140 animais, com grandes flutuações durante os diferentes períodos. Para efetuar uma modelagem muito fiel aos dados experimentais seria necessário simular uma população muito pequena, uma tarefa bastante problemática, não somente pela dificuldade de obter-se uma análise estatística confiável dos resultados, mas, sobretudo, porque as flutuações de uma população tão pequena destróem os elementos chaves do modelo: a dinâmica geral não é mais guiada pela dinâmica das tiras de bits, mas, ao contrário, pelo fator de Verhulst. Fica então difícil encontrar uma distribuição dos fenótipos que alcance um equilíbrio. Por todas essas razões decidimos descrever o comportamento de uma população mantida próxima de 10000 indivíduos, considerando esta a menor população ainda apta a um estudo baseado no nosso modelo e que permita uma abordagem estatística consistente.

Populações deste tamanho podem ser consideradas, por exemplo, mais pertinentes no caso do estudo de bactérias. De outro lado, o sucesso que o nosso estudo teve em reproduzir os resultados da referência [53], nos deixa confiantes que o tamanho da população não seja um elemento fundamental no

processo de modelização deste fenômeno. A apoiar esta hipótese vem também o resultado de algumas simulações rodadas com cerca de 250000 indivíduos. Não foi possível encontrar diferenças qualitativas, sugerindo então que não seja possível determinar uma dependência dos nossos resultados em efeitos de escala finita.

3.4 Resultados

3.4.1 Os dois estados extremos da dinâmica

Como primeiro passo vamos escolher uma variável significativa para descrever o estado do nosso sistema. Decidimos usar o valor médio da distribuição da tira de bits relativa ao traço fenotípico. Escolhemos este parâmetro, no lugar do valor efetivo do fenótipo, porque queremos manter o nosso foco em mudanças evolucionárias, e então variações em nível do genoma.

Na realidade, a escolha entre valor médio das tiras de bits e fenótipo é bastante indiferente, sendo os resultados qualitativamente idênticos (ver figura 3.3). Isso porque no nosso modelo a relação entre fenótipo e genoma é bem simplificada e direta, não levando em consideração fatores ambientais.

Para definir os estados do sistema temos que necessariamente ter uma abordagem estatística. Então consideramos o sistema como estando em um estado diferente quando a população sofreu uma variação no repertório genético suficiente para mudar o valor médio da distribuição das tiras de bits por um valor maior que o desvio padrão.

Como fica evidente na Fig. 3.4, o estado fundamental do sistema é uma distribuição gaussiana centrada em 26 com um desvio padrão de 2. Quando a seleção de caça é ativada, ela determina uma deriva na direção de valores menores. Selecionando um número máximo de animais caçados igual a 230 em cada passo temporal, a distribuição chega no seu menor valor médio (19). A partir deste valor, um aumento no número de animais caçados não causará outras mudanças nas distribuições. A distribuição com valor médio 19 pode ser considerada como um estado de equilíbrio extremo da dinâmica do sistema. Por meio dos dados relativos a mais que 1000 simulações, concluímos que não é possível encontrar estados incluídos entre estes dois valores extre-

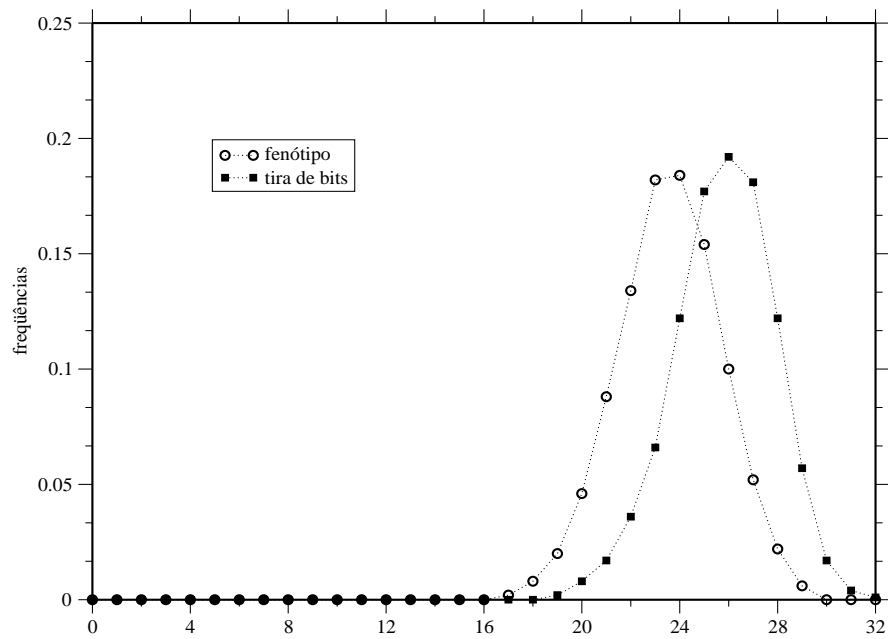


Figura 3.3: Diferenças entre o valor da tira de bits e o fenótipo. O comportamento é qualitativamente parecido, com o valor do fenótipo fortemente correlacionado, mas obviamente sempre menor que o valor da tira de bits.

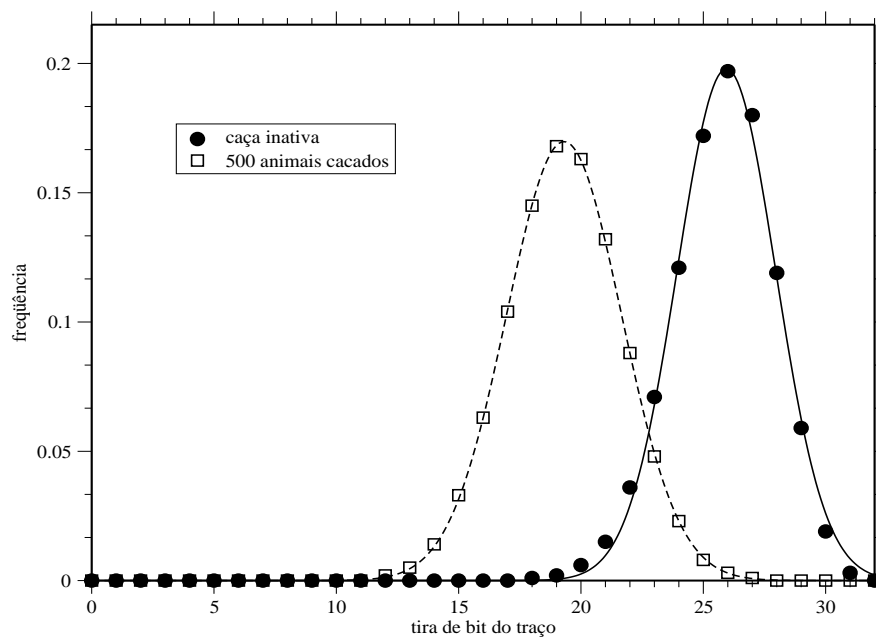


Figura 3.4: Distribuição dos valores das tiras de bits que representam o traço fenotípico para os dois estados extremos (sem caça e com uma pressão de caça igual a 500). Os dados são bem interpolados por funções Gaussianas com valor médio de 25.9 e 19.3, e desvio padrão, respectivamente, de 2 e 2.4.

Parâmetro de Verhulst	130000
População inicial	2000
Idade mínima de reprodução	8
Idade máxima de reprodução	32
Número de filhos	1
Máximo valor de doenças ativas	3
Número de mutações	1
Locos dominantes	6

Tabela 3.1: Principais parâmetros usados nas simulações. O número de mutações e locos dominantes é igual nas duas tiras de bits, a que representa a estrutura de idade e a do traço fenotípico.

mos, 19 e 26. Disso segue que, variando-se o número de animais caçados, não é possível visitar todos os estados possíveis no espaço definido pelo valor da tira de bits. Existe uma região proibida, entre 21 e 23, nunca visitada pelo valor médio da distribuição da tira de bits. De fato, o modelo apresenta só dois estados estacionários diferentes, aqueles que definimos como extremos.

De um ponto de vista puramente biológico, consideramos o sistema em estados diferentes se, partindo de uma condição de equilíbrio inicial com uma distribuição centrada em 26, é observada uma variação estatisticamente relevante no repertório genético da população. Com essa definição, o valor médio da distribuição da tira de bits resulta efetivamente ser um bom parâmetro. Mais, apresentando-se só dois valores, a descrição do fenômeno se simplifica notavelmente. Enfim, a existência de uma descontinuidade entre estes dois estados sugere a presença de uma transição tipo primeira ordem entre eles. Isso resultará evidente se for possível definir um efetivo valor de limiar para o número de animais caçados. No parágrafo seguinte veremos que esse é o caso.

3.4.2 A transição descontínua

Realizamos cem diferentes simulações para cada conjunto de parâmetros, para investigar o comportamento do repertório genético da população em

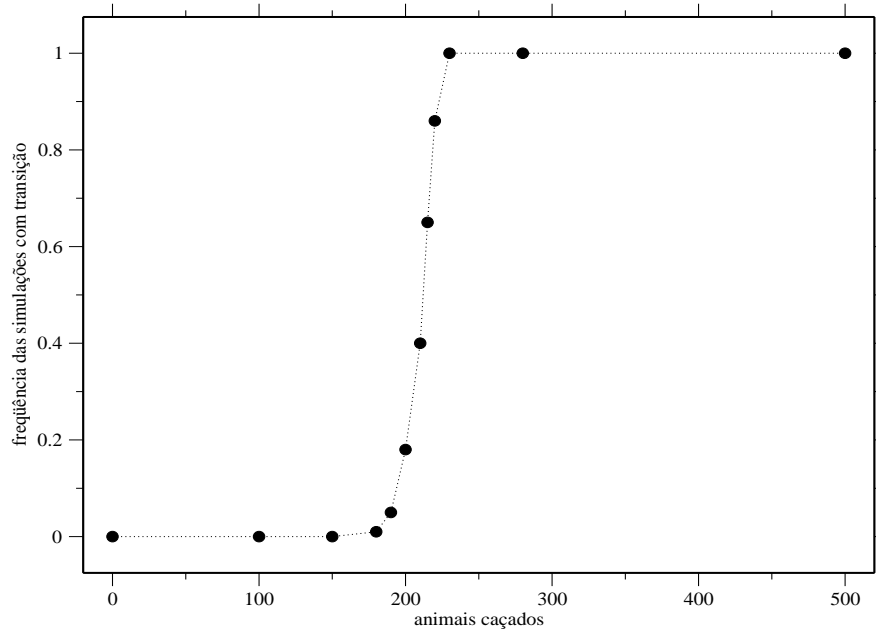


Figura 3.5: Probabilidade de alcançar a transição em função do número de animais caçados.

função dos números de animais caçados. Os resultados são apresentados na figura 3.5, onde para cada valor do número de animais caçados é apresentada a probabilidade de obter-se uma transição, ou seja, uma mudança apreciável, no repertório genético da população.

Parece evidente que, para valores menores que 200, a fração de simulações com um valor médio da tira de bits do traço fenotípico menor que 23 não é relevante. Com 200 mortos, em 20% das simulações o modelo sofre uma transição na distribuição da tira de bits, alcançando um valor médio perto de 19. Aumentando o número de animais caçados, o número de simulações que passa por uma transição também aumenta, e, para valores maiores que 230, todas as simulações terminam com uma distribuição com valor médio perto de 19.

É possível descrever o mesmo fenômeno registrando diretamente o comportamento do parâmetro de ordem (ver Figura 3.6). Examinando esses da-

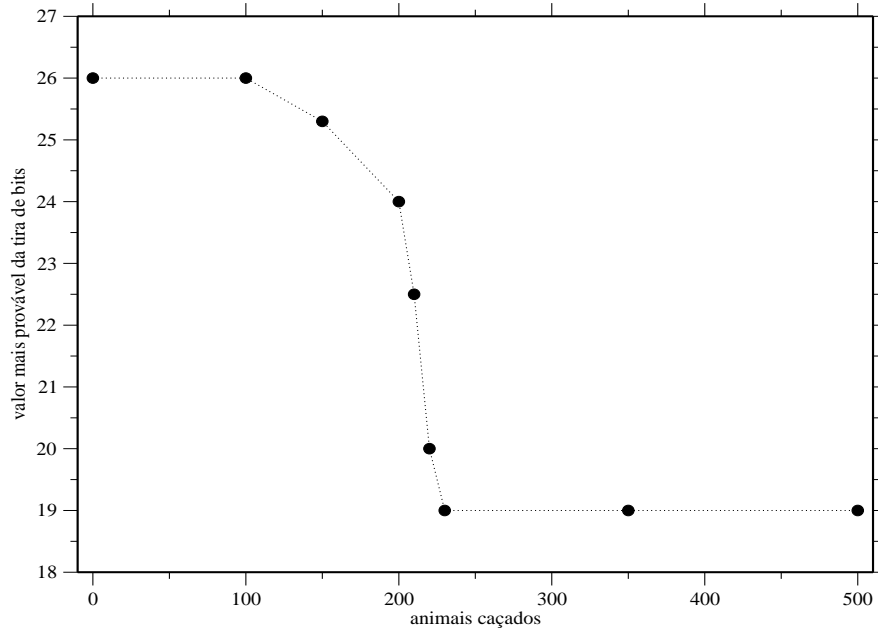


Figura 3.6: Transição do parâmetro de ordem do sistema ao passar do valor de limiar de 210 animais caçados.

dos observa-se novamente que o valor do numero máximo de animais caçados igual a 210 pode ser considerado o valor de limiar da transição.

3.4.3 Como a dinâmica representa uma microevolução

Um estudo do comportamento dinâmico do modelo permite apreciar como seria possível representar com sucesso um processo de microevolução. De fato, olhando a figura 3.7, pode-se notar como a evolução do sistema se comporta de forma realmente abrupta, e 100 passos temporais apenas já são suficientes para levar o valor da tira de bits de um estado ao outro. Por exemplo, caçando-se 500 animais, todas as simulações alcançam a transição em menos de 100 passos. Estes fatos testemunham a rapidez com a qual acontece a transição descrita até agora. Exatamente por estas razões podemos afirmar que o nosso modelo representa com sucesso o fenômeno de microevolução que foi a motivação do nosso trabalho. No entanto, uma transição rápida não é

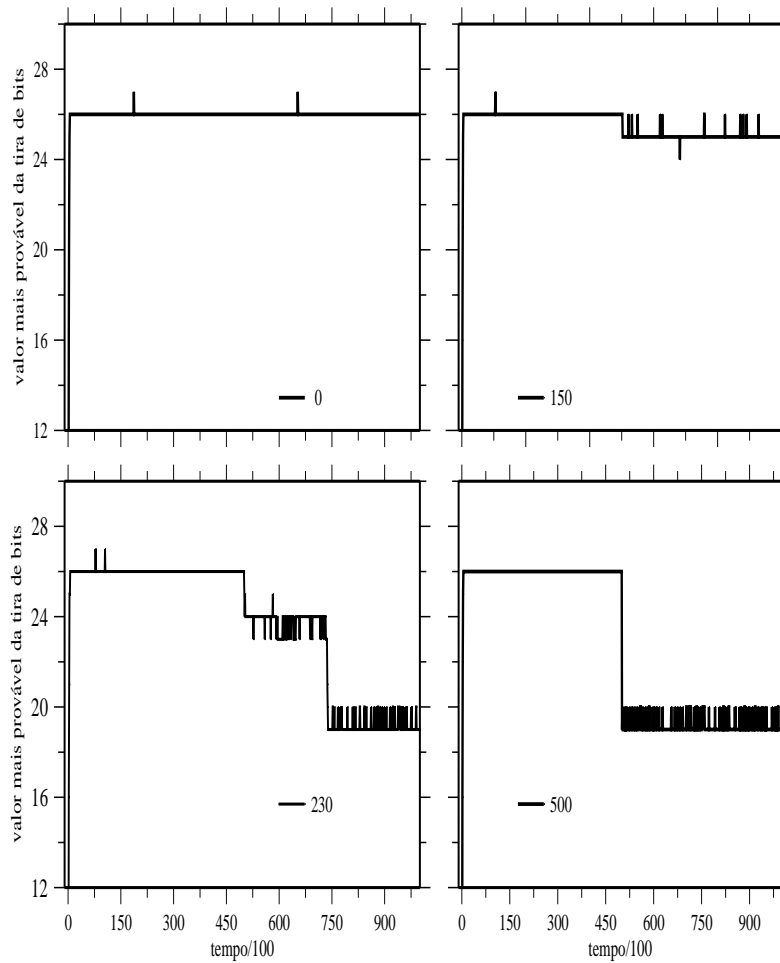


Figura 3.7: O valor mais provável da distribuição da tira de bits do traço fenotípico em função do tempo: com a caça desativada, alcança o valor de 26 e não muda mais. Caçando-se 150 animais, a partir do passo 50000, o valor passa para 25. Para 230, muda para 24 antes de se estabilizar no estado de equilíbrio de 19. Na última simulação, caçando-se 500 exemplares, verifica-se uma rápida transição para 19.

o único resultado possível deste tipo de dinâmica. Em algumas simulações, onde o número de animais caçados está entre 210 e 230, o valor da tira de bits que representa o traço fenotípico se torna metaestável antes de alcançar o estado estacionário, e a escala temporal da transição torna-se grande.

