

Danilo Germano Muniz da Silva

# SEXO NO ESPAÇO



São Paulo

2015

Danilo Germano Muniz da Silva

# SEXO NO ESPAÇO

Tese apresentada ao Instituto de Biociências da Universidade de São Paulo como parte dos requisitos para obtenção do título de Doutor em Ciências, na área "Ecologia".

**Orientador:** Glauco Machado

**Co-orientador:** Paulo Roberto Guimarães Jr.

Departamento de Ecologia, Universidade de São Paulo, Brasil

São Paulo

2015

## FICHA CATALOGRÁFICA

Silva, Danilo Germano Muniz da  
Sexo no espaço / Danilo Germano Muniz da Silva; Orientador  
Glauro Machado. --. São Paulo, 2015.  
130 f.

Tese (Doutorado) – Instituto de Biociências da Universidade de  
São Paulo. Departamento de Ecologia.

1. Distribuição espacial. 2. Modelos baseados em indivíduos. 3. Seleção sexual. 4. Sistemas de acasalamento. 5. Teoria de redes I. Universidade de São Paulo. Instituto de Biociências. Departamento de Ecologia. II. Título.

*Missão dada, é missão cumprida.*

*Coronel Antunes (Tropa de Elite, 2007)*

*Space is disease and danger wrapped in darkness and silence.*

*Dr. Henry McCoy (Star Trek, 2009)*

## Sumário

Resumo .....	8
Abstract.....	9
<b>Introdução geral</b> .....	10
Apresentação geral da tese.....	14
Referências.....	16
<b>Capítulo 1: A sexual network approach to sperm competition in a species with alternative mating tactics</b> .....	18
Abstract.....	19
Introduction.....	20
Material and methods.....	22
Results .....	29
Discussion.....	34
Acknowledgements.....	38
References.....	38
Supplementary material .....	42
<b>Capítulo 2: A spatially explicit model of mate choice and assortative mating</b> .....	48
Abstract.....	49
Introduction.....	50
Multinomial model of mate choice .....	52
Binomial model of mate choice .....	54
Individual based simulation .....	55
Analysis of the simulation data.....	56
Results .....	58
Discussion.....	62
References.....	64
Supplemental Material .....	68
<b>Capítulo 3: Past mating frequency and social environment lead to strategic movements in a scramble competition polygyny</b> .....	85
Summary.....	86
Introduction.....	87
Methods .....	88
Results .....	91
Discussion.....	95
Acknowledgements.....	98
References.....	98

<b>Capítulo 4: Mate sampling influences the intensity of sexual selection and the evolution of costly sexual ornaments.....</b>	<b>102</b>
Abstract.....	103
Introduction.....	104
Methods .....	105
Results .....	110
Discussion.....	113
Acknowledgements.....	115
References.....	115
Supplemental Material .....	118
<b>Conclusões gerais .....</b>	<b>127</b>
Referências.....	130

## Agradecimentos

Eu queria começar agradecendo a minha família: meu pai Coca, minha mãe Gina, meu irmão Yuri, minha amada namorada Priscila e minha sogra Maria. Por todo o apoio, todo o amor, toda a paciência e por (aparentemente) acreditarem que eu entendo alguma coisa de biologia e que eu tenho algum futuro. Eu queria agradecer ao meu pai, que sempre foi sincero comigo, sempre dividindo comigo o conhecimento dele sobre como o mundo funcionava ao invés de inventar historinhas pra criança. Apesar do meu pai não ser cientista, se alguém me estimulou a gostar de ciência desde cedo foi ele. É sério, parando pra pensar agora eu não me lembro de nenhuma vez que ele me contou alguma história da carochinha. Eu quero agradecer a minha mãe, que sempre me ensinou a ser teimoso e perseverante e ir atrás do que eu queria e sempre me apoiou. Foi ela quem fez um *inception* na minha cabeça e me convenceu que eu queria fazer o curso técnico de informática no COTUCA. Ao final do curso, eu descobri que queria ser biólogo, mas sem esse curso eu não teria escrito essa tese. Se a vida é uma grande equação, a minha tem uns parâmetros bem malucos. Eu queria agradecer ao meu irmão pelo companheirismo, preocupação e especialmente pelos *Stray Cats*. Ele vai entender. Finalmente eu agradeço demais à Priscila, por todo o apoio, por acreditar em mim e por andar ao meu lado, rir comigo e dividir sua vida comigo. Como eu já disse antes, sem essas pessoas tudo teria sido uma dureza inacreditável.

Eu não escrevi essa tese sozinho, eu tive muito mais colaboradores do que eu jamais imaginei e sem cada uma dessas pessoas eu não teria aprendido tanto, produzido tanto, nem viajado ou me divertido tanto. Por isso eu sou muito grato a Bruno Buzatto (*Uncle Wiggily*), Eduardo Santos, Rogélio Macías, Martha Baena, Shinichi Nakagawa, Paulo Guimarães (Miúdo) e Glauco Machado. Eu também queria fazer uma menção mais do que honrosa a dois colaboradores que com os quais eu estou escrevendo artigos que não entraram na tese: Louise Alyssa e Andres Rojas, com os quais eu também aprendi e me diverti muito.

Agora eu preciso tecer meus agradecimentos às pessoas maravilhosas da LAGE (Laboratório Assustadoramente Grande de Ecologia), este ambiente intelectualmente estimulante onde se pode conversar sobre ecologia, desenhos japoneses, evolução, *Mad Max*, equações diferenciais, sexo, teoria de jogos, brigas conjugais, jogos de RPG e todo tipo de maluquice. Em primeiro lugar, eu agradeço demais ao Glauco, que me incluiu nesse laboratório e ao longo de todo o doutorado foi um orientador super dedicado e um grande

amigo. Ele não apenas me ensinou milhões de coisas como soube me orientar (ahá!) sobre quais tipos de conhecimentos eu deveria buscar, como busca-los, quais cursos fazer e como alcançar meus objetivos acadêmicos. Graças ao Glauco eu tenho participado de cursos de campo todo ano desde 2012 (mais detalhes adiante) e tenho aprendido demais com isso. Também sou muito grato ao Miúdo, meu co-orientador, por todo o aprendizado sobre redes, piadas sobre a UNICAMP e adjacências e por também ser um grande amigo. Para fechar a trilogia dos donos do morro, não posso esquecer-me de agradecer ao Gandalf (*a.k.a.* Paulo Inácio). Por algum tempo, eu frequentei as reuniões do LET com ele e seus alunos e certamente isso me ajudou muito com as simulações e as equações. Além disso, o Paulo esteve envolvido em quase todas as disciplinas que eu fiz durante o doutorado, sem as quais essa tese teria uma cara totalmente diferente, muito menos geral e muito menos ambiciosa. Continuando a falar da LAGE, eu sou muito grato aos meus irmãos de laboratório Andres, Billy, Edu, Érika, Ernesto, Louise, Lygia, Musgo, Pedrinho, Rachel, Renatcho, Rosa, Soly, Tais e Vivian por gerarem um ambiente tão divertido e instrutivo e gerar reuniões com mais de doze pessoas dedicadas a falar sobre sexo. Também queria agradecer a todos os demais lagueanos, lagueanas e simpatizantes por gerar o ambiente que eu descrevi acima.

Ao longo dos últimos quatro anos eu fui monitor/tutor/professor de alguns cursos e eu queria muito agradecer aos coordenadores dessas disciplinas: Alexandre Adalardo (Curso de R da USP), Paulo Inácio e João Batista (Curso de Ajuste de Modelos da USP), Glauco (Curso de Campo da USP), Paulo Inácio (de novo!) e Roberto Kraenkel (Curso de Biologia Matemática do IFT), Diogo Meyer (Disciplina de Mecanismos Evolutivos da USP) e finalmente Paulinho Enrique e Paulo Estéfano (Curso de Campo da Amazônia). Eu sou muito grato a todos eles por terem me convidado (ou aceitado meu autoconvite) para participar dos cursos. Cada um destes cursos foi uma oportunidade incrível para interagir com uma diversidade estrambólica de alunos e aprender demais com eles. Eu sou muito grato a todos os alunos com os quais eu interagi nestes cursos e espero sinceramente que eles também tenham aprendido algo com a experiência. Também queria muito agradecer àqueles que foram monitores junto comigo de alguns destes cursos, especialmente Soly, Cris e Ayana. Eu incomodei demais essas três e sou muito grato pela paciência.