3.5 Quais informações o modelo fornece?

3.5.1 O interesse por uma dinâmica não trivial

Este capítulo apresentou como é possível enfrentar estudos de processos microevolucionários com o uso de modelos microscópicos de dinâmica populacional. Dando atenção à escolha dos ingredientes, conseguimos definir um modelo simplificado capaz de imitar os principais elementos da dinâmica do repertório genético de uma população, mostrando as estruturas dominantes do seu espaço de fase. Notamos como a competição entre seleção sexual e um estresse externo, determinado por uma caça baseada no valor do fenótipo dos indivíduos, gera uma transição no repertório genético da população no momento em que é alcançado um valor de limiar de animais caçados. Estes resultados testemunham uma surpreendente transição descontínua em um modelo biológico, um fato relativamente incomum.

Neste sistema, altamente não linear e acoplado com um estresse externo que age de forma estocástica, esta fonte de flutuações aleatórias gera um novo estado que não existia na ausência de tal acoplamento [63].

3.5.2 A reconstrução de dados experimentais e a possibilidade de fazer algum tipo de previsão

O nosso estudo permite descrever a dinâmica de um fenômeno biológico aplicando uma linguagem típica de uma ciência dura, como a física estatística. Isso é importante para começar um processo de modelização do fenômeno e para individualizar as formas para uma definição abstrata do mesmo. É este um primeiro passo indispensável para poder-se aplicar conceitos matemáticos mais complexos e continuar no caminho de uma descrição interdisciplinar destas problemáticas.

Além disso, é efetivamente possível confrontar com sucesso os resultados do modelo com dados experimentais. Na nossa simulação, o valor de limiar do número de animais caçados é igual a 0.02% da população; este valor é exatamente aquele registrado nos dados de campo [62], onde os efeitos de microevolução resultaram evidentes.

Ademais, o modelo confirma como os mecanismos minimais que os biólogos apontaram como responsáveis pelos efeitos de microevolução [53] são efetivamente suficientes para obter-se uma variação significativa no repertório genético da população. Ou seja, o modelo age como um teste experimental das condições determinadas nos estudos biológicos.

Por fim, o nosso estudo permite sugerir algumas estratégias de gerenciamento ambiental. O fato de ter-se identificado uma transição de fase descontínua adverte que é necessário prestar muita atenção na regulamentação do número de animais caçados. Uma escolha um pouco diferente deste valor pode levar de regimes, do ponto de vista evolucionário, bem gerenciados e inócuos à situações catastróficas, no momento em que este valor se aproxima perigosamente do valor de limiar. Um comportamento similar já tinha sido descrito em um trabalho precedente, onde resultava evidente como um pequeno aumento nas taxas de pesca do bacalhau podia destruir populações estáveis [61]. Ademais, podemos concluir, em concordâncias com um trabalho precedente baseado no modelo Penna [57] e com as opiniões de biólogos [51], que não é suficiente ter um limite mínimo nos valores do tamanho de algum traço dos animais caçáveis mas é também oportuno impor um limite máximo para evitar uma evolução da população em direção oposta à polarização causada pela coleta. De fato, é importante salvar os indivíduos de tamanho maior não só em espécies onde a fertilidade é diretamente ligada a esta característica [57]. Em geral, sendo o tamanho maior ou o crescimento mais rápido correlacionado com a saúde e a qualidade do patrimônio genético dos indivíduos, esta medida resulta importante para evitar o declínio dos indivíduos com valor adaptativo, em termos ideais, maior.

Os resultados obtidos por meio deste modelo, que descreve fatos puramente dinâmicos do sistema - a análise de propriedades coletivas em sistemas estatísticos de muitos agentes submetidos a um estresse externo - sem entrar nos detalhes biológicos, permitem definir alguns mecanismos não evidentes e

podem ajudar na definição de estratégias para a solução de problemas que, em princípio, parecem explicáveis só em termos ecológicos e biológicos.

Capítulo 4

Competição e dinâmica evolutiva: o fenômeno da especiação simpátrica

πόλεμος πάντων μὲν πατήρ ἐστι,
πάντων δὲ βασιλεὺς ...

Eraclito, *Fr.53*

4.1 O problema biológico da especiação

Coerentemente com a moderna definição de espécie, a especiação é o processo que explica a geração de duas populações reprodutivamente isoladas, ou seja, privadas de qualquer forma de fluxo genético. Em linha com a moderna teoria da evolução, também o processo da especiação quer ser explicado como o resultado de uma dinâmica microscópica governada pelo antagonismo entre seleção natural e a recombinação de formas mutáveis.

A conjectura mais aceita pelo pensamento dominante considera os eventos de especiação gerados por barreiras geográficas, ou por qualquer outro mecanismo de isolamento físico, que separa alguma fração da população de uma espécie. A causa do surgimento do isolamento reprodutivo é então reconduzida à deriva genética aleatória (*random genetic drift*): a ausência de qualquer mecanismo que obrigue os indivíduos a manter a compatibilidade

reprodutiva, junto à falta física de fluxo genético, causa uma divergência na população, independentemente de qualquer particular força evolutiva. Este processo, chamado especiação alopátrica, é hoje em dia referendado por várias evidências biogeográficas, onde espécies irmãs tendem a ser separadas por barreiras geográficas [64, 65, 66].

Não obstante a partir dos anos 50 esta visão puramente geográfica do fenômeno da especiação se estruture como um dogma, trabalhos mais recentes suportaram uma nova classificação dos possíveis modos da especiação, mais baseada nos mecanismos que na geografia. Estas conjeturas supõem que a especiação é também possível na ausência de qualquer mecanismo de isolamento físico: uma população caracterizada por um único *continuum* de organismos que se reproduzem entre eles pode dividir-se em duas subpopulações reprodutivamente isoladas. Para ser preciso, estas conjeturas não são novas, datando do tempo de Darwin [3], se bem que de uma perspectiva em parte diferente daquela dos nossos tempos. Um grande interesse neste tema cresceu a partir de estudos de laboratório que mostraram que a seleção disruptiva podia conduzir ao isolamento reprodutivo [67]. A seguir, um modelo com uma dinâmica de um loco bialélico mostrou como, em um ambiente com dois nichos ecológicos, pode existir um polimorfismo também no caso de uma única população com acasalamento aleatório [68]. A sustentação destas investigações teóricas chegaram diretamente da natureza através de dados observacionais diretos, como no caso da mosca da maçã *Rhagoletis pomonella* [69] e das evidências filogenéticas dos peixes ciclídeos de alguns pequenos lagos vulcânicos em Camarões [70]. Por fim, resultados mais recentes demonstraram a especiação simpátrica *in vitro* [71] ou com simulações baseadas em vidas artificiais [72]. Uma boa revisão geral sobre o tema da especiação simpátrica se encontra no artigo de Via [74].

4.1.1 Uma vasta literatura de modelos

A complexidade dos modelos usados para abordar o problema da especiação torna as teorias verbais dominantes nesta área. Essas, por outro lado, são satisfatórias só ao explicar a origem do isolamento reprodutivo em alopatría. No caso simpátrico, ao contrário, não resultam ser suficientemente

convincentes e impõe-se a construção de um rico aparato argumentativo e de modelos gerais convincentes capazes de dar estrutura e plausibilidade às teorias verbais (para uma revisão geral do problema ver Ref. [65, 73]). Nesta perspectiva, o uso de modelos matemáticos e computacionais está criando uma nova e interessante área de pesquisa.

Sendo típico do estudo da especiação simpátrica uma grande complexidade de detalhes e uma forte dependência das condições iniciais e de contorno, uma abordagem tradicional, que tem como objetivo previsões quantitativas, é substituída por previsões de natureza qualitativa, relativas a viabilidade de alguma dinâmica particular num bem definido esquema de regras. Desta forma, as simulações matemáticas e computacionais viram um instrumento de análise empírica e de suporte no teste de conjecturas.

Podemos concluir que o objetivo deste tipo de trabalho, ou seja as suas previsões, é testar regras e mecanismos capazes de gerar eventos de especiação. Quer-se desenvolver modelos que sugerem o conjunto de condições ecológicas e genéticas, em termos de balanço de forças evolutivas, que facilitam eventos de especiação em simpatria. Encontrar estas condições nas populações reais daria um apoio definitivo à plausibilidade da especiação simpátrica. Ademais, a utilidade destes estudos está na capacidade de organizar teorias verbais, especificando as interrelações e a consistência de mecanismos alternativos, guiados por um princípio geral de economia de pensamento.

Dentre a grande variedade destes modelos, restringimos a nossa análise ao caso do isolamento reprodutivo gerado pela ecologia e concentramos a nossa atenção em modelos nos quais a seleção natural dependente da distribuição dos recursos naturais origina a especiação simpátrica. Um primeiro tipo de modelo representava ambientes heterogêneos com dois nichos onde estava presente um mecanismo de adaptação a esta distribuição discreta de recursos. Neste caso, a especiação era causada por uma seleção fixa no tempo que favorecia os fenótipos dos dois extremos do possível intervalo fenotípico (seleção disruptiva) e causava o isolamento reprodutivo. Um modelo com estes ingredientes apareceu recentemente na literatura [75], e um similar, mas baseado em uma dinâmica microscópica, foi também estudado nos artigos [58, 60]. A idéia principal na qual são fundamentados estes modelos sugere uma nova questão, se para explicar a coexistência de espécies em simpatria seriam ne-

cessários nichos discretos. Uma resposta negativa foi dada por uma análise [76] que tentou demonstrar como a especiação pode surgir da competição por recursos distribuídos de forma contínua. Em tal situação, o meio ambiente apresenta um único nicho ecológico e, conseqüentemente, a população sofre uma seleção estabilizante que favorece fenótipos intermediários. Neste cenário, a introdução de um mecanismo de competição gera uma seleção disruptiva que se enfraquece com a divergência das espécies. Esta competição é gerada através de uma força de seleção que é mais forte para fenótipos mais freqüentes e mais fraca para aqueles raros (seleção dependente da freqüência). Este mecanismo é interessante por duas razões principais. Primeiro, esta freqüência de acontecimentos é controlada por uma interação dependente das freqüências e não precisa de nenhuma seleção disruptiva imposta pelo meio externo. Ademais trata-se de um mecanismo mais geral, que pode acontecer em um conjunto bem mais vasto de possíveis habitat ecológicos. Por fim, como sublinhado em um recente artigo [78], esta competição considera realisticamente a deterioração nas condições ambientais e se atualiza constantemente de acordo das mudanças na distribuição fenotípica.

As duas forças seletivas apresentadas até agora (seleção estabilizante e seleção dependente das freqüências) não são ainda suficientes para gerar especiação neste modelo. Para obter-se um processo estável é necessário prevenir o aparecimento de fenótipos intermediários através da divergência da compatibilidade reprodutiva. Isto é obtido através da evolução do acasalamento preferencial, uma estratégia de acasalamento não aleatória. No caso de uma distribuição de recursos unimodal, este mecanismo é também estritamente necessário para o surgimento de uma primeira divergência fenotípica. Todas estas idéias foram recentemente desenvolvidas na Ref. [76] e são o ponto de partida da nossa análise, que propõe estudar esses resultados para um modelo microscópico mais complexo.

4.1.2 Recentes críticas ao modelo de Dieckmann e Doebeli

Um trabalho recente produzido por Polechova e Barton [79] encontra explicações alternativas à relativa facilidade com a qual o modelo de Dieckmann

e Doebeli [76], tanto para modelos assexuados como para modelos sexuados, obtém eventos de especiação. No caso assexuado propõe que a formação do aglomerado fenotípico é na realidade um transiente e que no caso sexuado é um efeito do tamanho limitado do intervalo de fenótipos, sendo que Dieckmann e Doebeli determinaram o fenótipo a partir de apenas 5 locos. Argumenta também que o acasalamento preferencial sem custos seletivos (fato que será explicado adiante, na seção 4.2.2) facilita grandemente fenômenos de especiação simpátrica.

Estas conclusões são o fruto de uma análise inteiramente determinista. Este fato nos sugeriu uma atitude precavida em relações às conclusões obtidas. Em sistemas dessa natureza, a ausência de estudo dos fenômenos de flutuações, como bem exemplificado por vários casos marcantes na teoria dos fenômenos críticos, não pode ser considerado um simples detalhe, nem é possível reduzir os efeitos estocásticos aos de uma simples perturbação. Com o nosso estudo queremos então definir um modelo microscópico, que contemple flutuações, que possa testar estas críticas e ademais introduzir ulteriores elementos de confiança sobre a robustez da hipótese de que a especiação simpátrica possa emergir por meio da competição por recursos distribuídos de forma contínua.

Com essa finalidade, o fenótipo da população será sempre determinado por uma dinâmica microscópica em nível do genótipo. Concentraremos também muita atenção em garantir uma alta variabilidade genética, implementando mecanismos de cruzamento, dominância e locos homo/heterozigóticos. Enfim, estudaremos uma população estruturada por idade, sendo interessante ver se este fenômeno, presente em qualquer tipo de população, afeta de alguma forma a dinâmica da especiação.

4.1.3 Populações assexuadas

Na continuação descreveremos resultados relativos a populações tanto sexuadas quanto assexuadas. No primeiro caso, estamos considerando eventos de especiação simpátrica no sentido estrito do termo, que coincide com o surgimento do isolamento reprodutivo. No segundo caso, estamos investigando a emergência de polimorfismos, indicando com este termo a subdivisão de

uma população fenotipicamente monomorfa em dois ou mais aglomerados fenotípicos distintos.

4.2 O modelo

Como já discutido, estamos interessados num modelo estruturado em um nível microscópico, no sentido de uma representação em nível de indivíduos, e em definir as características de cada agente por meio de uma dinâmica gerada em nível genômico. Tal caracterização tem que permitir a descrição de uma população estruturada por idade e com uma alta variabilidade fenotípica. Para conseguir realizar estes objetivos o modelo Penna modificado com a introdução de um traço fenotípico, como já apresentado no capítulo precedente, na descrição da população das cabras das Ram Mountain, é um ótimo ponto de partida. Descreveremos dois tipos de populações: diplóide e sexuada e haplóide e assexuada. Como já deveria estar claro para o leitor, o caso sexuada refere-se a uma população diplóide onde cada indivíduo é representado por duas duplas de tiras genéticas. A primeira permite implementar a estrutura de idade e é a responsável pelas mortes causadas por envelhecimento. A segunda, somando todas as suas mutações ativas, permite associar a cada indivíduo um fenótipo. O modelo assexuada é implementado exatamente da mesma forma mas a natureza haplóide dos indivíduos, onde a única causa de variabilidade é devida ao processo de mutação, simplifica notavelmente as regras reprodutivas. Uma descrição mais precisa pode ser encontrada na seção 2.3. Com este modelo podemos indagar os efeitos que a estrutura de idade e os principais mecanismos de variabilidade genética (cruzamento, locos homo/heterozigotos e dominância), e portanto fenotípica, têm sobre o fenômeno da especiação.

4.2.1 Seleção natural e competição

O ponto mais característico do modelo é a introdução de um adicional fator de morte que representa as formas de atuação da seleção natural em um ecossistema real. Este fator é caracterizado por três elementos diferentes. O primeiro, dependente da densidade, é responsável por limitar o número total

da população, inspirando-se num crescimento logístico. Isso quer dizer que a probabilidade de morte (\mathcal{M}) é proporcional ao tamanho total da população a cada passo temporal, dividido por um parâmetro K que controla a capacidade de um determinado ecossistema em sustentar indivíduos (capacidade de carga):

$$\mathcal{M} \propto \frac{\sum_{y=0}^{32} N_y}{K} \quad (4.1)$$

onde com N_y indicamos o número de indivíduos com fenótipo y .

A esta primeira causa de morte junta-se um fator dependente das frequências que quer representar a competição entre indivíduos. Isso considera como, em situações reais, a tendência a ocupar regiões mais favorecidas no espaço dos fenótipos é contrastada por um aumento da competição entre indivíduos. Este componente dinâmico da seleção representa a retro-alimentação entre indivíduos e ecossistema e leva em consideração a incessante evolução do meio ambiente. Tal competição é codificada introduzindo na eq.4.1 uma dependência da distância fenotípica:

$$\mathcal{M} \propto \frac{\sum_{y=0}^{32} N_y \cdot \exp(-(x-y)^2/2C^2)}{K} \quad (4.2)$$

A competição diminui com a distância fenotípica seguindo uma função gaussiana com desvio C . x é o valor do fenótipo do indivíduo que está sujeito à pressão seletiva e a soma é feita sobre o índice y que varre todo o espaço fenotípico.

Finalmente, é implementada uma componente estática da seleção. Esta componente desenha as condições ambientais gerais do território e define o único nicho ecológico onde vivem os agentes. É causa de uma seleção direcionada que empurra constantemente a população em direção a um máximo do espaço dos fenótipos. Resumindo, a expressão completa para este fator de morte, como se apresenta na simulação de Monte Carlo, é:

$$\mathcal{M} = \frac{\sum_{y=0}^{32} N_y \cdot \exp(-(x-y)^2/2C^2)}{K \cdot \exp(-(x-16)^2/2S^2)} \quad (4.3)$$

O componente estático da paisagem adaptativa é representado por uma gaussiana com desvio S e direciona a população para um valor adaptativo máximo igual a 16. Na simulação, a cada passo temporal, é sorteado um número aleatório entre 0 e 1; o indivíduo sobrevive se este valor é maior que o valor da probabilidade de morte \mathcal{M} .

4.2.2 Acasalamento preferencial positivo

O acasalamento preferencial (*assortative mating*) tem um papel relevante para a obtenção da especiação. No nosso modelo este tipo de acasalamento é obtido permitindo o acoplamento só no caso em que a distância entre os dois parceiros, no espaço dos fenótipos, seja inferior (preferencial positivo) a um valor predefinido, controlado pelo parâmetro A [78].

Como já mencionado, no caso do modelo de Dieckman-Doebeli, o acasalamento preferencial é obtido sem custos seletivos. Isso significa que é definido um mecanismo que, normalizando as probabilidade de reprodução, garante um igual sucesso reprodutivo para todos os indivíduos. Esta atenção é devida ao fato que o acasalamento preferencial pode causar uma seleção sexual desfavorável aos fenótipos raros, que têm uma possibilidade menor de encontrar um companheiro. Este produto secundário do acasalamento preferencial tem um efeito estabilizante sobre as distribuições dos fenótipos e chega a causar maior dificuldade no surgimento de eventos de especiação [77]. Pelo fato que este efeito secundário pode estar presente na natureza e sendo que, ao contrário, um mecanismo de normalização das probabilidade reprodutivas é biologicamente justificável só no caso de populações estritamente monogâmicas, no nosso modelo não implementamos este mecanismo.

Por fim, não consideramos o problema da evolução natural do acasalamento preferencial [76]; ao contrário, por razões de simplicidade, assumimos a regra de acasalamento agora apresentada como fixa e diretamente ativada ao princípio da simulação. Mesmo com este mecanismo esquemático, a procura dos parâmetros que permitem obter especiação ainda permanece um problema difícil.

K	100000
População inicial	1000
Idade mínima de reprodução	6
Idade máxima de reprodução	32
Número de filhos	2
Máximo valor de doenças ativas	3
Probabilidade de mutação na tira estruturada por idade	1
Probabilidade de mutação na tira do traço fenotípico	0.01

Tabela 4.1: Principais parâmetros usados nas simulações de populações assexuadas.

4.3 Resultados

4.3.1 Populações assexuadas

Começamos a apresentação dos nossos resultados mostrando alguns dados obtidos com uma população assexuada. Neste caso mais simples, a população evolui rapidamente na direção do fenótipo com valor adaptativo maior (16), que corresponde ao valor da capacidade de carga maior. Uma vez alcançado este valor, duas evoluções posteriores são igualmente possíveis.

Na primeira, o valor 16 é um ponto fixo no espaço dos fenótipos e os traços fenotípicos da população se estruturam em uma distribuição gaussiana centrada neste valor. Na segunda, o valor 16 ainda não é um máximo do valor adaptativo (como na figura 4.1) e a população fica sujeita a uma seleção disruptiva que causa o aparecimento de dois distintos conjuntos fenotípicos. Neste último caso, a tendência a evitar uma região superpopulada do espaço fenotípico é mais forte que a vantagem obtida alcançando-se o valor do fenótipo ótimo. Este fenômeno acontece quando $C < S$, como pode ser apreciado na tabela 4.2.

Este não é o único comportamento possível: para alguns valores suficientemente pequenos de C , a população divide a sua distribuição fenotípica antes que o seu valor médio alcance o fenótipo com maior valor adaptativo. Um exemplo deste comportamento é ilustrado, no caso de populações sexu-

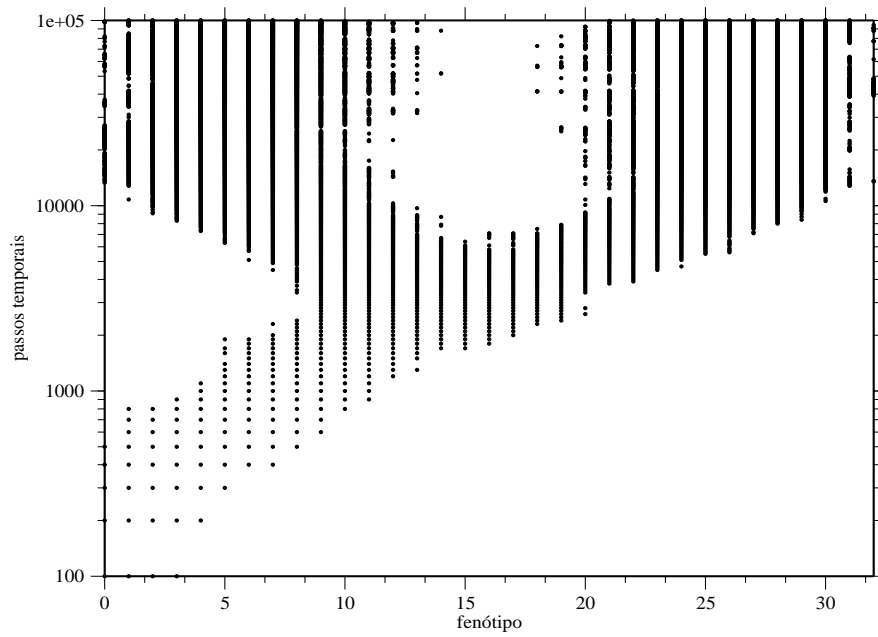


Figura 4.1: Evolução temporal da distribuição fenotípica de uma população assexuada. Os valores do desvio padrão da competição (C) e da componente estática (S) são, respectivamente, 20 e 25.

	40	1	1	1	1	1	1	1
C	35	1	1	1	1	1	1	2
	30	1	1	1	1	1	2	2
	25	1	1	1	1	2	2	2
	20	1	1	1	2	2	2	2
	15	1	1	2	3	3	3	2
	10	1	3	3	3	3	3	3
	5	x	x	x	x	x	x	x
		10	15	20	25	30	35	40
							S	

Tabela 4.2: Geração de polimorfismos variando os valores do desvio padrão da competição (C) e da componente estática (S). Cada número na tabela corresponde ao número de picos isolados da distribuição fenotípica, enquanto uma cruz indica a ocorrência de polimorfismos instáveis, como aquele apresentado na figura 4.2. Nesta tabela mostramos os resultados produzidos por uma única simulação, para cada conjunto de parâmetros. Em realidade realizamos numerosas simulações geradas a partir de diferentes condições iniciais. Em todos os casos a definição de dois morfos fenotípicos diferentes era possível só no caso $C < S$, e a dependência geral dos resultados nos parâmetros é parecida, embora haja uma grande variação no número efetivo de picos, sejam isolados ou não. Os dados são recolhidos depois de 100000 passos temporais.

adas, na figura 4.6. Este particular desenvolvimento temporal é relacionado com uma influência da competição mais forte que causa bifurcações prematuras. Efetivamente, a força da competição aumenta ao diminuir C , como se evidencia na tabela 4.2, onde, para tais valores, aumentam os eventos de ramificação. Por exemplo, para $C = 10$, em apenas 100000 passos temporais a população experimenta dois eventos de bifurcação, cada um causando a formação de dois distintos aglomerados fenotípicos, que geram uma distribuição final de três modos. Aumentando ainda mais a competição, a população sofre uma interação tão forte que é impossível obter um polimorfismo estável: os eventos de ramificação são tão numerosos que a distribuição fenotípica resulta instável, caracterizada por um grande número de picos conectados por fenótipos intermediários. Uma apresentação destas diferentes distribuições finais é visível na figura 4.2.

Todos estes resultados são coerentes com aqueles obtidos na referência [76], mas o fato do fenótipo ser determinado através de um genoma governado por uma dinâmica microscópica determina uma fenomenologia bem mais rica. O resultado destes autores, que prevê bifurcações para $C < S$, é confirmado nos nossos estudos, onde podemos também mostrar como as ramificações podem acontecer repetidamente, levando a estados polimórficos com mais de dois aglomerados fenotípicos, ou até a comportamentos mais complexos, como refletido no desenvolvimento temporal mais rico do nosso modelo. Um outro fato de interesse é que dados obtidos com simulações muito mais demoradas mostram como a estrutura aglomerada dos fenótipos da população é biologicamente relevante. De fato, também depois de 10^6 passos de Monte Carlo o polimorfismo permanece (ver figura 4.3). Por esta razão, embora não seja possível determinar a solução para a distribuição de equilíbrio final por meio de simulações numéricas, uma estrutura tão longa não pode ser considerada uma simples fase transiente, como afirmado em [79], mas pelo menos um tipo de estado quase estacionário com uma escala temporal comparável com aquelas de interesse biológico.

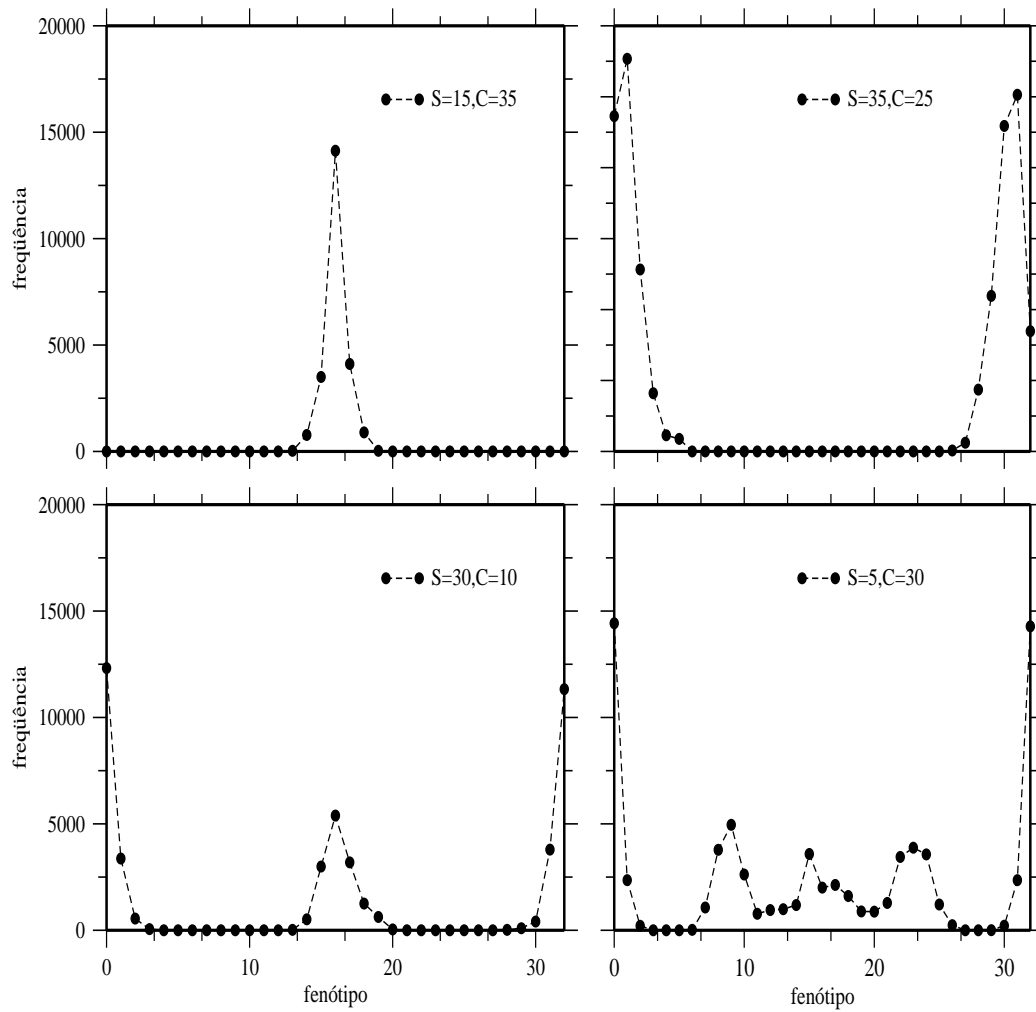


Figura 4.2: Os principais tipos de distribuições finais relativas ao resultados apresentados na tabela 4.2.

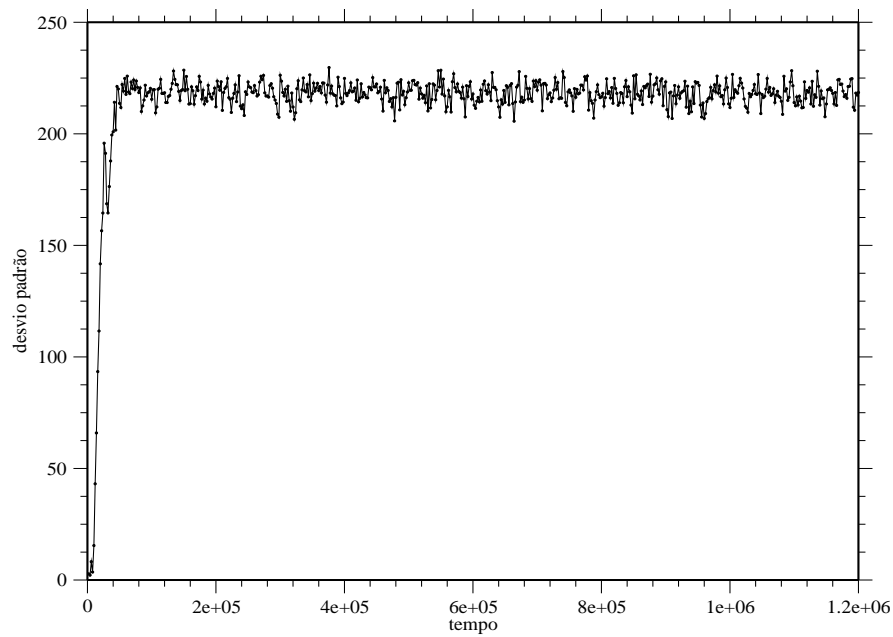


Figura 4.3: O desvio padrão da distribuição dos fenótipos para uma simulação demorada. Esta grandeza pode ser considerada um bom parâmetro de ordem do sistema [78]. O aparecimento de uma distribuição fenotípica com muitos picos corresponde a um pulso repentino no valor do desvio padrão. Um evidente sinal de que a estrutura de aglomerados fenotípicos é mantida vem do fato de que, depois da primeira variação relevante, perto de $t = 30000$, o valor médio do desvio não muda mais.

K	100000
População inicial	1000
Idade mínima de reprodução	8
Idade máxima de reprodução	32
Número de filhos	4
Máximo valor de doenças ativas	3
Probabilidade de mutação na tira estruturada por idade	1
Probabilidade de mutação na tira do traço fenotípico	0.01
Locos dominantes	16

Tabela 4.3: Principais parâmetros usados nas simulações de populações sexuadas. O número de locos dominantes é igual nas duas tiras de bits.