A todos que me ajudaram nessa longa estrada do doutorado: meu muito obrigado. Eu aprendi muitas vezes mais do que esperava nesse período e termino o doutorado animado com as possibilidades, ainda mais interessado em entender como o mundo funciona e com milhões de ideias na cabeça.

## Resumo

O objetivo geral desta tese foi investigar como a distribuição espacial dos indivíduos influencia a organização dos sistemas de acasalamento e a intensidade da seleção sexual. A tese contém dois capítulos em que analisamos dados empíricos e dois capítulos em que abordamos questões mais teóricas mais gerais. No primeiro capítulo, avaliamos como a distribuição espacial de haréns defendidos por machos territoriais influencia a rede de competição espermática que existe entre machos do opilião *Serracutisoma proximum*. No segundo capítulo, propusemos um modelo probabilístico de escolha de parceiros que leva em consideração o fato de que as fêmeas estão restritas a amostrar apenas alguns machos da população, e que essa restrição é imposta principalmente pela distribuição espacial dos indivíduos. Nosso modelo foi mais acurado do que modelos alternativos que ignoram o espaço. No terceiro capítulo, investigamos os movimentos de busca de parceiros do besouro *Leptinotarsa undecimlineata*. Encontramos que tanto machos quanto fêmeas se movimentam estrategicamente, saindo de onde estão quando não obtêm cópulas buscando áreas próximas e com muitos parceiros em potencial. Finalmente, no quarto capítulo, voltamos a explorar o efeito de restrições espaciais sobre a escolha de parceiros. Usamos simulações baseadas em indivíduos para investigar como a restrição espacial influencia a seleção sexual e a evolução de ornamentos sexualmente selecionados. Encontramos que quanto maior o número de parceiros que as fêmeas podem amostrar durante a escolha, mais intensa é a seleção sexual, o que permite a evolução de ornamentos mais exagerados nos machos. Além disso, analisamos um conjunto de dados publicados sobre a intensidade de seleção sexual e observamos que estes se ajustam às previsões do modelo teórico. Concluímos que o espaço é muito mais importante para a organização de sistemas de acasalamento e para a intensidade da seleção sexual do que se acreditava previamente. Esperamos que as idéias propostas aqui encontrem terreno fértil na mente do leitor e que gerem novos e estimulantes desdobramentos no campo teórico e empírico.

## Abstract

The main goal of this thesis was to investigate how the spatial distribution of individuals influences the organization of the mating systems and the intensity of sexual selection. The thesis contains two empirical chapters and two chapters with a theoretical scope. In the first chapter, we investigate how the spatial distribution of harems defended by territorial males influences the sperm competition network among males of the harvestman *Serracutisoma proximum*. In the second chapter, we proposed a probabilistic model of mate choice that includes the spatial constraint in the analyses of mate choice. This model takes into account the fact that females are restricted to sample only some males available in the population, and that this restriction is imposed primarily by the spatial distribution of individuals. In the third chapter, we investigate the mate searching movements of the leaf beetle *Leptinotarsa undecimlineata*. We found that both males and females move strategically, leaving the host plant when they do not obtain copulations and seeking neighboring areas with many potential mates. Finally, in the fourth chapter, we explore the effect of spatial restrictions on mate choice. We used individual based simulations to investigate how spatial restrictions influence sexual selection and the evolution of sexually selected ornaments. We found that the higher the number of potential partners female can sample during mate choice, the more intense is sexual selection, which leads to the evolution of more exaggerated male ornaments. In addition, we analyzed a published dataset of intensity of sexual selection and found that the data adjust to the predictions made by our simulations. We conclude that the space is much more important than previously thought to mating systems organization and the intensity of sexual selection. We hope the ideas proposed here can flourish in the minds of the readers, stimulating both empirical and theoretical follow ups.

# Introdução geral<sup>1</sup>

*No practical biologist interested in sexual reproduction would be led to work out the detailed consequences experienced by organisms having three or more sexes; yet, what else should he do if he wishes to understand why the sexes are, in fact, always two?*

R. A. Fisher (1930)

Não, infelizmente esta tese não fala sobre sexo entre alienígenas... Apesar da foto na capa, o assunto da tese na verdade é como a distribuição espacial dos indivíduos influencia diferentes aspectos dos sistemas de acasalamento, como a escolha de parceiros, a busca por parceiros, a competição espermática e a evolução de características sexualmente selecionadas. Um leitor atento poderá notar que, enquanto os capítulos ímpares usam "espécies de verdade" como modelo de estudo, os capítulos pares são essencialmente teóricos. Esses capítulos teóricos quase não fazem referências a "espécies de verdade" e são baseados principalmente em equações e organismos virtuais que vivem dentro do computador. Usando simulações, exploramos as conseqüências de limitações espaciais (no sentido cartesiano, não no sentido sideral) para o estudo e evolução de sistemas de acasalamento. As equações e "bichos virtuais" foram baseados nas características mais elementares dos sistemas de acasalamento reais, incluindo a presença de dois sexos separados, a escolha de parceiros e alguma herdabilidade de características sexualmente selecionadas. Apesar da limitação de apenas dois sexos parecer frustrante, especialmente quando a ficção científica nos apresenta espécies com mais de dois sexos (e.g., Asimov 1972), a teoria aponta que devemos esperar apenas dois sexos na maior parte dos casos (Parker et al. 1972, Lehtonen & Kokko 2011), seja na Terra ou em outros planeta onde, por ventura, exista vida. Portanto, se algum dia descobrirem alienígenas que fazem sexo, possuem sexos separados e escolhem parceiros, espero sinceramente que as conclusões dos capítulos pares desta tese sejam úteis para o estudo de espécies no espaço (agora no sentido sideral!).

---

<sup>1</sup> Este texto nunca vai ser publicado em nenhum outro lugar. Portanto, tomei bastante liberdade em relação à forma como apresento a teoria que embasa minha tese. Espero que minha informalidade não ofenda o leitor. As referências seguem o formato do periódico *Advances in Space Research*. Afinal, por que não?

Deixando os alienígenas de lado, um dos primeiros textos a abordar como o espaço pode ser importante para os sistemas de acasalamento é também o texto mais clássico de toda a literatura de seleção sexual: o artigo de Emlen & Oring publicado em 1977 e intitulado *Ecology, Sexual Selection, and the Evolution of Mating Systems*. Além de contar com mais de 4 mil citações de acordo com o Google Scholar em 2015, este é o artigo que estabelece os conceitos básicos da teoria de sistemas de acasalamento. O objetivo inicial de Emlen & Oring (1977) era gerar um sistema conciso de definições de sistemas de acasalamento focado principalmente em aves e mamíferos. Em seu esforço, eles não apenas classificaram nove formas diferentes de obtenção de parceiros sexuais, como estabeleceram um corpo teórico que faz previsões sobre como a intensidade da seleção sexual em uma população varia em função de condições ambientais. As principais condições ambientais abordadas foram a distribuição espacial e temporal de indivíduos sexualmente receptivos do sexo limitante, que normalmente são as fêmeas. Basicamente, se as fêmeas ficarem receptivas todas ao mesmo tempo, qualquer macho tem muita dificuldade em monopolizar mais que uma fêmea e, com isso, quase todos os machos conseguem copular. O mesmo ocorre se as fêmeas estiverem muito espalhadas espacialmente. Nesse caso, um macho que encontre uma fêmea e copule com ela terá que se movimentar muito antes de encontrar a próxima. Novamente, isso dificulta que um macho monopolize muitas fêmeas. Portanto, se fêmeas receptivas estão muito concentradas no tempo ou espalhadas no espaço, a previsão é que os indivíduos sejam monogâmicos por falta de opção e que a variação no sucesso reprodutivo dos machos seja baixa. Por outro lado, quando as fêmeas estão agregadas espacialmente, é possível que um macho monopolize grupos de fêmeas mantendo outros machos afastados. Nessa situação, existe potencial para a poliginia, em que alguns machos copulam com muitas fêmeas. De forma similar, quando as fêmeas se tornam receptivas de forma assincrônica, um macho tem o potencial para monopolizar múltiplas fêmeas, uma de cada vez, e assim acumular cópulas com várias fêmeas.