4.3.2 Populações sexuadas

A evolução de populações sexuadas e diplóides é condicionada por uma variabilidade maior, gerada pelos efeitos da recombinação e da segregação Mendeliana. Este fato é motivo de ulteriores dificuldades em obter diferenciações fenotípicas distintas por causa da constante geração de fenótipos intermediários. Por essa razão, se o processo de acasalamento é puramente aleatório não é mais possível obter uma divisão estável da distribuição em dois diferentes grupos fenotípicos, qualquer que seja o valor de S e de C . Pelo contrário, introduzindo um acasalamento preferencial, não só é possível obter diferentes morfos fenotípicos como também efetiva especiação, com populações sexualmente isoladas, como representado na figura 4.4.

O espaço dos parâmetros onde tais eventos são possíveis resulta mais confinado que no caso assexuado. A condição $C < S$, embora seja ainda necessária, não é mais suficiente. Das numerosas simulações efetuadas foi possível obter especiação para aquelas com os parâmetros incluídos nestas regiões (ver figura 4.5):

$$S \geq 15, 6 \leq C \leq 8, A = 4;$$

$$S = 40, C = 6, 3 \leq A \leq 5$$

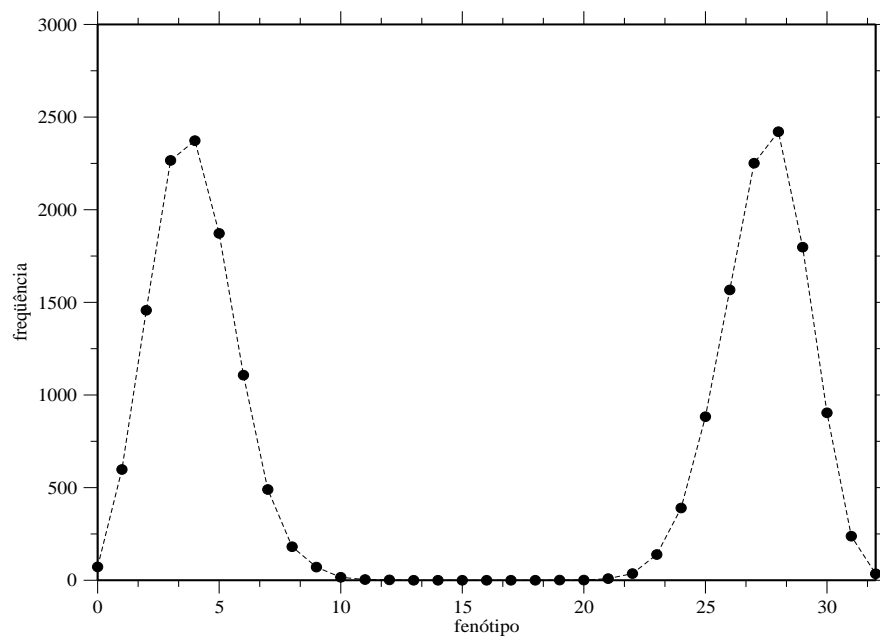


Figura 4.4: Um exemplo de especiação em uma população sexuada. A figura mostra a distribuição fenotípica depois de 100000 passos temporais ($S = 40$, $C = 8$, $A = 4$).

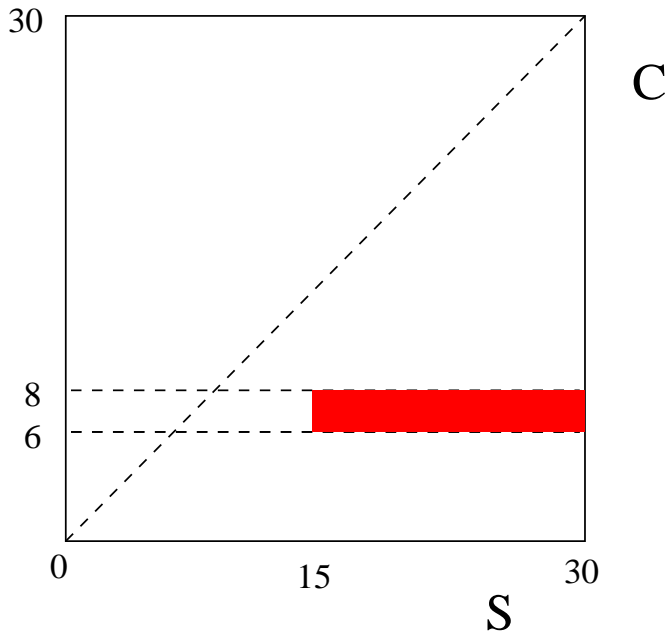


Figura 4.5: Diagrama de fases para $A = 4$. É evidenciada a região onde é possível obter especiação.

As simulações mostram como o sistema não é muito sensível ao valor de S . Se tal valor é suficientemente maior que um particular valor de C a população foge da região com o máximo da capacidade de carga. Por outro lado, o valor de C não pode ser nem pequeno demais, caso no qual levaria a uma distribuição caracterizada por picos conexos e não bem definidos, nem grande demais, no qual a competição não resultaria suficientemente forte por evitar uma distribuição gaussiana centrada em 16. Por fim, valores de A pequenos demais podem limitar totalmente a variabilidade do fenótipo em uma região confinada perto do 0 ou até prevenir a especiação, por causa dos efeitos penalizadores da seleção sexual sobre fenótipos raros [77]. Para valores maiores, ao contrário, a escolha de acasalamento não é suficientemente severa para impedir o aparecimento de fenótipos intermediários. Em qualquer caso, também neste cenário mais restrito, é ainda possível encontrar situações particularmente favoráveis nas quais dois eventos de especiação já aconteceram só depois de 100000 passos temporais (ver figura 4.6).

Em todos estes casos os nossos resultados são comparáveis com aqueles

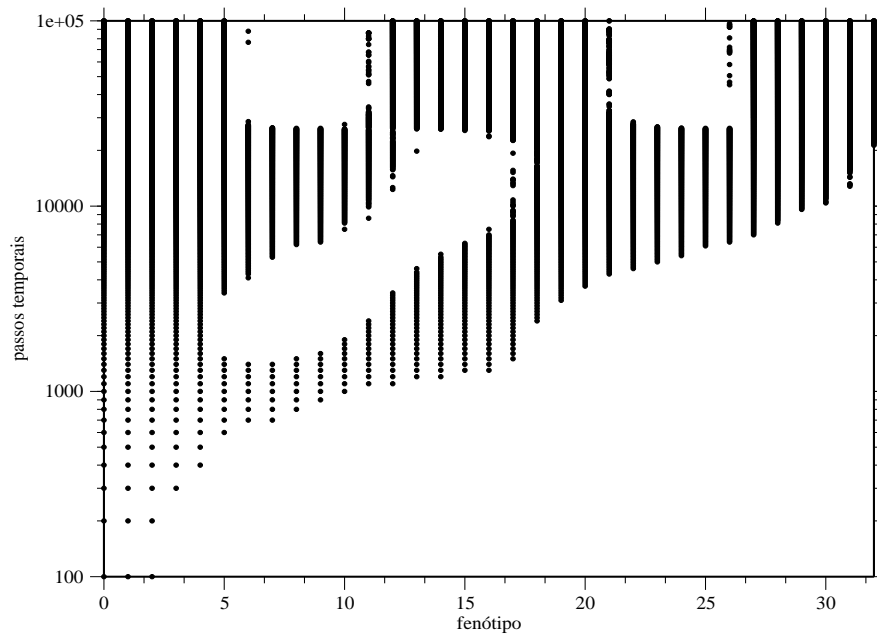


Figura 4.6: Uma situação particular onde é possível obter rapidamente dois eventos de especiação ($S = 40$, $C = 6$, $A = 4$).

obtidos por Dieckmann e Doebeli [76], mas os parâmetros necessários para obter especiação resultam mais restritivos.

Em relação às críticas do artigo de Polechova e Barton [79], que levantou dúvidas sobre a robustez destes resultados no caso do uso de um intervalo de valores fenotípicos maiores, realizamos uma série de simulações caracterizadas por um traço fenotípico grande. Uma vez que, para modelos de até 160 locos, não surgiram diferenças, é lícito supor que os resultados obtidos não são dependentes do intervalo de valores do fenótipo. A figura 4.7 representa a distribuição final de um modelo a 160 locos, onde é possível notar a geração de 5 novas espécies distintas.

Como última curiosidade, em relação à dependência das condições iniciais, podemos afirmar que, em geral, começando com populações com um fenótipo perto de 16, é mais fácil obter especiação.

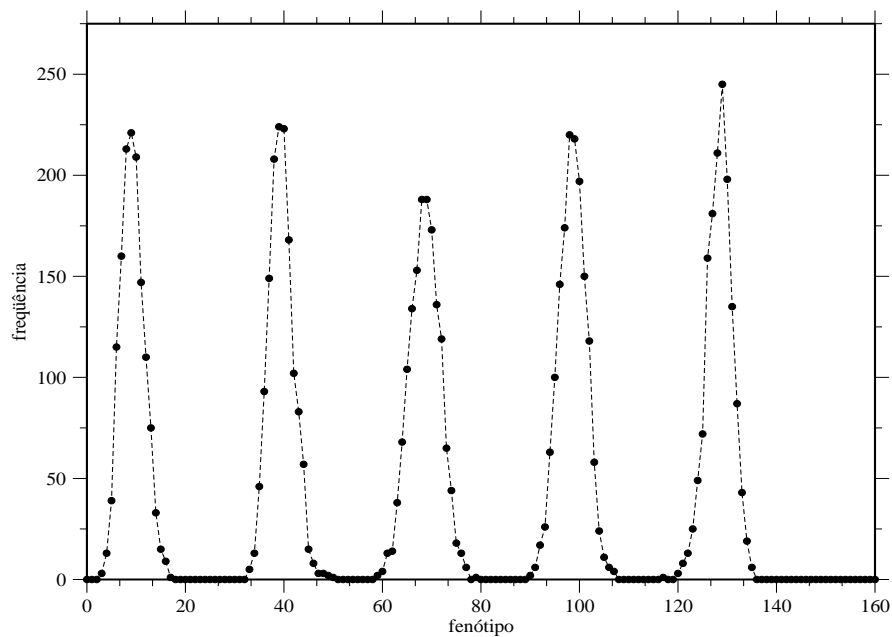


Figura 4.7: Uma simulação com uma tira de bits, relativa ao traço fenotípico, de 160 locos ($S = 160$, $C = 8$, $A = 4$). Este resultado sugere que não existe uma dependência entre especiação e intervalo de valores do fenótipo, não obstante uma recente afirmação em contrário [79].

4.4 Considerações finais

A representação de um caráter fenotípico individual usando uma segunda tira de bits é uma extensão natural do clássico modelo Penna que permite o estudo da especiação simpátrica. O nosso modelo, estruturado seguindo estas idéias, corrobora a hipótese de que a especiação simpátrica pode emergir por competição de recursos naturais distribuídos de forma contínua. Com o uso do mesmo modelo testamos também a robustez desta hipótese quanto à inclusão de outros ingredientes, como o envelhecimento biológico, e a dependência dos resultados nos valores de algumas variáveis constitutivas do modelo.

Estudando uma população estruturada por idade concluímos que os resultados precedentes não são invalidados pela possível existência de indivíduos extremamente longevos. O efeito do envelhecimento, que é presente em qualquer população, não parece causar complicações nos processos estudados e, pelo contrário, chega até a contribuir para acelerar a dinâmica evolutiva.

Na população assexuada a introdução de um fenótipo determinado por um genótipo de 32 locos confirma os resultados da referência [76] num cenário enriquecido. Ademais, o fato de que a estrutura aglomerada seja ainda presente depois de mais que 10^6 passos temporais, determina a relevância biológica desta demorada fase do sistema.

A especiação simpátrica, no sentido estrito do termo, é obtida por uma população sexuada onde foram implementados os mecanismos de cruzamento, dominância e locos homo/heterozigotos. Os nossos resultados, obtidos com um modelo que não inclui mecanismos de normalização da probabilidade de reprodução, são resistentes aos efeitos devidos à seleção sexual causada pelo acasalamento preferencial. Por fim, a agilidade do modelo Penna na manipulação das tiras de bits permitiu o cálculo de simulações com um intervalo do fenótipo grande, com resultados que sugerem a independência entre o fenômeno da especiação e o intervalo fenotípico, não obstante recentes afirmações em contrário [79].

De uma perspectiva mais geral, este estudo quer outra vez afirmar a opinião de que os modelos baseados em agentes e que admitem flutuações, como o da referência [76] e o nosso, são os testes mais promissores para problemas

evolucionários em dinâmicas populacionais. O efeito fundamental das flutuações na criação e destruição de estados do sistema que apresentam ordem e simetrias diferentes não permite tratar as flutuações como simples perturbações e é o motivo que causa a pobreza de valor preditivo dos resultados obtidos por vias deterministas em relação ao comportamento seja dos modelos estocásticos, seja da realidade biológica.

Capítulo 5

Especiaçãoção e macroevolução: micro e macroevolução são desacopladas ?

“io sono una giustizia che non ho da distinguere, non ho da far differenze; ... e però, quando dono, non vedo a chi dono; quando toglio, non vedo a chi toglio: acciò che in questo modo io vegna a trattar tutti equalmente e senza differenza alcuna.”

Giordano Bruno, *Spaccio de la bestia trionfante*

5.1 Os modos microevolucionários são extrapoláveis a escalas geológicas?

A interpretação original que Gould e Eldredge [80] deram, com os seus trabalhos dos anos 70, à antigas observações paleontológicas tiveram um profundo impacto, não só sobre os estudos mais próximos a esta área da biologia, mas diretamente sobre alguns aspectos fundamentais da teoria da evolução. A teoria do equilíbrio pontuado causou uma mudança nos paradigmas com os quais os dados paleontológicos eram analisados e também uma mudança radical nas formas de pensar em biologia. Um das mais interessantes con-

seqüências de alguns conceitos desenvolvidos nesses trabalhos pioneiros, foi a afirmação da idéia de que seria necessário dissociar os mecanismos microevolucionários dos macroevolucionários.

O equilíbrio pontuado baseia-se no reconhecimento do fato de que “*stasis is data*”. Como consequência deste fato, que responsabiliza pelos períodos de *stasis* nos registros fósseis não uma perda de informação mas o *modo* real da evolução, a dinâmica evolutiva vem sendo descrita por um processo caracterizado por longos intervalos de *stasis*, sem atividade, e abruptas eclosões de atividade, na forma de eventos de especiação ou extinção, não tratando-se, pois, de uma dinâmica puramente gradualista. Ou seja, tal interpretação reconhece como algumas mudanças evolutivas podem ser produzidas mais pelo sucesso de algumas espécies em relação a outras que de uma transformação direcional e progressiva das linhagens (ver figura 5.1).

Como já discutido no parágrafo 1.1, a teoria evolutiva de Darwin interpretava o mundo vivo como o produto de uma única dinâmica microscópica, guiada por seleção e mutações, supondo uma causalidade nas mudanças evolutivas direcionada pela seleção natural que opera unicamente ao nível individual. Além disso, tal causalidade vinha naturalmente extrapolada a todos os tempos, esperando que, sozinha, a seleção natural Darwiniana pudesse explicar totalmente as mudanças em grande escala da história da vida [4, 81].

Em contraste, o equilíbrio pontuado sugere que a interposição de níveis quebra esta redução causal e dissocia a evolução microscópica da macroscópica. A causalidade é suprimida por um modelo hierárquico com ações simultâneas ao nível genético, dos organismos e das espécies.

Desta perspectiva, o problema central da macroevolução é entender, através de um estudo direto das espécies, quais sobrevivem e fazem melhor que as outras. Ao contrário, a abordagem tradicional era focada em como a seleção natural fora capaz de produzir adaptação em um processo em nível de população, ou seja indivíduos, gerando uma evolução puramente gradualística.

5.2 Estudos precedentes

O primeiro e famoso modelo capaz de sugerir como os fenômenos evolutivos podem ser caracterizados por um comportamento de equilíbrio pontuado e

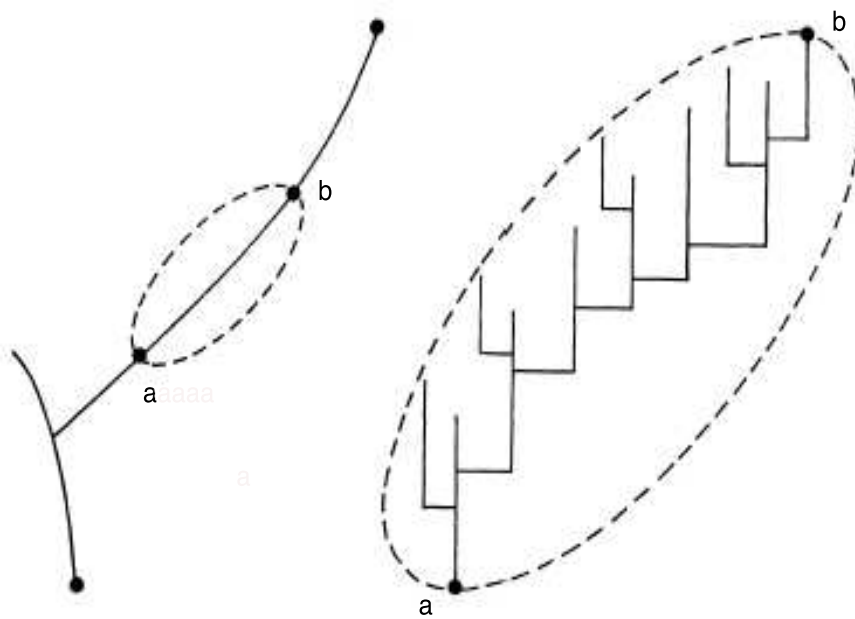


Figura 5.1: Na escala apropriada um aparente gradualismo, como no intervalo entre a espécie ancestral a e o descendente b , pode ser visto como um processo de eventos pontuados. Característica deste tipo de dinâmica são os períodos de *stasis* e as ramificações sem direções privilegiadas. Esta hipótese é confirmada por dados fósseis que apresentam ramificações pontuadas com a sobrevivência do ancestral. Se as espécies fossem simplesmente transformadas umas nas outras não existiria coexistência entre as formas antigas e as novas.

que lançou a hipótese de que fossem organizados em um estado crítico é o modelo de Bak-Sneppen [83]. Este modelo realiza uma análise a longo termo dos processos evolutivos baseando-se na idéia de uma macroevolução, onde os agentes são espécies e só as interações interespecíficas são consideradas, com a hipótese de que o número de espécies seja fixo e a originação de novas proibida.

Na nossa abordagem, tentaremos descrever o mecanismo completo da evolução por meio da seleção natural em um nível populacional, ou seja dos indivíduos, como na visão do Darwinismo mais tradicional. Modelos com dinâmicas estruturadas em um nível populacional [84, 85] ou enfocados nas relações entre micro e macroevolução [91], também com uma estrutura complexa que representa a organização hierárquica em diferentes níveis tróficos [93], estão presentes na literatura. Apesar disso, são todos baseados em uma interação de tipo presa-predador, que pode unicamente descrever as dinâmicas de extinção, e, não apresentando um mecanismo de originação de novas espécies, não exploram o problema especificado pela teoria do equilíbrio pontuado.

Ao contrário, nós queremos analisar um modelo onde a interação representa a seleção natural responsável pelos eventos de especiação. O nosso fim é testar se com apenas esta força podemos descrever toda a fenomenologia que caracteriza a macroevolução. O objetivo é definir um mecanismo que gere naturalmente uma seleção disruptiva e cause autonomamente a diferenciação no interior de uma única espécie. Por querer um mecanismo baseado unicamente na seleção, não estamos interessados em simples modelos estocásticos. Ao contrário, implementamos uma força seletiva [76] que se auto modifica, baseada em uma seleção dependente da frequência, que permite a coexistência e a ramificação de diferentes espécies. Podemos resumir dizendo que não consideramos o valor adaptativo de espécies inteiras, mas um mecanismo que gera espécies autonomamente. É por esta razão que não realizamos um simples preenchimento das espécies extintas (como por exemplo nas Ref.[83, 84, 88, 91]) e, em conseqüência, o número delas não é fixo. Com esta estratégia podemos também testar se é possível construir um modelo simples que unifique em um único cenário adaptação e diferenciação.

Esta abordagem unifica também as três escalas temporais [86] que carac-

terizam a evolução: a escala rápida da dinâmica de populações (controlada pela seleção natural), a escala lenta da evolução (controlada pelo processo de mutação) e a escala ultra lenta da macroevolução (a escala temporal da dialética especiação/extinção). Embora não exista nenhuma dúvida de que os eventos de especiação e extinção aconteçam por interação entre seleção natural e mutação (escalas rápida e lenta), como já apontado, não existe um acordo geral se a macroevolução pode ser vista como a simples soma de eventos discretos de especiação gerados pela dinâmica de população.

Enfrentaremos esta questão comparando os resultados do nosso modelo, nascido desta visão unificadora da evolução, com as propriedades estatísticas quantitativas observadas no registro fóssil: comportamento livre de escala, em alguns intervalos, de algumas distribuições [87], uma série temporal de eventos de extinção que indica equilíbrio pontuado [81], extinções de massa [82] e correlação a longo prazo [90]. Veremos que, embora seja reproduzida a natureza livre de escala do fenômeno, não emergem em nossos resultados todas as características conexas ao equilíbrio pontuado.

5.3 O modelo

O modelo, por razões de simplicidade, não é desenvolvido no espaço dos genótipos, como nos modelos dos capítulos precedentes ou como, por exemplo, em relação a estes problemas de macroevolução, nas referências [84, 85], mas no espaço mais manejável das estratégias, onde com estratégias referimo-nos ao termo como usado na Teoria dos Jogos. Em tal espaço, um indivíduo é representado por um número inteiro, sua estratégia x ($0 \leq x \leq P$), que descreve todas as características fenotípicas que determinam o seu sucesso biológico. A cada passo temporal, um indivíduo gera uma descendência com a mesma estratégia do seu ancestral, eventualmente mutada por um fator aleatório ± 1 com uma probabilidade μ , definida e fixa desde o começo da simulação. Cada agente continua a viver até o acontecimento da morte causada por seleção, implementada utilizando o mesmo mecanismo descrito no capítulo 3, com a única diferença da ausência da componente estática que causava a seleção direcionada. Permanece então só a componente dependente da densidade, responsável por limitar o tamanho da população, e uma de-

pendente das frequências, que descreve o fato de que a tendência a ocupar regiões mais favorecidas no espaço das estratégias é balanceada por um aumento da competitividade entre indivíduos. Para resumir, a seleção natural é implementada através de uma probabilidade de morte que tem a forma:

$$\mathcal{M} = \frac{1}{K} \cdot \sum_{y=0}^P N_y \cdot \exp\left(-\frac{(x-y)^2}{2C^2}\right) \quad (5.1)$$

A cada passo temporal é sorteado um número aleatório entre 0 e 1; o indivíduo sobrevive se o número é maior que \mathcal{M} . A intensidade da competição diminui com a distância no espaço das estratégias em função do desvio C , e o parâmetro K controla a capacidade de carga do sistema. x é o valor da estratégia do indivíduo que está sofrendo a pressão seletiva, e a soma é feita sobre o índice y que mapeia o espaço inteiro das estratégias. Com N_y indicamos o número de indivíduos com estratégia y .

Sem dúvida, a seleção apresentada em 5.1 causa alguns efeitos de contorno na distribuição da população, pelo fato que os agentes encontram uma competição menor nas bordas do espaço das estratégias. O efeito mais relevante é a geração de picos sobrepovoados próximos aos extremos dos valores da estratégia. Para evitar estas conseqüências, que de alguma forma poderiam influenciar a dinâmica dos eventos de especiação/extinção, usamos condições de contorno periódicas, que foram implementadas da seguinte forma:

$$\left\{ \begin{array}{l} \mathcal{M} = \frac{1}{K} \cdot \sum_{y=0}^P N_y \cdot \exp\left(-\frac{\Psi^2}{2C^2}\right) \\ \Psi = \min \left\{ \text{mod}_P [|x - y + P|], \text{mod}_P [|y - x + P|] \right\} \end{array} \right\} \quad (5.2)$$

Esta função de seleção é diretamente modificada pelo processo evolutivo e pode produzir a auto organização de um número variável de diferentes aglomerados fenotípicos, permitindo a ramificação e extinção dos mesmos. Por esta razão, não leva o sistema a um ponto final optimal, mas o deixa em um estado dinâmico de mudança permanente.

Pelo fato de estar representando uma população assexuada, a caracterização biológica de espécie, definida para indivíduos sexuais como uma população reprodutivamente isolada, tem que ser substituída por uma definição

mais operacional, baseada em uma diferenciação funcional entre grupos fenotipicamente distintos. Por esta razão, referimo-nos a uma espécie como um grupo de indivíduos que compartilha a maioria das características fenotípicas mas difere por alguns traços. De acordo com esta definição, o algoritmo usado associa diferentes espécie a diferentes aglomerados de indivíduos que apresentam uma pequena distância no espaço das estratégias - uma unidade já é suficiente - entre eles. Ou seja, o espaço entre dois aglomerados não pode ser ocupado por indivíduos.

Embora heterogeneidades espaciais ou a predação possam ter uma influência relevante nas dinâmicas da população e, em particular, nas frequências dos eventos de ramificação, não são consideradas no modelo. Porém, a inclusão de uma seleção estática que define as condições gerais da ecologia e pode causar uma seleção direcionada, como aquela da equação 4.3, não muda crucialmente a dinâmica do modelo.

5.3.1 Dependência dos parâmetros

A dependência do modelo no valor de alguns parâmetros fundamentais pode ser resumida em alguns comportamentos simples. Em geral, os parâmetros de interesse são só dois, e são aqueles que efetivamente controlam o processo de ramificação. Ao contrário, a capacidade de carga (K) e o número de estratégias possíveis (P) não são cruciais: o primeiro só regula o tamanho da população e o valor de P é correlacionado com o número médio de espécies viventes, que cresce com P seguindo uma relação linear simples.

Em contraste, a taxa de mutação (μ) influencia diretamente a probabilidade de ramificação. Este parâmetro controla a variabilidade de estratégias da população, e por isso dele depende a largura da distribuição das estratégias e a rapidez do processo evolutivo. Fixaremos o valor de μ com a finalidade de permitir uma simulação realista, onde as mutações ocorrem raramente, em acordo com o fato de que este parâmetro controla a dinâmica lenta na escala temporal de um processo evolutivo. Por esta razão, escolhemos utilizar o valor $\mu = 0.005$.

Ainda sobrevive um parâmetro, C , cujo valor é responsável por controlar a competição e, em consequência, a força que separa um aglomerado em dois

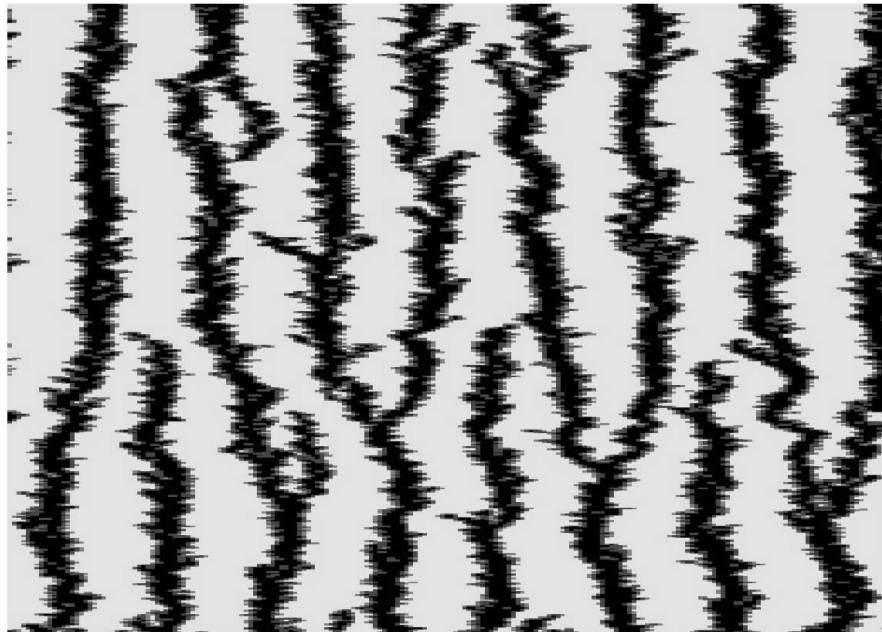


Figura 5.2: Evolução temporal da população. O eixo horizontal representa o espaço das estratégias, o vertical o tempo. Os parâmetros usados são: $K = 10000$, $\mu = 0.005$, $P = 200$ e $C = 15$.

diferentes. Esta deriva cresce inversamente com C , causando o aparecimento de um número maior de eventos de ramificação. Existe uma relação simples entre o número médio de espécies e o valor de C :

$$N(C) \propto C^{-1}$$

Ajustamos então este último parâmetro procurando um equilíbrio entre uma dinâmica de bifurcações muito lenta, que é estabelecida por valores grandes de C e não é apropriada a uma análise estatística, e pequenos valores de C , com os quais a população é submetida a uma seleção tão forte que é impossível definir uma árvore evolutiva estável. Nessa última situação, de fato, os eventos de ramificação são tão numerosos que a distribuição não pode ser bem definida, com um grande número de picos conexos por estratégias intermediárias, e não resulta possível fazer-se uma análise dos aglomerados de estratégias. Um exemplo de árvore evolutiva, gerada por valores padrão dos parâmetros, pode ser visto na Figura 5.2.

Por fim uma curiosidade. Implementamos também simulações com populações estruturadas por idade, da mesma forma que as populações descritas no capítulo precedente, querendo provar a robustez do nosso modelo à introdução deste elemento adicional. Não encontramos um comportamento claramente diferente nos dados assim obtidos.

5.4 Resultados

5.4.1 Análise estatística das distribuições de frequências

Deixamos rodar simulações muito demoradas com 10^7 passos temporais. Deses dados obtivemos, como primeiro resultado, a distribuição de frequências dos tempos de vida $E(t)$ das diferentes espécies. Trata-se de uma medida central que pode ser comparada com os resultados provenientes de dados observacionais. Isso porque é possível achar uma simples relação entre tempos de Monte Carlo e tempos reais: um tempo de Monte Carlo corresponde a um ciclo reprodutivo. Normalmente nos dados experimentais a unidade de

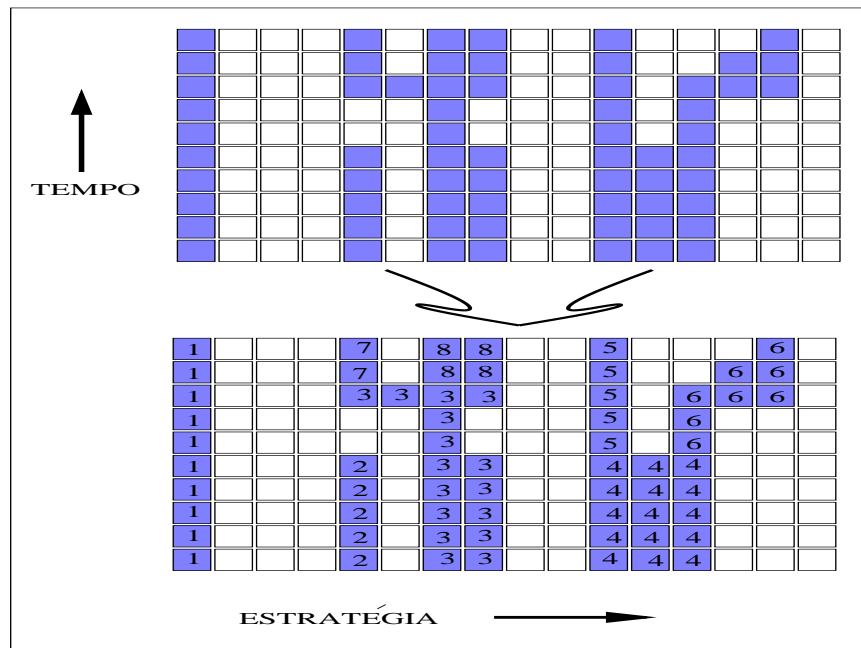


Figura 5.3: O algoritmo para o cálculo do tempo de vida dos aglomerados. A primeira imagem mostra a evolução temporal da distribuição de indivíduos com um determinado valor de estratégia. Na segunda, um número etiqueta a evolução de cada aglomerado.

tempo é o ano. Podemos então passar de um tempo medido em ciclos reprodutivos ($t_{c.r.}$) para um tempo medido em anos (t_a) através de uma simples dilatação: $t_{c.r.} \rightarrow \alpha t_a$. Depois desta transformação, não só podemos ainda distinguir entre o caráter exponencial ou de lei de potências das distribuições, mas também o valor dos expoentes das leis de potência resulta o mesmo.