Neste momento, alguém poderia dizer "*mas por que essa fixação com machos monopolizando fêmeas?*" ou "*qual a diferença que isso faz na prática?*". Se um macho puder monopolizar (e copular com) muitas fêmeas, seu sucesso reprodutivo será alto. Mais do que isso, em uma população com o mesmo número de machos e fêmeas, quando um macho monopoliza várias fêmeas, outros machos necessariamente ficam sem copular. É dessa variação no sucesso reprodutivo entre machos que pode surgir a seleção sexual (Darwin 1971, Shuster & Wade 2003). Se o macho que copula com várias fêmeas é geneticamente diferente dos que não copulam, os genes deste macho "*super-copulador*" se tornam mais

freqüentes na próxima geração. Assim, características que aumentam a capacidade de macho monopolizar muitas fêmeas, tais como chifres e outros armamentos usados em brigas, podem evoluir por seleção sexual. Note que o argumento é construído em termos de monopolização, de forma que o foco está mais nas disputas entre machos do que na escolha de parceiros pelas fêmeas. Entretanto, os mecanismos clássicos de seleção sexual propostos por Darwin são a disputa entre membros do mesmo sexo e a escolha por membros do sexo oposto (Darwin 1871, Jones & Ratterman 2009). Com isso, parece que o argumento central de Emlen & Oring (1977) deixa de lado um componente importante da seleção sexual, que é a escolha de parceiros. Será que podemos estender o argumento para entender como a distribuição temporal e espacial dos indivíduos influencia a escolha de parceiros?

O efeito da distribuição temporal sobre a escolha de parceiros é relativamente simples: se as fêmeas ficam todas receptivas ao mesmo tempo, os machos preferidos rapidamente ficam "ocupados" ou esgotam seu esperma após algumas cópulas. Dessa forma, se além das fêmeas ficarem receptivas sincronicamente, elas precisarem copular logo (por algum motivo qualquer), nem todas as fêmeas podem copular com os machos favoritos, abrindo oportunidades de cópula para machos de menor qualidade. O efeito da distribuição espacial sobre a escolha de parceiros é um pouco mais difícil. Se as fêmeas estão agregadas, é mais fácil ou mais difícil para elas escolher seus parceiros? A própria pergunta é contra intuitiva, mas pensando um pouco fica claro que a distribuição dos machos também tem um papel no processo de escolha. Afinal, uma fêmea só pode escolher copular com um macho uma vez que se encontre com ele, e o encontro depende dos dois não viverem muito longe um do outro. Portanto, a distribuição espacial dos indivíduos, sua densidade e mobilidade influenciam quem se encontra com quem. O mesmo se aplica a quem será rejeitado ou aceito como parceiro sexual por cada indivíduo. Poderia me estender um pouco mais nesse argumento, mas estaria adiantando o assunto dos capítulos dois e quatro. Sendo assim, vou mudar um pouco de assunto agora. Gostaria de salientar, porém, que a escolha da fêmea é tão negligenciada na teoria sobre sistemas de acasalamento que o processo nem é mencionado no capítulo sobre fenologia da seleção sexual no livro *Mating Systems and Strategies* (Shuster & Wade 2003), uma das melhores sínteses sobre o tema.

Neste momento, poderia enveredar por uma discussão sobre como a distribuição espacial dos indivíduos influencia quais machos brigam contra quais outros machos pelo acesso às fêmeas ou recursos que elas precisam, como sítios de oviposição ou chocolates. Entretanto, não explorei diretamente esse assunto na minha tese e, portanto, vou me ater a outra forma de competição entre machos, que, nas espécies com fertilização interna, acontece depois da cópula. Pois é, apesar da cópula ser considerada um *score* por muita gente que conheço (eu jamais usaria linguajar tão rude!), a competição por fertilizações pode continuar após a cópula (Eberhard 2009, Parker & Birkhead 2013). Quando múltiplos machos copulam com a mesma fêmea, seus ejaculados podem competir pela fertilização dos óvulos em um processo conhecido como competição espermática (Parker 1970). Ou seja, após a cópula, existe toda uma nova arena de competição entre machos cuja intensidade também deve ser influenciada pela distribuição espacial. Se uma fêmea está limitada a copular com os machos disponíveis em sua vizinhança, um macho vai acabar competindo espermaticamente de forma mais intensa com seus vizinhos. Com isso, chegamos ao assunto do primeiro capítulo da minha tese. Mas antes de finalmente apresentar formalmente o conteúdo dos capítulos, existe mais um assunto sobre o qual preciso falar, as redes, que foram a ferramenta metodológica usada para abordar várias questões ao longo da minha tese.

Até agora chamei sua atenção para como a distribuição espacial dos indivíduos, ou seja, onde cada indivíduo está localizado, influencia com quem os indivíduos copulam e com quem eles competem. Em resumo, passamos um bom tempo discutindo como o espaço influencia quem interage com quem e falamos de dois tipos de interações: acasalamento e competição por parceiros. Qualquer conjunto de elementos que interagem entre si pode ser descrito como uma rede (Newman 2010). A teoria de redes essencialmente fornece ferramentas analíticas para caracterizar como os elementos que compõem a rede, chamados nós, estão conectados entre si. Essas ferramentas já foram usadas para descrever todo tipo de rede, mas em ecologia elas são frequentemente usadas para analisar redes de interação entre espécies, tais como teias tróficas, vários tipos de mutualismo e também relações entre parasitas e hospedeiros (Ings et al. 2009). De maneira semelhante, uma população sexual pode ser descrita como um conjunto de indivíduos que podem ou não estar conectados por cópulas, formando uma *rede sexual*. Nesse sentido, se um sistema de acasalamento pode ser definido como um padrão de cópulas entre indivíduos, a teoria de redes fornece um ferramental analítico ideal para o estudo de sistemas de acasalamento de uma forma quantitativa. E foi exatamente isso que fizemos aqui.

## **Apresentação geral da tese**

Nesta tese, eu e um grande número de amigos nos perguntamos como o espaço e a distribuição espacial dos indivíduos influencia a organização dos sistemas de acasalamento e a intensidade da seleção sexual. Em cada um dos quatro capítulos que compõem a tese, abordamos um aspecto diferente dessa pergunta geral. Na seqüência, farei uma breve descrição do conteúdo de cada um dos capítulos, com foco nas questões gerais que orientaram cada um deles. Não apresentarei os resultados aqui, pois farei um breve sumário do que descobrimos nas Conclusões Gerais da tese.

No primeiro capítulo, abordamos a competição espermática entre machos do opilião *Serracutisoma proximum* (Gonyleptidae), uma espécie com morfos masculinos alternativos, em que machos do morfo grande tipicamente defendem territórios contendo haréns, enquanto machos do morfo pequeno invadem os haréns para copular furtivamente. Machos do morfo grande podem permanecer por meses defendendo o mesmo harém, enquanto machos do morfo pequeno são mais vágéis e podem visitar e invadir muitos haréns ao longo da estação reprodutiva. Dado que as fêmeas geralmente aceitam copular com machos do morfo pequeno, a competição espermática deve ser intensa nessa espécie de opilião. Para estimar a intensidade de competição espermática enfrentada por machos de cada morfo, modificamos uma métrica de redes que, no nosso sistema de estudo, representa o grau de promiscuidade das fêmeas. Usando observações detalhadas em campo, investigamos (1) como os morfos masculinos diferem na intensidade de competição espermática que eles enfrentam, (2) quais fatores influenciam a intensidade de competição espermática enfrentada pelos machos territoriais e (3) como a distribuição espacial dos haréns influencia a rede de competição espermática. Este trabalho foi publicado na revista *Behavioral Ecology*.

No segundo capítulo, voltamos nossa atenção para a escolha de parceiros. Supusemos que, antes de escolher um parceiro sexual, um indivíduo tem que "amostrar" indivíduos do sexo oposto na população. Também supusemos que essa amostra pode ser bastante restrita, devido ao custo envolvido na busca de parceiros, por exemplo. Dessa forma, se pensarmos em uma população em que as fêmeas são seletivas e os machos não, uma fêmea com acesso a poucos machos pode acabar copulando com um macho subótimo, ou seja, um macho que não possui as características fenotípicas preferidas pelas fêmeas. Qual a importância disso? Se ignorarmos o fato de que as fêmeas não têm acesso a todos os machos, podemos subestimar a seletividade feminina ao analisar um conjunto de dados em que a maioria das fêmeas copula com machos subótimos. No segundo capítulo da tese, atacamos essa questão

usando equações e "bichinhos virtuais". Nosso objetivo era propor um novo modelo estatístico, baseado nas ferramentas de redes, que consiga estimar de maneira acurada a seletividade feminina mesmo quando a escolha de parceiros esteja limitada espacialmente. Dada a natureza predominantemente metodológica deste capítulo, pretendemos submeter o trabalho para a revista *Methods in Ecology and Evolution*.