Para efetuar esta análise estatística desenvolvemos um algoritmo simples de cálculo. No espaço das estratégias, substituímos os pontos ocupados por ao menos um indivíduo por um número e cada número identifica um aglomerado de estratégias. Este número se mantém com o mesmo valor até o acontecimento de uma ramificação. Nesse passo temporal, o número passa a ser igual ao maior número presente até aquele momento, aumentado por um (ver figura 5.3). Trata-se depois de contar a duração de um evento etiquetado

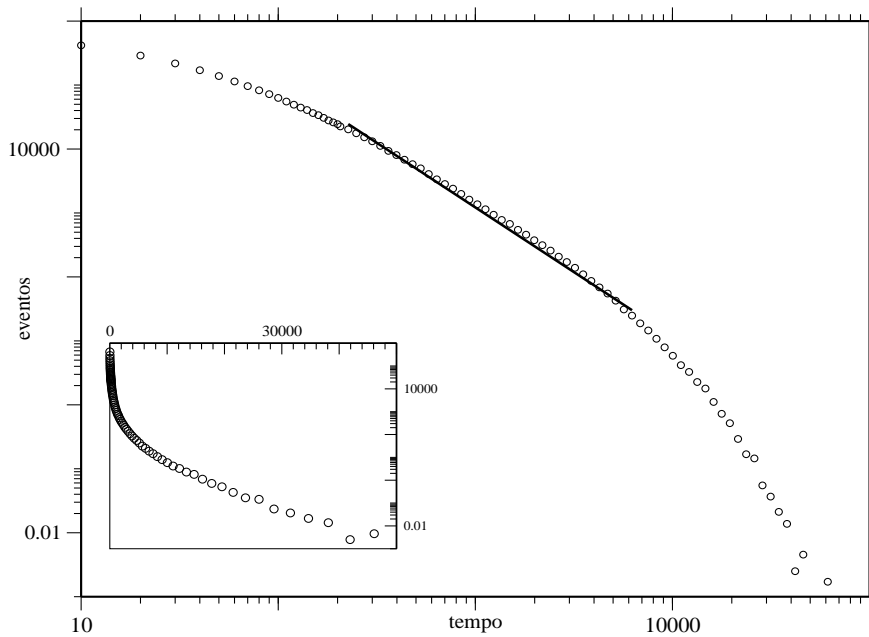


Figura 5.4: Distribuição em frequências do tempo de vida dos eventos de extinção, a partir da ramificação de uma nova espécie até a sua extinção. A figura menor mostra a distribuição em uma escala semi-logarítmica. Os parâmetros desta simulação e de todas as seguintes são: $K = 10000$, $\mu = 0.005$, $P = 200$ e $C = 15$.

pelo mesmo valor.

Dos dados mostrados na figura 5.4, é possível identificar uma lei de potência

$$E(t) \propto t^{-\gamma} ; \quad \gamma = 2.02 \pm 0.04$$

com duração de duas décadas.

O valor encontrado para o expoente γ é comparável com os registros dos eventos de extinção na natureza. Embora a interpretação desses dados esteja ainda em debate, uma lei de potência com um expoente -2 parece ser mais convincente que uma interpolação exponencial, ao menos para os tempos de vida mais breves [87, 8]. Um comportamento similar é confirmado por outros modelos: a referência [84] obtém o mesmo valor do expoente e as referências [91, 92] o desvio da lei de potência para tempos de vida muito demorados.

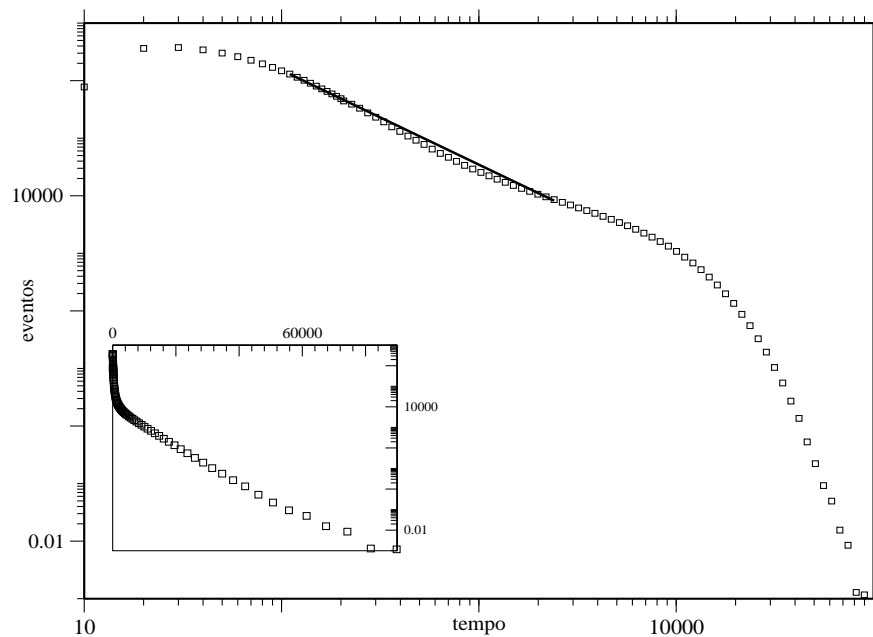


Figura 5.5: Distribuição em frequências dos intervalos de tempo entre dois eventos de especiação sucessivos. A figura menor mostra a cauda exponencial para valores grandes do tempo de vida.

Na figura 5.5, são apresentados dados relativos à distribuição de tempos de vida para o processo de originação de novas espécies.

O tempo de vida das originações representa o intervalo temporal entre um evento de especiação e o sucessivo, na mesma linhagem. Estes dados são mais difíceis de interpretar porque não existe uma análise estatística similar para os registros fósseis. Ademais, só modelos onde os eventos de especiação não coincidem com os eventos de extinção, como acontece nos modelos com preenchimento, mas são definidos por uma dinâmica interna, podem fornecer este tipo de dados. A distribuição dos tempos de vida das originações também fornece um comportamento na forma de lei de potência com expoente

$$\gamma_2 = 1.64 \pm 0.01.$$

Das nossas simulações obtivemos também a distribuição dos eventos de extinção como função do tamanho s . Com o termo tamanho consideramos o número de indivíduos que fizeram parte de uma espécie extinto desde sua

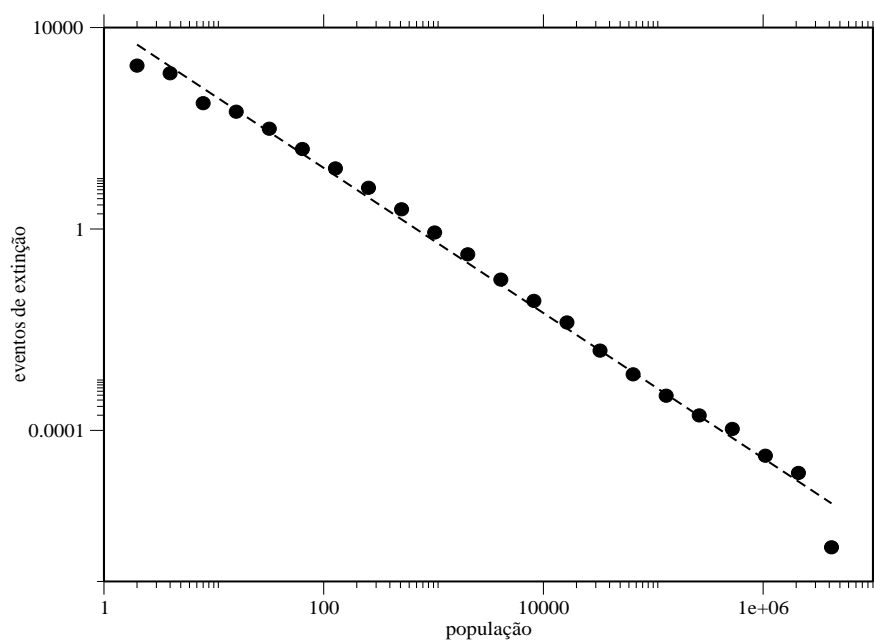


Figura 5.6: Distribuição em frequências por tamanho de extinção. Com tamanho entendemos o número total de indivíduos a partir do ponto de ramificação de uma espécie até sua extinção.

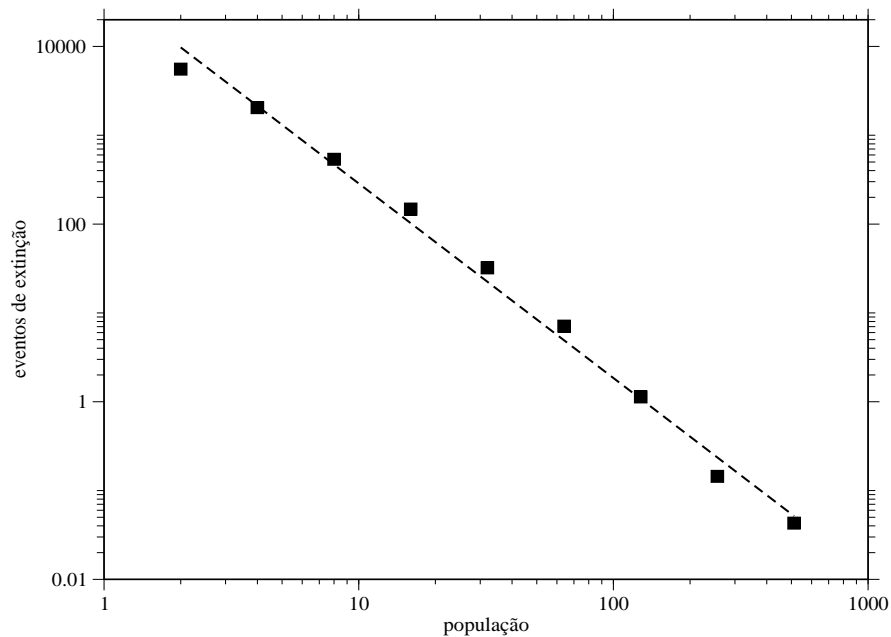


Figura 5.7: Distribuição em frequências dos tamanhos dos eventos de extinção. Neste caso, o número de indivíduos foi normalizado em função do tempo de duração da espécie relativo.

originação até o seu desaparecimento (ver figura 5.6). É possível modelar os dados com uma lei de potência:

$$E(s) \propto s^{-1.44 \pm 0.03}$$

É difícil comparar estes resultados com aqueles presentes na literatura [84, 88] porque usualmente o tamanho dos eventos é obtido contando-se o número de espécies ou famílias. Por fim, o estudo da distribuição dos tamanhos, mas com o número de indivíduos normalizados pelo tempo de vida da espécie (ver figura 5.7), também mostra uma lei de potência, mas com um expoente de valor -2.19 ± 0.07 .

5.4.2 Ausência de extinções de massa

Os resultados apresentados mostram um sistema caracterizado por distribuições que seguem um comportamento de leis de potência, ao menos por algumas décadas. Isto sugere que estamos lidando com um sistema que talvez

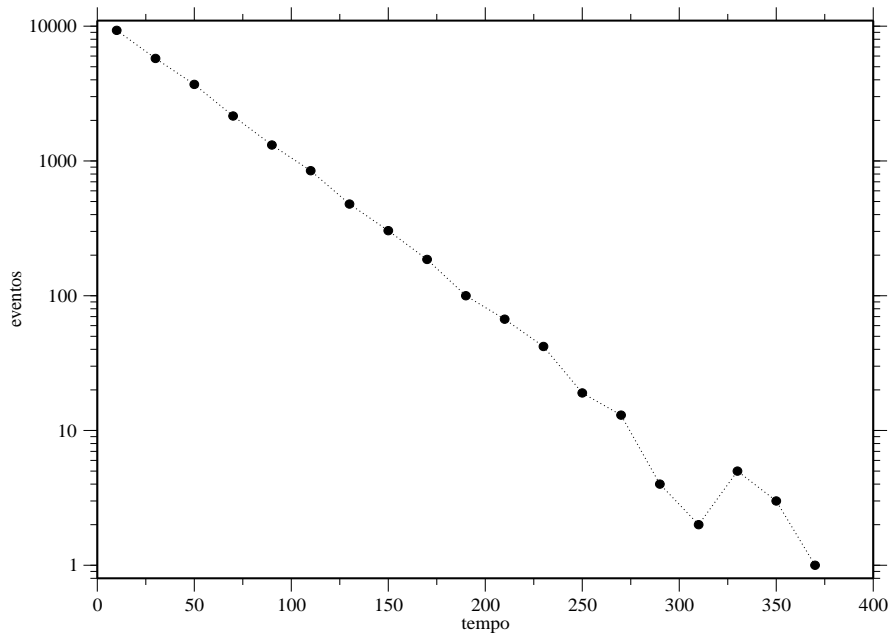


Figura 5.8: A distribuição da duração dos intervalos sem atividade (períodos de *stasis*) em escala semilogarítmica.

esteja situado num estado crítico, caracterizado por distribuições livres de escala no tempo e no espaço. Tudo isso estaria em acordo com alguns trabalhos que afirmam que a evolução biológica em uma ecologia de espécies interagentes evolui na direção de um estado crítico auto-organizado [83, 89].

Deixando de lado a questão de ser ou não auto-organizado, uma propriedade que usualmente discrimina de forma bastante eficaz sistemas puramente aleatórios, mas que conseguem exibir algumas propriedades de autosimilaridade (até fractalidade e ruído do tipo $1/f$), de sistemas em um estado crítico é a distribuição dos tempos de *stasis* [89].

Agora, analisando a distribuição das durações dos intervalos sem atividade (período de *stasis*) encontramos um comportamento claramente exponencial (ver figura 5.8), sugerindo, em consequência, que o sistema em exame não pode estar situado em um estado crítico.

Mais, a existência de uma escala, na forma da impossibilidade de descrever uma conexão contínua entre extinções pequenas, intermediárias e grandes, é facilmente percebida analisando-se a evolução temporal do número de

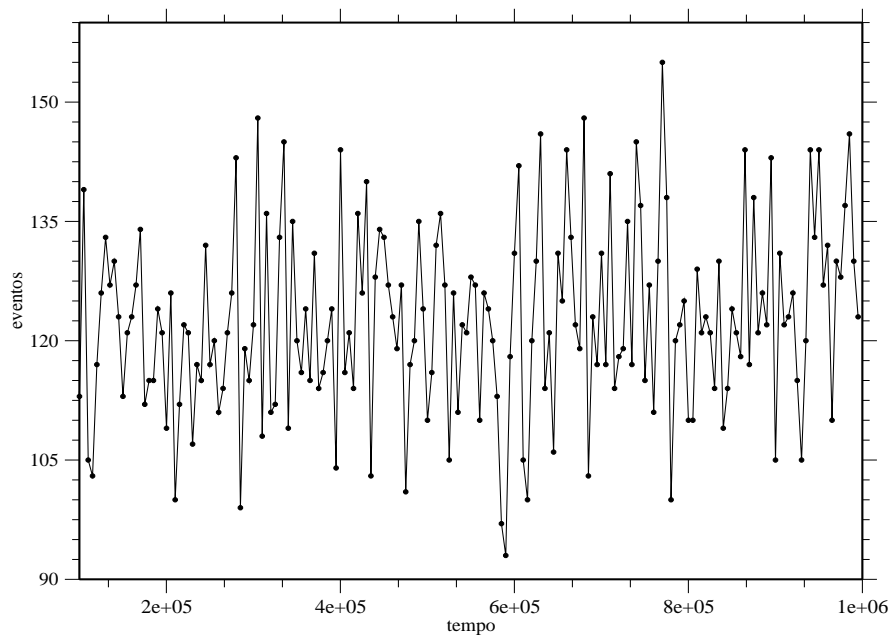


Figura 5.9: Evolução temporal do número de eventos de extinção coletando cada valor em um intervalo temporal de 5000 passos.

eventos de extinções em um dado intervalo temporal (fig.5.9). Não são localizáveis extinções de massa e a série temporal não apresenta o comportamento intermitente característico do fenômeno do equilíbrio pontuado.