No terceiro capítulo, trabalhamos novamente com dados empíricos. Percebemos que, se existe um comportamento sexual que está diretamente conectado ao espaço, este é a movimentação na busca por parceiros. Com isso, decidimos testar a hipótese de que os movimentos do besouro fitófago *Leptinotarsa undecimlineata* (Chrysomelidae) são motivados pela busca por parceiros sexuais. Machos e fêmeas dessa espécie vivem, alimentam-se e acasalam-se sobre suas plantas hospedeiras do gênero *Solanum*. Estudamos uma população de besouros que vive em uma área com alta concentração dessas plantas. Dessa forma, a movimentação dos indivíduos pode ser descrita como mudanças entre plantas cuja posição no espaço era conhecida. Com isso, nos perguntamos *quando* um besouro deveria mudar para outra planta e *para onde* ele deveria ir ao mudar. Nossas hipóteses foram de que as decisões de quando e para onde mudar deveriam ser influenciadas tanto pela distribuição espacial das plantas, quanto pelo contexto social (disponibilidade de parceiros) e sucesso (ou fracasso) copulatório prévio de cada indivíduo (incluindo machos e fêmeas) em cada planta. Vale salientar aqui que, para responder a pergunta de *para onde* os besouros devem mudar, adaptamos o modelo proposto no Capítulo 2, provendo assim a primeira aplicação empírica do método que criamos. Dada a generalidade das questões investigadas neste capítulo, pretendemos submeter o manuscrito para uma revista de maior abrangência, tal como o *Journal of Animal Ecology*.

Por fim, no quarto capítulo, voltamos à questão da amostragem de parceiros em potencial durante o processo de busca e escolha. Se limitações na amostragem realmente fazem as fêmeas copularem com machos subótimos, então existem consequências para os indivíduos da população e não apenas para a nossa habilidade de detectar a escolha feminina. Neste capítulo, voltamos a usar equações (poucas) e "bichinhos virtuais" (muitos) para responder como a amostragem de parceiros provocada por restrições espaciais influencia a intensidade da seleção sexual e a evolução de características masculinas provocada pela seleção sexual. Para não dizer que ficamos apenas nas elucubrações teóricas, aproveitamos um grande banco de dados disponível na literatura para testar se as previsões teóricas sobre a relação entre amostragem de parceiros e a intensidade da seleção sexual

geradas pelo nosso modelo têm apoio na "vida real". Este trabalho foi apresentado recentemente no congresso *Behaviour 2015*, realizado na Austrália, e, de acordo com a sugestão de várias pessoas que assistiram à apresentação, deveríamos focar em uma revista de ecologia e evolução com público amplo. A sugestão mais freqüente que recebemos foi a revista *Ecology Letters* e, portanto, é para lá que pretendemos submeter o manuscrito.

Ao longo dos quatro capítulos, usamos abordagens complementares, incluindo modelos probabilísticos, simulações baseadas em indivíduos e testes de hipótese com dados empíricos para investigar como o espaço e a distribuição espacial dos indivíduos influencia a organização dos sistemas de acasalamento e a intensidade da seleção sexual. Nos capítulos empíricos, usamos uma espécie cujo sistema de acasalamento é uma poliginia por defesa de recursos e outra cujo sistema de acasalamento é uma poliginia por competição desordenada (*scramble competition*). Nos capítulos teóricos, geramos modelos que são aplicáveis a vários tipos de sistema de acasalamento, mas o exemplo empírico apresentado no Material Suplementar do Capítulo 2 refere-se a um sistema de acasalamento socialmente monogâmico, com cópulas extra-par. No capítulo 4, o banco de dados que utilizamos inclui espécies com praticamente todos os tipos de sistema de acasalamento, desde a monogamia social até *leks*. Ao explorar a mesma questão geral em diferentes sistemas de acasalamento, espero conseguir convencê-lo de que o espaço é um componente crucial, porém até então negligenciado, nos estudos sobre seleção sexual, incluindo particularmente a competição espermática (Capítulo 1), a escolha de parceiros (Capítulo 2), a busca por parceiros (Capítulos 3 e 4) e a evolução de características sexualmente selecionadas (Capítulo 4).

## Referências

- Asimov, I., 1972. *The gods themselves*. Doubleday, New York.
- Darwin, C., 1871. *The descent of man and selection in relation to sex*. John Murray, London.
- Eberhard, W.G., 2009. Postcopulatory sexual selection: Darwin's omission and its consequences. *Proc. Natl. Acad. Sci. USA*. 106 Suppl 1, 10025–10032.
- Emlen, S., Oring, L., 1977. Ecology, sexual selection, and the evolution of mating systems. *Science*. 197, 215–223.
- Fisher, R.A., 1930. *The genetical theory of natural selection*. Oxford University Press, Oxford.

- Ings, T.C., Montoya, J.M., Bascompte, J., Blüthgen, N., Brown, L., Dormann, C.F., Edwards, F., Figueroa, D., Jacob, U., Jones, J.I., Lauridsen, R.B., Ledger, M.E., Lewis, H.M., Olesen, J.M., Van Veen, F.J.F., Warren, P.H., Woodward, G., 2009. Ecological networks — beyond food webs. *J. Anim. Ecol.* 78, 253–269.
- Jones, A.G., Ratterman, N.L., 2009. Mate choice and sexual selection : What have we learned since Darwin ? *Proc. Natl. Acad. Sci. USA.* 106, 10001–10008.
- Lehtonen, J., Kokko, H., 2011. Two roads to two sexes: Unifying gamete competition and gamete limitation in a single model of anisogamy evolution. *Behav. Ecol. Sociobiol.* 65, 445–459.
- Newman, M., 2010. *Networks: an introduction.* Oxford University Press, Oxford.
- Parker, G. A., Baker, R.R., Smith, V.G., 1972. The origin and evolution of gamete dimorphism and the male-female phenomenon. *J. Theor. Biol.* 36, 529–553.
- Parker, G.A., Birkhead, T.R., 2013. Polyandry : the history of a revolution. *Philos. Trans. R. Soc. Biol. Sci.* 368, 20120335.
- Shuster, S.M., Wade, M.J., 2003. *Mating systems and strategies.* Princeton University Press, Princeton.

# Capítulo 1

## A sexual network approach to sperm competition in a species with alternative mating tactics<sup>2</sup>

Danilo G. Muniz<sup>1,2</sup>, Paulo R. Guimarães Jr.<sup>2</sup>, Bruno A. Buzatto<sup>3</sup> & Glauco Machado<sup>2</sup>



1. Programa de Pós-graduação em Ecologia, Departamento de Ecologia, Instituto de Biociências, Universidade de São Paulo, São Paulo, Brazil
2. LAGE do Departamento de Ecologia, Instituto de Biociências, Universidade de São Paulo, São Paulo, Brazil
3. Centre for Evolutionary Biology, School of Animal Biology, The University of Western Australia, Crawley, WA, Australia

---

<sup>2</sup> Manuscrito formatado de acordo com as normas da revista *Behavioral Ecology*.

## **Abstract**

Alternative mating tactics are common among species exhibiting resource defense polygyny. While large territorial males aggressively defend harems, small sneaker males generally invade these harems to mate furtively. The result is a sexual network that provides information on the sperm competition intensity (SCI) faced by males of both morphs. Here we use metrics derived from the network approach to compare SCI between sneakers and territorials of the male-dimorphic harvestman *Serracutisoma proximum*. We also tested hypotheses about the influence of harem size and spatial distribution of harems on the SCI faced by territorial males. Sneakers faced, on average, higher levels of SCI than territorials, while the SCI faced by territorials was more variable than that of sneakers. Owners of large harems faced less intense sperm competition than owners of small harems, suggesting that sperm competition is more diluted among females in large harems. At the population level, sneakers concentrated their invasions on neighboring harems that are spatially aggregated. We argue that the spatial distribution of harems is an important element influencing the topology of the sexual network, and that the spatially-explicit approach we used here can bring new insights to the study of sperm competition and mating systems in a wide range of organisms.

**Key-words:** mating system, Opiliones, resource defense polygyny, sexual selection, social network, spatial distribution

## **Introduction**

In polygynous mating systems, a few males often monopolize access to most females, while most males have no access to females. In these systems, it is common for some males to adopt alternative mating tactics, such as female mimicry and/or sneaking into territories to copulate (Shuster and Wade 2003). Males adopting alternative mating tactics may be younger or smaller individuals, but several species present alternative male morphs (Oliveira et al. 2008). In these species, there is usually a large and armed morph known as major or territorial that defends territories and/or females, and a smaller and generally unarmed morph known as minor or sneaker that does not defend territories or females (Dominey 1984; Gross 1996; Oliveira et al. 2008). The presence of males adopting alternative mating tactics increases the likelihood that females will mate with multiple males, creating the opportunity for the sperm of two or more males to overlap inside the female's reproductive tract and thus compete for the fertilization of ova. This process, known as sperm competition, is widely recognized as an important process modulating the total strength and direction of sexual selection (Kvarnemo and Simmons 2013). In species with different male morphs, sneakers are frequently subject to sperm competition, since they usually mate with already mated females. Territorials, on the other hand, are subject to sperm competition only when sneakers invade their harems (Parker 1990).

Sperm competition has been quantified in two different and complementary ways: sperm competition risk and sperm competition intensity (Simmons 2001). Sperm competition risk is the probability of facing sperm competition resulting from female promiscuity, while sperm competition intensity is determined by the number of different males engaged in competition for a single set of ova (Simmons 2001). Theoretical models predict that within-species relative ejaculate investment should increase with increasing sperm competition risk and decrease with increasing sperm competition intensity (Parker et al. 1996, 2013; review in Kelly and Jennions 2011). In species with different male morphs, the prediction is that there should be an asymmetry in ejaculate investment between morphs. We would expect, therefore, sneakers to have proportionally higher ejaculate investment because they generally face higher sperm competition risk. However, if sneakers are too common or too effective in acquiring copulations, most territorials will also face sperm competition. As a result, the asymmetry in sperm competition disappears, and the prediction is of no difference in ejaculate investment between male morphs (Parker 1990; Simmons 2001).

Two important factors that could influence the sperm competition intensity faced by individual territorial males are the spatial distribution of harems in the population and the number of females in the harems (harem size). Larger harems are more easily invaded because territorials cannot monopolize all females effectively (e.g., Shuster 1987; Byers and Kitchen 1988; Munguía-Steyer et al. 2012). Thus, territorials whose harems have many females should be at higher sperm competition risk and intensity than territorials with fewer females in their harems. Also, when harems are spatially aggregated and sneakers are highly vagile, the same sneaker can invade several harems sequentially, imposing sperm competition on a great number of territorials (e.g., Gross 1991; Buzatto et al. 2011). Harem size and spatial distribution may also interact, so that a territorial defending a small harem close to large and highly attractive harems may be at higher sperm competition than a territorial defending an isolated small harem.

To evaluate the influence of harem size and spatial distribution of harems on the sperm competition intensity faced by territorials, we used a sexual network approach (Fortuna et al. 2008; McDonald et al. 2013), according to which a mating population can be viewed as a network where males and females are connected by copulations (Figure 1). Each individual may be characterized by its number of connections (i.e., sexual partners) and by the number of connections of its mating partners, which brings information about the sperm competition intensity (Figure 1). Additionally, the network structure that emerges from the interactions between all individuals in the population is a quantitative descriptor of the mating system, characterizing the degree of promiscuity of both males and females (Sih et al. 2009; McDonald et al. 2013; Pinter-Wollman et al. 2014). By combining detailed behavioral observations in the field with a multi-scale network analyses, we therefore addressed questions on sperm competition intensity from the perspective of individual males and also at the population level, incorporating the spatial structure of the harems.

Our model organism was the harvestman *Serracutisoma proximum* (Opiliones: Gonyleptidae), in which dimorphic males exhibit different reproductive tactics. Large males (hereafter 'territorials') use their long and sexually dimorphic second pair of legs in fights for the possession of territories on the vegetation along river margins (Buzatto and Machado 2008). Territories are visited by one or several females, who copulate with the territorial, lay their eggs on leaves and remain caring for the clutch (Buzatto et al. 2007). Small males (hereafter 'sneakers'), on the other hand, invade territories and furtively mate with egg-guarding females. This tactic is only possible because oviposition may last up to two weeks

and, although 80–90% of the eggs are laid in the first 24 h after copulation with the territorials, sneakers may sire some offspring if they successfully copulate with females that still retain unfertilized eggs in their ovaries (Buzatto et al. 2011). Even though sperm precedence and the forms of sperm use by females in harvestmen are unknown, sneakers usually copulate with already mated females (Buzatto et al. 2011). Moreover, 82% of the females in the population lay only one clutch of eggs during their lives (Buzatto et al. 2007), thus most of the sperm competition imposed on territorials by sneakers is restricted to a single set of unfertilized eggs present in the females' ovaries.

The first hypothesis we tested was that sperm competition intensity should be higher for sneakers than for territorials. We also predict that sperm competition faced by territorials should be more variable. Considering that sneaker vagility is limited and that harems are widely scattered along river margins (Buzatto et al. 2011), our second hypothesis was that the sexual network should be poorly connected and clearly arranged in many components, i.e., groups of nodes that are mutually interconnected (Proulx et al. 2005). Given that large harems are more likely to be invaded by sneakers (Munguía-Steyer et al. 2012) and that the same sneaker can invade several harems (Buzatto et al. 2011), our third hypothesis was that both sperm competition intensity for territorials and intensity of harem invasion should increase with harem size and the proximity to other harems. Finally, our fourth hypothesis refers to the entire population and postulates that females from nearby harems should mate with the same sneakers, thus establishing sexual connections between these harems. In this sense, aggregated harems should be hotspots of sneaker invasions when compared to isolated harems.

## **Material and methods**

### *Study population and dataset*

The studied population of *S. proximum* occurs in an Atlantic Forest area at Intervales State Park (24°14'S, 48°04'W; 800 m a.s.l.), state of São Paulo, southeastern Brazil. Specifically, our data describes the population living on the vegetation flanking the stream that follows the Caçadinha track. In the season when the data were collected, the frequency of sneakers in this population was of 20% (Buzatto et al. 2011). Males were classified as sneakers or

territorials following Knell (2009) and details of the procedure can be found in Buzatto et al. (2011).

The data we used is a subset from the data of Buzatto et al. (2011), comprising the behavioral observations from the period between January and February 2005, which corresponds to the peak of *S. proximum* mating season (Buzatto et al. 2007; Buzatto and Machado 2008). These behavioral observations were conducted in a 200 m transect with marks every 10 m on the vegetation flanking the stream. The exact location of 29 harems in this transect (to the nearest meter) was determined as the distance between each harem and the beginning of the transect. Each of these harems was inspected six times a day: twice between 08:30 and 12:00 h, twice between 14:00 and 17:30 h, and twice between 20:30 and 00:00 h. In each inspection, it was recorded female mating activities (with territorials or sneakers), mate guarding (always performed by territorials), distance between males (territorials and sneakers) and all females in the harem, and harem invasions (for more details on the sampling procedure see Buzatto et al. 2011).

In total, the behavioral observations comprised nearly 260 hours, and included information on 47 males (13 of which are sneakers) and 72 females that oviposited in the 29 harems monitored in the transect. For two females it was not possible to identify the territorial male with which they copulated, and thus these females were ignored in the harem level analyses, but not in the individual level analyses (see below). One of these females was observed near a territorial male only once, and thus her status as part of his harem could not be confirmed. The other female was never observed near a territorial male, and it is possible that she mated with a territorial male that died prior to the field observations.