Para finalizar, tentamos detectar a existência de correlações de longo alcance na série temporal dos eventos de extinção (figura 5.9). Em um processo caracterizado por correlações entre os diferentes dados da série temporal, seria esperada uma flutuação do tipo $F(t) \propto t^\alpha$ com $\alpha \neq 1/2$. Ademais, é possível relacionar α com o expoente que descreve o espectro de potência da série (β): $2\alpha = \beta + 1$. Por meio da análise de alguns dados paleontológicos [90], foi possível determinar com precisão o valor de β que resulta ser igual a 1. Este fato corresponde ao caso de flutuações autosimilares, descritas por um espectro de potência do tipo $1/f$.

Através do uso do método do DFA (*Detrended Fluctuation Analysis* [95]) analisamos várias séries temporais. Estas últimas foram escolhidas variando-se o valor do intervalo temporal de coleta dos eventos de extinção (entre 500 e 5000 passos) e o valor de μ ($0.03 \leq \mu \leq 0.07$). Todas as séries resulta-

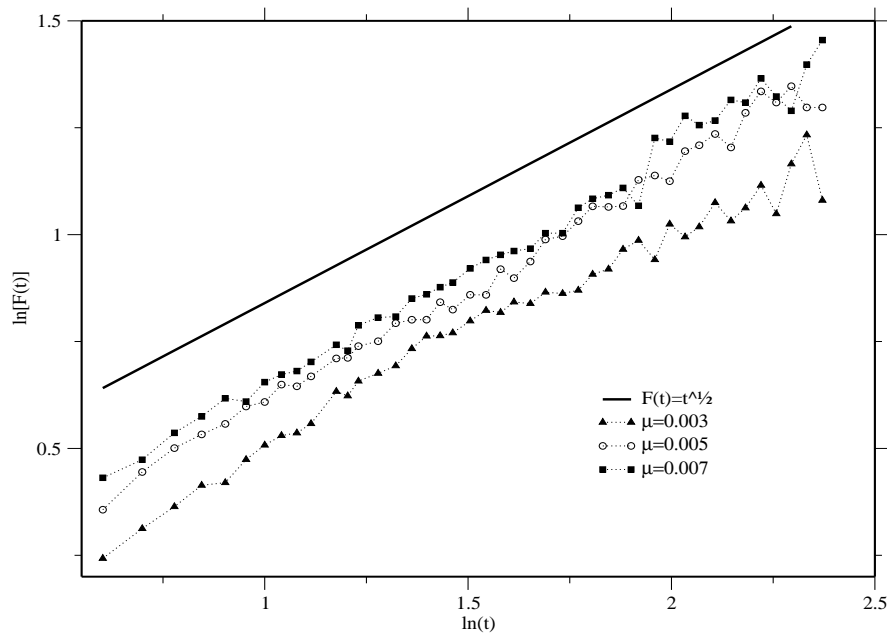


Figura 5.10: Cálculo do valor do expoente α para várias séries temporais de eventos de extinção: todas as regressões fornecem $\alpha = 1/2$. Representamos os dados relativos a três séries temporais, diferentes no valor de μ . Os valores de todas as séries foram obtidos coletando os dados em um intervalo temporal de 1000 passos.

ram caracterizadas por $\alpha = 1/2$ (ver Figura 5.10), ou seja, nenhum tipo de correlação temporal foi encontrada.

5.5 Conclusões

Este modelo, via uma dinâmica microscópica realista definida a partir unicamente da seleção natural e do processo de mutação, exhibe resultados promissores descrevendo, em acordo com as observações obtidas por registros fósseis, distribuições estatísticas do tipo leis de potência no tempo e no espaço. Embora estas predições concordem com os dados empíricos, o modelo não consegue restituir uma dinâmica pontuada com extinções de massa, onde correlações de longo alcance permitam a aglomeração dos eventos de extin-

ção.

Estes fatos são bastante intrigantes e necessitam de investigações futuras. Com o material obtido até agora podemos propor as seguintes possíveis interpretações.

Talvez o comportamento definido pelos nossos dados dependa fortemente da estrutura do modelo que, sendo baseado em uma seleção que se autorregula, não permite grandes flutuações no tamanho da população.

Alternativamente, podemos argumentar ter encontrado um elemento contrário à possibilidade de construir um modelo unificado que, implementando uma microdinâmica de especiação baseada unicamente na força da seleção natural, possa representar a estatística a longo prazo da evolução. Podemos afirmar, parafraseando Philip Anderson, que “*macroevolution is not applied population dynamics*” [23]. Para descrever fenômenos macroevolutivos seria então necessário definir outras dinâmicas, considerando também interações de ordem maior, como relações tróficas entre diferentes níveis ou uma dinâmica intrínseca adicionada àquela gerada pela seleção natural [80], como já apresentado na seção 1.1 do capítulo introdutório. Um possível atalho poderia ser incluir processos qualitativamente diferentes, relacionados unicamente com os eventos de extinção de massa, que possam ser introduzidos contemplando a representação de algum estresse externo, seja biótico ou abiótico [96].

De todo modo, também no caso em que este último cenário poder ser considerado o mais próximo à realidade, podemos afirmar que o modelo é capaz de descrever, através da implementação de verdadeiros eventos de especiação, as distribuições estatísticas relacionadas com a taxa de recolocação espontânea de uma espécie por outra.

Apêndice

Modelo Penna sexuado, o programa

Segue uma implementação em FORTRAN do clássico programa do Modelo Penna sexuado, inspirado em aquele apresentado na referência [9].

```
PROGRAM nbit
IMPLICIT NONE
INTEGER POPDIM, X, XX
PARAMETER (POPDIM=1000000, X=2, XX=64 )

C      array dimension, X: number of strings, XX=X*32: total number of alleles

INTEGER POPMAX, INIPOP, MAXSTEP, MEDSTEP, MINAGE, MAGE, FAGE,
&     LIM, MMUT, FMUT, BIRTH, PREGN, D, T, VERHU,
&     DOMINANCE(1:X), MSTEP, FA, MO, I, P, SEED, IBM, II,
&     MC, NC, G1(1:X), G2(1:X), GENE1(1:X), GENE2(1:X),
&     N, AGEI, CR, IMUT, AGE, WAGE, NMUT, BREED, PREG, POP, MPOP, FPOP,
&     MNUMBER(0:XX), FNUMBER(0:XX), BIT(0:32), MYMED(0:XX),
&     FYMED(0:XX), GEN1M(POPDIM, X), GEN2M(POPDIM, X),
&     GEN1F(POPDIM, X), GEN2F(POPDIM, X),
&     DATAM(POPDIM), DATAF(POPDIM), MUTAM(POPDIM), MUTAF(POPDIM),
&     EXHAUST(POPDIM), PREGNANCY(POPDIM), KO, W

REAL FACTOR, AV, VER, MYTOT

PARAMETER (POPMAX=400000, INIPOP=1000, MAXSTEP=5,
&     MEDSTEP=1000, MINAGE=8, MAGE=64, FAGE=64, LIM=3, MMUT=7,
&     FMUT=7, BIRTH=4, PREGN=1, D=12, SEED=10)

C      PARAMETERS: VERHULST POPULATION, INITIAL POP., MAXIMUM TIME,
C      NUMBER OF PASS-FROM THE END-INCLUDED IN AVERAGE,
C      REPRODUCTIVE AGE: MIN, MAX MALE, MAX FEMALE,
C      MAX.ACTIVE MUTATIONS, MALE MUT., FEMALE MUT., BIRTH RATE,
C      WAITING PERIOD, DOMINANT GENES, SEED>1

PRINT *, POPMAX, INIPOP, MAXSTEP, MEDSTEP, MINAGE, MAGE, FAGE, LIM, MMUT,
&     FMUT, BIRTH, PREGN, D, SEED
```

```

C *****INITIALIZATIONS*****

VER=0
KO=0
IBM=2*SEED-1
FACTOR=1.0/2147483647
AV=1.0/MEDSTEP
MSTEP=MAXSTEP-MEDSTEP
MPOP=INIPOP
FPOP=INIPOP

C   initialisation of practical vectors, bit is a
C   ortonormalized base for a 33d space, the rest is just 0

DO I=1,X
  DOMINANCE(I)=0
END DO
BIT(0) = 1
DO I=0,32
  IF(I.GT.0) BIT(I) = ISHFT(BIT(I-1),1)
  IBM = IBM*16807
END DO
DO I=0,XX
  MYMED(I) = 0
  FYMED(I) = 0
  MNUMBER(I) = 0
  FNUMBER(I) = 0
END DO

C   RANDOM INITIALIZATION OF THE DOMINANT LOCI

DO I=1,D
4   IBM = IBM*16807
   IF(IBM.LT.0) IBM = (IBM+2147483647)+1
   CR=IBM*FACTOR*X+1
   IF (CR.GT.X) GOTO 4
5   IBM = IBM*16807
   P = BIT(ISHFT(IBM,-27))
   IF(IAND(DOMINANCE(CR),P).NE.0) GOTO 5
   DOMINANCE(CR) = IOR(DOMINANCE(CR),P)
ENDDO
MNUMBER(0) = MPOP
FNUMBER(0) = MPOP
DO W=1,X
  DO I=1,MPOP
    GEN1M(I,W) = 0
    GEN1F(I,W) = 0
    GEN2M(I,W) = 0
    GEN2F(I,W) = 0
  ENDDO

```

```

ENDDO
DO I=1,MPOP
  DATAM(I) = 0
  DATAF(I) = 0
  MUTAM(I)=0
  MUTAF(I)=0
  EXHAUST(I)=0
  PREGNANCY(I)=0
ENDDO

C*****TIME LOOP *****

DO 7 T=1,MAXSTEP
  POP = FPOP + MPOP
  VER = (POP*2.0/POPMAX-1.0)
  IF (VER.GE.1) THEN
    VERHU = 2147483647
  ELSE
    VERHU = VER*2147483647
  ENDIF

C*****LOOP ON MALE INDIVIDUALS *****

  I = 1

C  START LOOP
8  AGE = DATAM(I)
   NMUT = MUTAM(I)

C  ACTIVE MUTATIONS -N>1-

   W=(AGE/32.)+1
   WAGE=AGE-(W-1)*32
   N = IAND(1,ISHFT(GEN1M(I,W),-WAGE)) +
&     IAND(1,ISHFT(GEN2M(I,W),-WAGE))+
&     IAND(1,ISHFT(DOMINANCE(W),-WAGE))
   IF (N.GE.2) NMUT = NMUT + 1
   MNUMBER(AGE) = MNUMBER(AGE) - 1
   AGE = AGE + 1

C  AGEING SELECTION

   IF (NMUT.GE.LIM.OR.AGE.EQ.(XX+1)) THEN

C  DEATH
   IF(MPOP.LE.1) GOTO 1
   DO W=1,X
     GEN1M(I,W) = GEN1M(MPOP,W)
     GEN2M(I,W) = GEN2M(MPOP,W)
   END DO
   DATAM(I) = DATAM(MPOP)

```

```

        MUTAM(I) = MUTAM(MPOP)
        MPOP = MPOP - 1
    ELSE

C     ALIVE
        MNUMBER(AGE) = MNUMBER(AGE) + 1
        DATAM(I) = AGE
        MUTAM(I) = NMUT
        I = I + 1

        ENDIF
        IF (I.LE.MPOP) GOTO 8

C*****LOOP ON FEMALE INDIVIDUALS *****

        I = 1
        FA = FPOP

C     START LOOP

9     AGE = DATAF(I)
        NMUT = MUTAF(I)
        PREG = PREGNANCY(I)

C     ACTIVE MUTATIONS -N>1-

        W=(AGE/32.)+1
        WAGE=AGE-(W-1)*32
        N = IAND(1,ISHFT(GEN1F(I,W),-WAGE)) +
&         IAND(1,ISHFT(GEN2F(I,W),-WAGE))+
&         IAND(1,ISHFT(DOMINANCE(W),-WAGE))
        IF (N.GE.2) NMUT = NMUT + 1
        FNUMBER(AGE) = FNUMBER(AGE) - 1
        AGE = AGE + 1

C     AGEING SELECTION

        IF (NMUT.GE.LIM.OR.AGE.EQ.(XX+1)) THEN

C     DEATH
        IF(FPOP.LE.1) GOTO 1
        DO W=1,X
            GEN1F(I,W) = GEN1F(FPOP,W)
            GEN2F(I,W) = GEN2F(FPOP,W)
        END DO
        DATAF(I) = DATAF(FPOP)
        MUTAF(I) = MUTAF(FPOP)
        FPOP = FPOP - 1
        IF (FPOP.GE.FA) THEN
            I = I + 1
        ELSE

```

```

        FA = FA - 1
    ENDIF
ELSE

C   ALIVE
    FNUMBER(AGE) = FNUMBER(AGE) + 1
    IF(PREG.GT.0) PREG = PREG - 1
    DATAF(I) = AGE
    MUTAF(I) = NMUT
    PREGNANCY(I) = PREG

C   ****   BREEDING
    IF (AGE.GE.MINAGE.AND.AGE.LE.FAGE.AND.PREG.EQ.0) THEN

C   FOUNDING A PARTNER, 20 ATTEMPTS
    DO 10 N=1,20
        IBM = IBM*16807
        IF(IBM.LT.0) IBM = IBM+2147483647+1
        II=IBM*FACTOR*MPOP+1
        IF(II.GT.MPOP.OR.II.LE.0) GOTO 10
        AGEI=DATAM(II)
        BREED=EXHAUST(II)
    IF(AGEI.GT.MAGE.OR.AGEI.LT.MINAGE.OR.BREED.NE.0) GOTO 10
    GOTO 11

C   PARTNER FOUND
10   CONTINUE
    GOTO 17

C   **   START REPRODUCTION

11   DO 12 N=1,BIRTH

C   HEROD SELECTION
    IBM = IBM*16807
    IF (IBM.LE.VERHU) GOTO 12

C   FEMALE CROSSING
141   IBM=IBM*65549
        CR=ISHFT(IBM,-27)
        IF (CR.LE.0) GOTO 141
        MC=BIT(CR)-1
        NC=NOT(MC)
241   IBM = IBM*16807
        IF(IBM.LT.0) IBM = (IBM+2147483647)+1
        CR=IBM*FACTOR*X+1
        IF (CR.GT.X) GOTO 241
        G1(CR)=IOR(IAND(GEN1F(I,CR),MC),IAND(GEN2F(I,CR),NC))
        G2(CR)=IOR(IAND(GEN1F(I,CR),NC),IAND(GEN2F(I,CR),MC))
        IF (CR.EQ.1) THEN
            DO W=2,X

```

```

        G1(W) = GEN1F(I,W)
        G2(W) = GEN2F(I,W)
    END DO
ELSE IF (CR.EQ.X) THEN
    DO W=1,(CR-1)
        G1(W) = GEN2F(I,W)
        G2(W) = GEN1F(I,W)
    END DO
ELSE
    DO W=1,(CR-1)
        G1(W) = GEN2F(I,W)
        G2(W) = GEN1F(I,W)
    END DO
    DO W=(CR+1),X
        G1(W) = GEN1F(I,W)
        G2(W) = GEN2F(I,W)
    END DO
END IF
IBM=IBM*65539
IF (IBM.LT.0) THEN
    DO W=1,X
        GENE1(W)=G1(W)
    END DO
ELSE
    DO W=1,X
        GENE1(W)=G2(W)
    END DO
ENDIF

C    FEMALE GAMETE MUTATION
    DO 13 IMUT=1,FMUT
242    IBM = IBM*16807
        IF(IBM.LT.0) IBM = (IBM+2147483647)+1
        CR=IBM*FACTOR*X+1
        IF (CR.GT.X) GOTO 242
        IBM=IBM*16807
        P=BIT(ISHFT(IBM,-27))
        GENE1(CR)=IOR(GENE1(CR),P)
13    CONTINUE

C    MALE CROSSING
142    IBM=IBM*65549
        CR=ISHFT(IBM,-27)
        IF (CR.LE.0) GOTO 142
        MC=BIT(CR)-1
        NC=NOT(MC)
243    IBM = IBM*16807
        IF(IBM.LT.0) IBM = (IBM+2147483647)+1
        CR=IBM*FACTOR*X+1
        IF (CR.GT.X) GOTO 243