### *Sexual networks*

A network is a set of elements, known as nodes, with connections between them, known as links (Newman 2003; Costa et al. 2007). To perform a multi-scale analysis, going from the individual to the population level, we used three networks describing different aspects of the mating system of *S. proximum* (Figure 1). Although the definition of nodes and links in these three networks differ, we will refer to them collectively as *sexual networks*. In the first network, nodes are males and females, and links represent the probability of copulation (Figure 1A). We used this network to investigate the patterns of sperm

competition intensity faced by individual males, including both territorials and sneakers. In the second network, nodes are harems and invader males, and a link represents the invasion of a harem by a male (Figure 1B). In this second network, used to investigate harem invasion, the invader males were usually sneakers, but at least some territorials invading neighboring territories also established sexual connections between harems. Finally, in the third network, nodes are harems and links represent that two harems share at least one invader male (Figure 1C). We used this network to investigate the pattern of harem invasion at the population level. In what follows we elaborate on the analyses performed on each of these networks.

#### *Male-female network and sperm competition intensity*

We built a weighted network in which nodes depict males and females in the population. In a weighted network, links between individuals assume continuous values representing how strongly connected are two elements (Costa et al. 2007). We built a sexual network in which the link weight represents the probability that two individuals mated, which ranges from 0 to 1. All territorials received a link weight of 1 with all females within their respective harems because we know that *S. proximum* females only lay eggs in a territory after copulating with its owner (Buzatto and Machado 2008). If a sneaker was observed copulating with a female, the link between these individuals also received a weight of 1 because we were sure they mated (solid lines in Figures 1A-B). When a sneaker was observed within 20 cm of a given female, we considered him to have a probability  $c$  of copulating with that female (dashed lines in Figures 1A-B). Given that harvestmen are unable to perceive long-range chemical, acoustic, and visual stimuli (Willemart et al. 2009), we suppose that the minimum distance for female detection is approximately two times the length of the sneaker's second pair of sensory legs (i.e., 20 cm). Moreover, our field observations indicate that: (1) sneakers directly move towards egg-guarding females within this distance, suggesting that they are able to locate potential mating partners using close range cues, (2) all sneakers that approached egg-guarding females attempted to copulate with them, and (3) sneakers were never rejected by the females (see Buzatto et al. 2011).

If the sneaker was observed more than once near the same female, his probability of copulating increased. We derived this probability using the binomial distribution as follow:  $w_{ij} = 1 - (1-c)^{N_{ij}}$ , where  $w_{ij}$  is the link weight between the sneaker  $i$  and female  $j$ , and  $N_{ij}$  is the number of times the sneaker  $i$  was seen near (< 20 cm) female  $j$ . The value of  $w_{ij}$  increases

towards 1 with increasing values of  $N$ . The probability  $c$  of a sneaker copulating with a given female once he is observed less than 20 cm from that female was defined as 0.7 because nearly 70% of the sneaker approaches to females observed in the field resulted in successful copulation (Buzatto et al. 2011). On the few occasions in which a territorial was observed invading a territory, we considered this male to have the same chance as a sneaker of copulating with a female and calculated their strength of interaction in the same way. The probability of copulation was set to 0 if the female closest to the invader (either sneaker or territorial) was mate guarded by the resident territorial male. During mate guarding, the territorial male remains less than 20 cm from a female, periodically touching her with his long second pair of legs (see Fig. 3b in Buzatto and Machado 2008), and most of the cases of sneaker detection and repulsion by territorials occurred when females were mate guarded (Buzatto et al. 2011).

Our approach to build the male-female network incorporates the uncertainty surrounding whether some interactions between males and females indeed culminate in copulation, and is solidly based on the reproductive biology of *S. proximum*. Sneakers usually invade a harem, copulate with a female, and then promptly abandon the area (Buzatto et al. 2011). In fact, according to our field observations, from all the times that a sneaker was observed copulating with a female ( $n = 32$ ), in 26 cases he was not recorded inside the harem in the previous inspection. Moreover, from all the times that a sneaker was recorded less than 20 cm from a female ( $n = 24$ ), in 17 cases (70%) he was not recorded close the female in the subsequent inspection, and in six cases (25%) he was recorded copulating with the female in the subsequent inspection (in only one case, the sneaker remained close to the female during two consecutive inspections). These findings support the notion that mating interactions between sneakers and females are fast and have low probability of being detected. Therefore, considering only the copulations actually observed in the field would greatly underestimate the role of sneakers on the network topology. To assess the sensitivity of our results to copulation probability, we performed all the analyses described below varying the value of  $c$  from 0 to 1 (with regular intervals of 0.1). Using the value 0 we are considering only the actually observed copulations in the networks, and all other values of  $c$  consider that a sneaker observed close to a female has some chance of copulating ranging from 10% to 100%.

In a weighted network, the *strength* of an individual node is the sum of the weights of all its links (Costa et al. 2007). In a sexual network context, the strength of an individual can

be interpreted as a measure of its mating success or promiscuity (McDonald et al. 2013). Given that sperm competition arises when a female copulates with more than one male within a limited period of time, allowing the ejaculates to overlap inside the female's reproductive tract, the strength of females in our analysis is an estimator of sperm competition. Hence, the mean strength of the partners of a male is an estimate of the sperm competition intensity faced by that male. This metric is known as *mean neighbor strength* in the network approach (Costa et al. 2007), but as a mnemonic device we will call it SCI (after 'sperm competition intensity') from now on.

To test the first hypothesis that sperm competition risk should be higher for sneakers and that its variation should be higher for territorials, we calculated the SCI of all males in the network. We compared the mean SCI of territorials and sneakers using a Welch's t-test, and compared the variances in SCI using a variance ratio test (Zar 2010). To test the hypothesis that SCI for territorials should increase with harem size and the distance to other harems, we calculated the mean distance of each harem to the five closest neighboring harems. The number five is an arbitrary choice, but the results do not change qualitatively with the number of harems varying from three to ten (Supplementary Material S1). We evaluated the effect of harem size and distance to other harems on SCI using a multiple linear regression. We standardized all predictor variables, thus regression coefficients are comparable between each other and can be interpreted as effect sizes (Schielzeth 2010).

To test the second hypothesis that the male-female network should be poorly connected due to spatial constraints and arranged in many components, we conducted a randomization analysis. We compared the number of components of the observed network with the number of components of an ensemble of 5,000 simulated networks generated by a null model. To build the null model, we kept the number of connections of each sneaker and assumed that he could be connected with any female in the population. Using this approach, we created a scenario where there is no constraint in the capacity of individual sneakers to impose sperm competition to any territorial male in the population, i.e., there was no effect of the spatial distribution of harems on the sneaker copulations. For each simulated network we counted the number of components, generating a null distribution of this metric. We calculated the p-value as the proportion of simulated networks in which the number of components was equal or higher than in the observed network. We ran this analysis in two different ways: firstly considering only the observed copulations as links, and secondly including both the observed and the probabilistic copulations.

### *Harem-invader network and harem invasion*

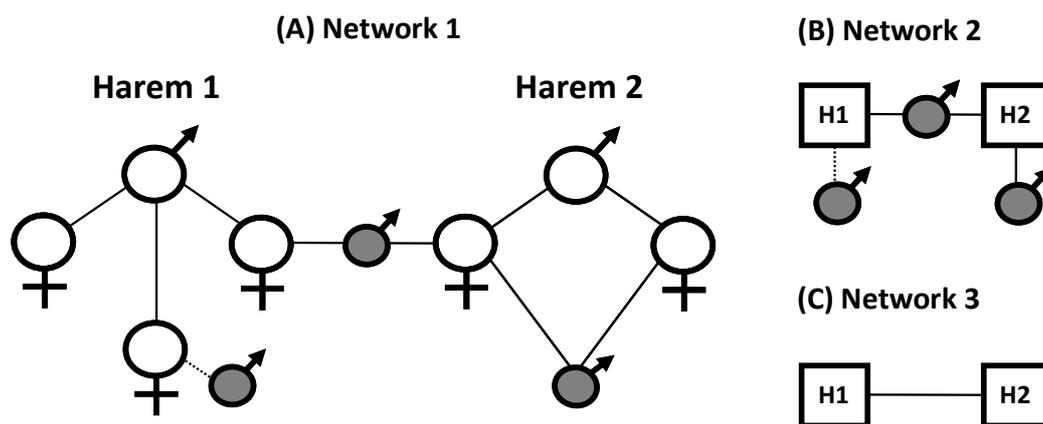
Based on the male-female network, we built a second network where the nodes were harems and invader males (Figure 1B). In this network, a harem and an invader male were connected if the invader had a link weight higher than zero with at least one female in that harem. We defined invader male as any male that was not the territorial resident of a harem, so that both sneakers and territorials from neighboring territories could be scored as invaders. We calculated the link weight  $w_s$  between a harem and an individual male as the sum of the link weights  $w_f$  for all  $g$  females with which the invader was connected according to:  $w_s = \sum_{f=1}^g w_f$ . Therefore, the strength of each harem is a measure of the intensity of harem invasion and is calculated as the sum of all its link weights.

To test the hypothesis that the intensity of harem invasion should increase with harem size and the proximity to other harems, we used again the mean distance to the five closest harems as a measure of distance to other harems. Then we used a multiple linear regression and standardized the predictor variables as described in the previous topic.

### *Harem-harem network*

To investigate the effect of the spatial distribution of harems on the connections between them, we built another network in which the nodes were the harems. In this network, two harems were connected when at least one male mated with females in the two harems (Figure 1C). The link weight ( $w_{kl}$ ) between harems  $k$  and  $l$  was calculated with two different approaches. In the first approach, we used the equation  $w_{kl} = \sum_{s=1}^m (w_{sk} + w_{sl}) \cdot a_{kl}$ , where  $w_{sk}$  and  $w_{sl}$  are the link weights between the male  $s$  and the harems  $k$  and  $l$ . The term  $a_{kl}$  is equal to 1 if  $w_{sk} > 0$  and  $w_{sl} > 0$ , and zero otherwise. Thus, the link weight between the harems  $k$  and  $l$  is the sum of  $w_{sk}$  and  $w_{sl}$  for all  $m$  shared males between the two harems. This first link weight indicates how strongly connected are two harems, since it increases with the number of shared males and the number of females with which these males are connected to. In the second approach, we built an unweighted network, in which we considered two harems to be connected if at least one invader male had a  $w_s > 0$  in both harems. Thus, in this second approach a link indicates if two harems share at least one invader male.

To test the hypothesis that spatially aggregated harems should share the same invaders, we built a matrix of distances between all harems in the population. We calculated the distance between each pair of harems as the length (in meters) of the shortest path between the two harems through the vegetation at the river margins. When two harems were located on different sides of the river we considered the distance between the margins to be 5 m (Buzatto and Machado 2008) and considered that the individuals could only cross to the other side in places where the foliage connects the two margins (see scheme in Figure 2). Since a network can be represented as a matrix, we tested our hypothesis using a Mantel test (Manly 2006) correlating the matrix of distances between harems and the matrix describing the harem-harem network. We used the Pearson correlation coefficient as the test statistic and ran 5,000 permutations to determine the p-value. Since differences in the way we calculated the link weight between two harems led to differences in the topology of the harem-harem network, we performed two Mantel tests, one for each topology of the harem-harem network.



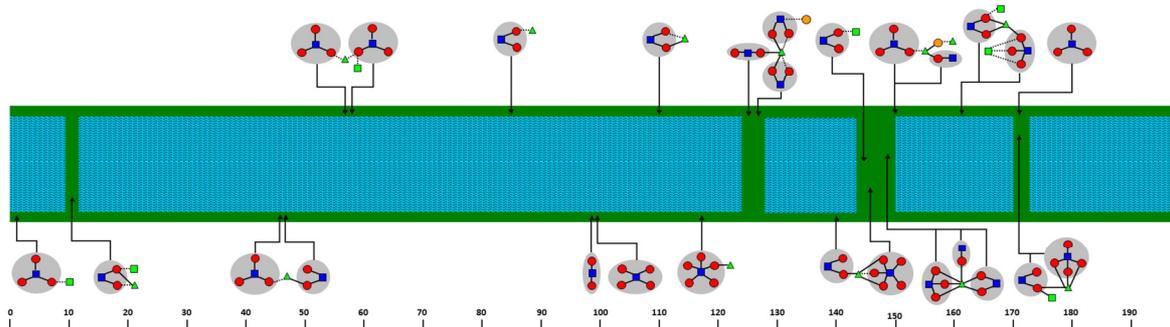
**Figure 1.** Hypothetical sexual network depicted in the three different ways used in our study. (A) Network 1: male-female network in which females and territorials are depicted in white and sneakers in grey. This is the most detailed network, from which we can extract information on sperm competition at the individual level. (B) Network 2: a harem-invader network, in which each square is a harem and sneakers are again depicted in grey. In this network, we lose information of which females are connected to each male, but it becomes easier to evaluate harem invasion. (C) Network 3: harem-harem network in which each harem is depicted as a square. In this network, we retain only the information about males that establish sexual connections between harems. In the networks (A) and (B), solid lines depict cases in which males were actually observed copulating with the females. Dotted lines depict cases in which invaders were seen close to the females, but no copulation was observed. In networks (B) and (C), H1 and H2 represent harem 1 and harem 2, respectively.

We performed all calculations and analyses described above using the software R 3.0.1 (R Development Core Team 2013).

## Results

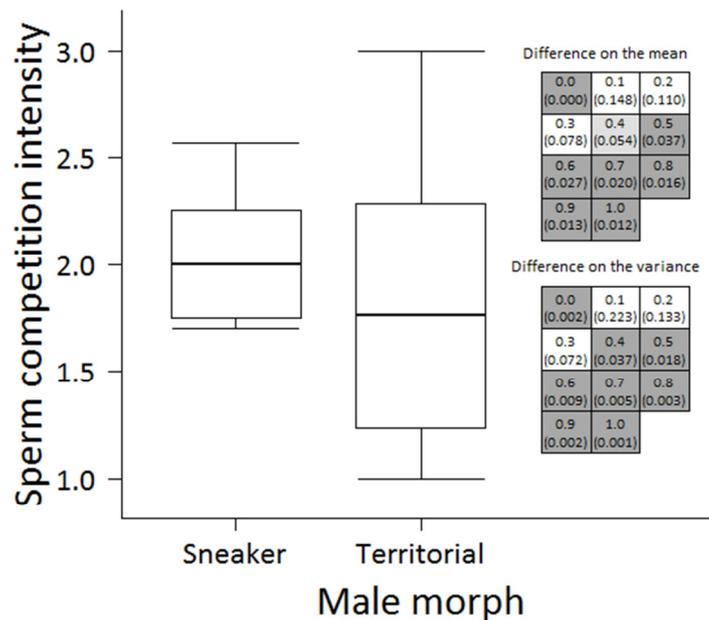
### *Male-female network: sperm competition intensity*

The male-female network of *S. proximum* was composed of 27 harems (Figure 2). Nearly 91% of the territorials (31 out of 34) and all sneaker males faced some level of sperm competition (i.e.,  $SCI > 1$ ). Among males holding territories, 89% (23 out of 27) faced some sperm competition. The mean SCI of sneakers was 13% higher than that of territorials when  $c = 0.7$  ( $t = 2.11$ ,  $df = 42.39$ ,  $p = 0.02$ ; Figure 3). The same qualitative result was obtained when  $c = 0$  and when  $c \geq 0.5$  (Figure 3; Supplementary Material S2). The standard deviation of territorials was 100% higher than that of sneakers when  $c = 0.7$  ( $F = 0.23$ ,  $df = 12; 33$ ,  $p = 0.005$ ; Figure 3). Once more, the same qualitative result was obtained when  $c = 0$  and when  $c \geq 0.4$  (Figure 3; Supplementary Material S2).



**Figure 2.** Spatially explicit schematic view, from above, of the male-female network of *Serracutisoma proximum* where each dot represents an individual: circles = females, triangles = sneakers, and squares = territorials. The red circles depict females in harems where a territorial owner was unequivocally identified, and the orange circles depict females without an assigned territorial owner. The blue squares depict territorials that held a harem (delimited by a grey shadow), and the green squares depict territorials that did not hold a harem and acted as invaders. Solid lines connecting males and females represent cases in which copulations were observed, and dotted lines represent cases in which invaders were seen close ( $< 20$  cm) to the females, but no copulation was observed. The specific position of each harem in the transect is indicated by an arrow. The green background represents the marginal vegetation, whereas the blue background represents the river. Notice that there are some vegetation bridges that connect the two margins of the river. The ruler at the bottom indicates the distance (meters) from the beginning of the transect. To calculate proximity between harems, we considered the shortest walking distance through the vegetation.