```

```

G1(CR)=IOR(IAND(GEN1M(II,CR),MC),IAND(GEN2M(II,CR),NC))
G2(CR)=IOR(IAND(GEN1M(II,CR),NC),IAND(GEN2M(II,CR),MC))
IF (CR.EQ.1) THEN
  DO W=2,X
    G1(W) = GEN1M(II,W)
    G2(W) = GEN2M(II,W)
  END DO
ELSE IF (CR.EQ.X) THEN
  DO W=1,(CR-1)
    G1(W) = GEN2M(II,W)
    G2(W) = GEN1M(II,W)
  END DO
ELSE
  DO W=1,(CR-1)
    G1(W) = GEN2M(II,W)
    G2(W) = GEN1M(II,W)
  END DO
  DO W=(CR+1),X
    G1(W) = GEN1M(II,W)
    G2(W) = GEN2M(II,W)
  END DO
END IF
IBM=IBM*65539
IF (IBM.LT.0) THEN
  DO W=1,X
    GENE2(W)=G1(W)
  END DO
ELSE
  DO W=1,X
    GENE2(W)=G2(W)
  END DO
ENDIF

C   MALE GAMETE MUTATION
DO 16 IMUT=1,MMUT
244   IBM = IBM*16807
      IF(IBM.LT.0) IBM = (IBM+2147483647)+1
      CR=IBM*FACTOR*X+1
      IF (CR.GT.X) GOTO 244
      IBM=IBM*16807
      P=BIT(ISHFT(IBM,-27))
      GENE2(CR)=IOR(GENE2(CR),P)
16    CONTINUE

C   FUSION OF GAMETI WITH 50/50 FEMALE AND MALE OFFSPRING

IBM=IBM*16807
IF (IBM.LT.0) THEN
  FNUMBER(0) = FNUMBER(0) + 1
  FPOP = FPOP + 1
  IF (FPOP.GT.POPDIM) GOTO 101

```

```

        DATAF(FPOP) = 0
        MUTAF(FPOP)=0
        PREGNANCY(FPOP)=0
        DO W=1,X
            GEN1F(FPOP,W) = GENE1(W)
            GEN2F(FPOP,W) = GENE2(W)
        ENDDO
    ELSE
        MNUMBER(O) = MNUMBER(O) + 1
        MPOP = MPOP + 1
        IF (MPOP.GT.POPDIM) GOTO 101
        DATAM(MPOP) = 0
        MUTAM(MPOP)=0
        EXHAUST(MPOP)=0
        DO W=1,X
            GEN1M(MPOP,W) = GENE1(W)
            GEN2M(MPOP,W) = GENE2(W)
        ENDDO
    ENDIF

12    CONTINUE

C    **            END OF REPRODUCTION
        EXHAUST(II)=1
        PREGNANCY(I)=PREGN

C    ****            END OF BREEDING
    ENDIF

17    I = I + 1

    ENDIF

C*****END LOOP ON FEMALE INDIVIDUALS *****
C*****END OF INDIVIDUALS LOOP *****

        IF (I.LE.FA) GOTO 9
        DO 15 I=1,MPOP
            EXHAUST(I)=0
15    CONTINUE

C    COLLECTING OUTPUT

        IF (T.LT.MSTEP) GOTO 7
        DO 20 I=0,XX
            MYMED(I) = MYMED(I) + MNUMBER(I)
            FYMED(I) = FYMED(I) + FNUMBER(I)
20    CONTINUE

C*****END OF TIME LOOP, END OF DYNAMIC*****
7    CONTINUE

```

```
C*****WRITING OUTPUT*****
DO I=0,XX
  MYTOT = MYTOT + MYMED(I) + FYMED(I)
ENDDO
PRINT *, "*****t,mpop,fpop,TOTpop*****",T,MPOP,FPOP,MPOP+FPOP
PRINT *, "*****MEDTOT*****", MYTOT*AV
PRINT *, "age,final Mnumber,f.Fnumber,MMed ul.T.,FMed ul.T"
PRINT 100 , (I,MNUMBER(I),FNUMBER(I),AV*MYMED(I), AV*FYMED(I),
&      I=0,XX)
100 FORMAT (1X,I3,2I9,2F9.2)

STOP

C*****CONTROL SEQUENCES*****
1 PRINT*, 'ERROR, EVERYBODY DIED'
STOP

101 PRINT*, 'ERROR, ARRAY OVERFLOW'
STOP

END
```


Bibliografia

- [1] E. Majorana, *Il valore delle leggi statistiche nella fisica e nelle scienze sociali*, Scientia **36**, 58 (1942).
- [2] E. Schrödinger, *What is life?*, Cambridge University Press, (1948).
- [3] C. Darwin, *On the origin of species*, J. Murray, London (1859).
- [4] S.J. Gould, *The Structure of Evolutionary Theory*, Belknap Press, Cambridge, MA (2002).
- [5] J. Maynard Smith *Evolutionary Genetics*, Oxford Univ. Press, Oxford (1989).
- [6] M. Kimura, Nature, **217**, 624 (1968). M. Kimura, *The neutral theory of molecular evolution*, Cambridge Univ. Press, Cambridge (1983).
- [7] S. Kauffman, *At home in the universe*, Oxford Univ. Press, New York (1995).
- [8] B. Drossel, Adv, Phys. **50**, 209 (2001).
- [9] S. Moss de Oliveira, P.M.C. de Oliveira and D. Stauffer, *Evolution, Money, War and Computers*, Teubner (1999).
D. Stauffer, S. Moss de Oliveira, P.M.C. de Oliveira and J.S. Sá Martins, *Biology, Sociology, Geology by Computational Physicists*, Elsevier, Amsterdam (2006), no prelo.
- [10] J.P. Crutchfield and P. Schuster Editors, *Evolutionary Dynamics*, Oxford Univ. Press, New York (2003).

- [11] J. Brown and G. West, *Scaling in Biology*, Oxford University Press, New York (2000).
V.M. Savage, J.F. Gillooly, W.H. Woodruff, G.B. West, A.P. Allen, B.J. Enquist and J.H. Brown, *Functional Ecology* **18**, 257 (2004).
- [12] P.M.C. de Oliveira, e-print: cond-mat/0101170 (2001).
- [13] L. Peliti, e-print: cond-mat/9712027 (1997).
- [14] M. Eigen, *Naturwissenschaften* **58**, 465 (1971).
- [15] C.R. Shalizi, K.L. Shalizi and R. Haslinger, *Phys. Rev. Lett.* **93**, 118701 (2004).
- [16] J.R. Banavar and A. Maritan, *Comp. Phys. Comm.* **146**, 129 (2002).
- [17] J.M. Yeomans, *Statistical Mechanics of Phase Transitions* Oxford Univ. Press, New York, (1993).
- [18] M. Kardar, Notas de aula para o curso ministrado no MIT, encontradas em: <http://web.mit.edu/8.334/www/index.html>
- [19] V. Schwämmle, K. Luz-Burgoa, J. S. Sá Martins, S. Moss de Oliveira, e-print: q-bio.PE/0508016 (2005).
K. Luz-Burgoa, e-print: q-bio.PE/0504006 (2005).
- [20] H.J. Jensen, *Self-Organized Criticality*, Cambridge Univ. Press, UK (1998).
- [21] P.M.C. de Oliveira, *Computing Boolean Statistical Models*, World Scientific, New York (1991).
- [22] J. Cole, *Brit. J. Phil. Sci.* **51**, 187 (2000).
- [23] P.W. Anderson, *Science* **177**, 393 (1972).
- [24] G. Parisi, *Physica A* **263**, 557 (1999).
- [25] T.J.P. Penna, *J. Stat. Phys.* **78**, 1629 (1995).

-
- [26] B. Charlesworth, *Evolution in age-structured populations*, Cambridge University Press, Cambridge (1994).
- [27] G.C. Williams, *Evolution*, **11**, 398 (1957).
- [28] E.B. Edney and R.W. Gill, *Nature*, **220**, 281 (1968).
- [29] P.B. Medawar, *Modern Quarterly*, **2**, 30 (1946).
- [30] P.B. Medawar, *An unsolved problem in biology*, Lewis, London (1952).
- [31] M.R. Rose, *Evolutionary theory of Aging*, Oxford University Press, Oxford (1991).
- [32] J. Maynard Smith, *A. Rev. Ecol. Syst.*, **9**, 31 (1978).
- [33] G.A. Parker and J. Maynard Smith, *Nature*, **348**, 27 (1990).
- [34] L. Partridge and N.H. Barton, *Nature*, **362**, 305 (1993).
- [35] L.D. Mueller, *Proc. Nat. Acad. Sci. USA*, **84**, 974, (1987).
- [36] Z. Medved, *Biol. Rev.* **65**, 375 (1990).
- [37] S. Moss de Oliveira, D. Alves, and J.S. Sá Martins, *Physica A* **285**, 77 (2000).
- [38] S. Moss de Oliveira, J.S. Sá Martins, P.M.C. De Oliveira, K. Luz-Burgoa, A. Ticona, and T.J.P. Penna, *Comp. Sci. Eng.* **6**, 74 (2004).
- [39] S. Cebrat, *Physica A* **258**, 493 (1998).
- [40] T.J.P. Penna, S. Moss de Oliveira and D. Stauffer *Phys. Rev. E* **52**, R3309 (1995).
- [41] T.J.P. Penna, D. Stauffer, *Z.Phys. B* **101**, 469 (1996).
- [42] K. Malarz, *Int. J. Mod. Phys. C* **11**, 309 (2000).
- [43] R.M.C. de Almeida and G.L. Thomas, *Int. J. Mod. Phys. C* **11**, 1209 (2000).

- [44] A.Raab, J. Stat. Phys. **91**, 1055 (1998).
- [45] J.S. Sá Martins and S. Cebrat, Theory Biosci. **119**, 156 (2000).
- [46] A. Laszkiewicz, S. Cebrat and D. Stauffer, Advance in Complex Systems, **8**, 7 (2005).
- [47] J.B. Coe and Y. Mao, Phys. Rev. E **72**, 051925 (2005).
- [48] M.V. Ashley, M.F. Willson, O.R.W. Pergams, D.J. O'Dowd, S.M. Gende, and J.S. Brown, Biol.Conserv. **111**, 115 (2003).
- [49] S.R. Palumbi, Science **293**, 1786 (2001).
- [50] R. Law, ICES J.Mar.Sci. **57**, 659 (2000)
- [51] D.O. Conover, S.B. Munch, Science **297**, 94 (2002).
- [52] F. Fenner and P.J. Kerr, *The evolutionary biology of Viruses*, edited by S.Morse, Raven Press Ltd, New York (1994).
- [53] D.W. Coltman, P. O'Donoghue, J.T. Jorgenson, J.T. Hogg, C. Strobeck and M. Festa-Bianchet, Nature **426**, 655 (2003)
- [54] S. F. Elena, V.S. Cooper and R.E. Lenski, Science **272**, 1802 (1996).
- [55] D. Papadopoulos, D. Schneider, J. Meier-Eiss, W. Arber, R.E.Lenski and M.Blot, Proc. Nat. Acad. Sci. **96**, 3807 (1999).
- [56] D.L. Hartl, *A primer of population genetics*, Sinauer Associates, Sunderland, MA (1988).
- [57] T.J.P. Penna, A. Racco, and A.O. Sousa, Physica A **295**, 31 (2001).
- [58] J.S. Sá Martins, S. Moss de Oliveira, and G.A. de Medeiros, Phys. Rev. E **64**, 021906 (2001).
- [59] A. Ticona, T.J.P. Penna, Braz. Jour. Phys. **33**, 619 (2003).
- [60] K. Luz-Burgoa, S. Moss de Oliveira. J.S. Sá Martins, D. Stauffer, and A.O. Sousa, Braz. Jour. Phys. **33**, 623 (2003).

-
- [61] S. Moss de Oliveira, T.J.P. Penna and D. Stauffer, *Physica A* **215**, 298 (1995).
- [62] D.W. Coltman, M. Festa-Bianchet, J.T. Jorgenson and C. Strobeck, *Proc. R. Soc. Lond. B* **269**, 165 (2002).
- [63] W. Horsthemke and R. Lefever, *Phys. Let. A* **64**, 19 (1977).
- [64] E. Mayr *Animal Species and Evolution*, Belknap Press, Cambridge, MA (1963).
- [65] J.A. Coyne and H. A. Orr, *Speciation*, Sinauer Associates, Sunderland, MA (2004).
- [66] W.R. Rice, and E.E. Hostert, *Evolution* **47**, 1637 (1993).
- [67] J.M. Thoday and B.B. Gibson, *Nature* **193**, 1164 (1962).
- [68] J. Maynard Smith, *Am. Nat.* **100**, 637 (1966).
- [69] G.L. Bush, *Evolution* **23**, 237 (1969).
- [70] U.K. Schlieven, D. Tautz and S. Pääbo, *Nature* **368**, 629 (1994).
- [71] M.L. Friesen et al., *Evolution* **58**, 254 (2004).
- [72] S.S. Chow, C.O. Wilke, C. Ofria, R.E. Lenski and C. Adami, *Science* **305**, 84 (2004).
- [73] M. Turelli, N.H. Barton and J.A. Coyne, *Trends in Ecol. & Evol.* **16**, 330 (2001).
- [74] S. Via, *Trends in Ecol. & Evol.* **16**, 381 (2001).
- [75] A.S. Kondrashov and F.A. Kondrashov, *Nature* **400**, 351 (1999).
- [76] U. Dieckmann and M. Doebeli, *Nature* **400**, 354 (1999).
- [77] M. Kirkpatrick and S.L. Nuismer, *Proc. R. Soc. Lond. B* **271**, 687 (2004).
- [78] F. Bagnoli and C. Guardiani, *Physica A* **347**, 534 (2005).

- [79] J. Polechova and N.H. Barton, *Evolution* **59**, 1194 (2005).
- [80] S.J. Gould and N. Eldredge, *Paleobiology* **3**, 115 (1977).
- [81] S.J. Gould and N. Eldredge, *Nature* **366**, 223 (1993).
- [82] D.M. Raup, *Science* **231**, 1528 (1986).
- [83] P. Bak and K. Sneppen, *Phys. Rev. Lett.* **71**, 4083 (1993).
- [84] P.M.C. de Oliveira, J.S. Sá Martins, D. Stauffer and S. Moss de Oliveira, *Phys. Rev. E* **70**, 051910 (2004).
- [85] P.A. Rikvold and R.K.P. Zia, *Phys. Rev. E* **68**, 031913 (2003).
- [86] N.C. Stenseth and J. Maynard Smith, *Evolution* **38**, 870 (1984).
- [87] M.E.J. Newman and R.G. Palmer, e-print: adap-org/9908002.
M. E. J. Newman and R. Palmer, *Modeling Extinction*, Oxford University Press, New York (2003).
- [88] R.V. Solé and S.C. Manrubia, *Phys. Rev. E* **54**, R42 (1996).
- [89] R.V. Solé and J. Bascompte, *Proc. R. Soc. London B* **263**, 161 (1996).
- [90] R.V. Solé, S.C. Manrubia, M. Benton and P. Bak, *Nature* **388**, 764 (1997).
- [91] D. Chowdhury, D. Stauffer and A. Kunwar, *Phys. Rev. Lett.* **90**, 068101 (2003).
- [92] D. Chowdhury and D. Stauffer, *Physica A* **318**, 461 (2003).
- [93] D. Chowdhury and D. Stauffer, *Physica A* **340**, 685 (2004).
- [94] M. He and G. Yu, *Int. J. Mod. Phys. C* **17**, (2005).
- [95] C.K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, A.L. Goldberger, *Phys Rev E* **49**, 1685 (1994). C.K. Peng, S. Havlin, H.E. Stanley, A.L. Goldberger, *Chaos* **5**, 82 (1995).
<http://www.physionet.org/physiotools/dfa/>

[96] M.E. Newman, Proc. R. Soc. London B **263**, 1605 (1996).

O trabalho e o material apresentado nesta tese foi publicado nos seguintes artigos:

- E. Brigatti, J. S. Sá Martins e I. Roditi, *Scaling properties of the Penna model*, Eur. Phys. J. B **42**, 431 (2004).
- E. Brigatti, J. S. Sá Martins e I. Roditi, *Sharp gene pool transition in a population affected by phenotype-based selective hunting*, Eur. Phys. J. B **45**, 529 (2005).
- E. Brigatti, J. S. Sá Martins e I. Roditi, *Evolution of polymorphism and sympatric speciation through competition in a unimodal distribution of resources*, e-print: q-bio.PE/0505017
- V. Schwammle e E. Brigatti, *Speciational view of macroevolution: are micro and macroevolution decoupled?*, e-print: q-bio.PE/0509032