The SCI of territory owners was related only to harem size when  $c = 0.7$ , but the observed effect was not in the direction we predicted (Table 1): territorials with larger harems faced less intense sperm competition (Figure 4). The negative relationship between SCI and harem size was also significant when  $c \geq 0.3$ , and marginally significant when  $c < 0.3$  (Figure 4; Supplementary Material S3).



**Figure 3.** Comparison of sperm competition intensity faced by territorial and sneaker males of *Serracutisoma proximum*. The bold line represents the mean, the box represents the standard deviation, and the whiskers indicate the range. The boxes on the right show the results of the sensitivity analysis on the copulation probability when a male was observed close ( $< 20$  cm) to a female. Each box shows a copulation probability value and between parentheses the p-value. Values are shown for the tests of difference between the means and the variances. Lower significance values are highlighted with darker boxes.

#### *Male-female network: effect of the spatial structure*

Considering only copulations actually observed, the male-female network was arranged in 20 components (Figure 2). Fifteen components were exclusively composed of a territorial male and one to five females inside his harem, which may or may not have been invaded by other males (Figure 2). The other five components were composed of two ( $n = 3$ ) or three harems ( $n = 2$ ) interconnected by invader males (Figure 2). When considering the probabilistic copulations ( $c > 0$ ), the harems were arranged in 17 components (Figure 2), being eight of them composed of two ( $n = 6$ ) or three harems ( $n = 2$ ) interconnected by invader males (Figure 2). The number of components in the observed male-female network

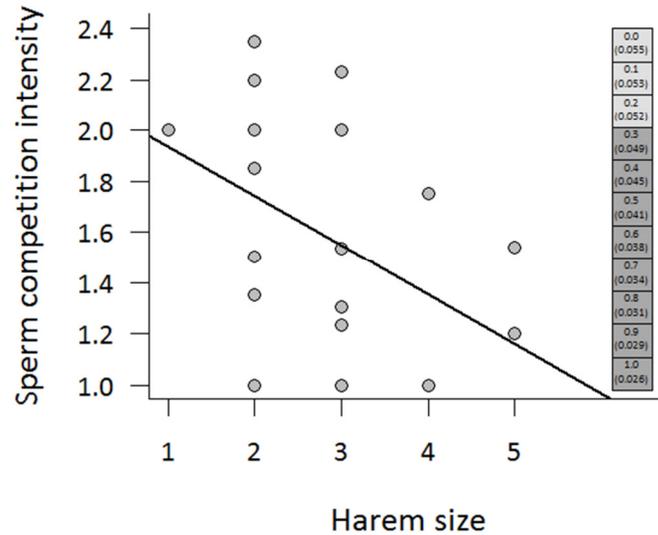
was higher than expected by our null model, both when we considered only the observed copulations (median = 14 components; range = 11-19;  $p < 0.001$ ) and when we included the probabilistic copulations (median = 10 components; range = 4-16;  $p < 0.001$ ).

**Table 1.** Summary of the linear models relating sperm competition intensity (SCI) faced by territorial males of the harvestman *Serracutisoma proximum* and harem invasion to the mean distance to the five closest harems (D) and to the harem size (H). D x H is the interaction between distance and harem size. SCI was estimated as the mean neighbor strength in a male-female sexual network. Harem invasion was estimated as the strength of each harem in a harem-invader sexual network. Predictor variable coefficients are presented as estimate  $\pm$  standard error.

Response variable	Predictor	Coefficient	t value	p
<b>Sperm competition</b>	Intercept	1.63 $\pm$ 0.07	22.52	<0.01
	H	-0.19 $\pm$ 0.09	-2.21	0.04
	D	-0.12 $\pm$ 0.07	-1.6	0.12
	H:D	-0.09 $\pm$ 0.12	-0.76	0.46
<b>Harem invasion</b>	Intercept	1.49 $\pm$ 0.19	7.95	<0.01
	H	-0.02 $\pm$ 0.22	-0.11	0.92
	D	-0.27 $\pm$ 0.19	-1.41	0.17
	H:D	-0.46 $\pm$ 0.3	-1.55	0.13

#### *Harem-invader network*

From all harems where we could identify a territorial resident, 24 (88.9%) were invaded by other males, 13 of them belonging to the sneaker morph and seven to the territorial morph (Figure 2). Individual sneakers invaded up to three harems, but did not necessarily mate with all females within each harem (Figure 2). No harem was invaded by more than one sneaker, but the same harem could be invaded by a sneaker and a territorial male ( $n = 4$ , Figure 2). The only female that mated with two sneakers was in one of the two harems where we could not identify the territorial resident. For  $c = 0.7$ , harem invasion was not correlated with harem size or with the distance to other harems (Table 1). However, for  $c \leq 0.3$ , harem invasion was negatively correlated with the distance to other harems.



**Figure 4.** (A) Relationship between sperm competition intensity faced by territorial males of the harvestman *Serracutisoma proximum* and harem size (number of females per harem). The line represents predicted values from the model presented on Table 1. The boxes on the right show the results of the sensitivity analysis on the copulation probability when a male was observed close (< 20 cm) to a female. Each box shows a copulation probability value and between parentheses the p-value for the relationship between sperm competition intensity and harem size.

#### *Harem-harem network*

The mean ( $\pm$ SD) distance between harems in the population was  $64.9 \pm 51.5$  m, but the mean distance between harems that shared at least one invader male was only  $3.8 \pm 6.9$  m. All eight males that connected harems were sneakers (Figure 2). For all values of  $c$ , the matrix of distances among harems was negatively correlated to the two matrices of harem-harem network (Table 2). Thus, nearby harems were more likely to be invaded by the same males and become sexually connected.

**Table 2.** Results of the Mantel test correlating the matrix of distances among harems and two matrices of representing the harem-harem network. In the network 1, the link weight between two harems indicates how strongly connected by invader males are two harems, and in the network 2, the link weight indicates if two harems share at least one invader male. Copulation probability indicates the probability of a sneaker copulating with a female once he invades a territory, but no actual copulation is observed. All correlations are highly significant ( $p < 0.001$ ).

Copulation probability ( <i>c</i> )	Pearson coefficient ( <i>r</i> )	
	Network 1	Network 2
<b>0.0</b>	-0.19	-0.19
<b>0.1</b>	-0.2	-0.22
<b>0.2</b>	-0.2	-0.22
<b>0.3</b>	-0.21	-0.22
<b>0.4</b>	-0.21	-0.22
<b>0.5</b>	-0.21	-0.22
<b>0.6</b>	-0.21	-0.22
<b>0.7</b>	-0.21	-0.22
<b>0.8</b>	-0.21	-0.22
<b>0.9</b>	-0.21	-0.22
<b>1.0</b>	-0.21	-0.22

## Discussion

Here we used a network-based approach to investigate the influence of harem size and spatial distribution of harems on the sperm competition intensity (SCI). We used a network metric based on the mean neighbor strength of each male as a proxy of SCI, and found that average SCI faced by sneakers of the harvestman *S. proximum* was slightly higher than that faced by territorials. However, the variation in the SCI was much higher for territorials than for sneakers. These results support both predictions of our first hypothesis. We also found that the male-female network is arranged in many isolated components, supporting our second hypothesis that there is a strong effect of the spatial distribution of harems on the sneaker copulations. Owners of large harems faced less SCI than owners of small harems, which contrasts with our third hypothesis that predicted that SCI for territorials should increase with harem size. Additionally, the distance to other harems had no influence on the SCI and on the intensity of harem invasion. However, sneakers concentrated their invasions on neighboring harems, which supports our fourth hypothesis that predicted that aggregated harems should be hotspots of sneaker invasions when compared to isolated harems. These results emerge when we build the male-female network including as links only the copulations actually observed in the field, and also most values of copulating probability. In what follows, we explore the implications of these findings in terms of sperm competition and sexual selection in mating systems with alternative mating tactics.

According to our field observations, 89% of territory owners experienced some level of sperm competition, despite the fact that sneakers composed only 20% of the males in the population (Buzatto et al. 2011). Therefore, even a small percentage of sneakers in the population may increase the SCI faced by territorials. In fact, the average SCI for territorial males was 1.76 (Figure 3), suggesting that most females in the harems mate polyandrously. When the probability of sneak copulations is high, sperm competition is expected to be intense for both territorials and sneakers, and asymmetries in gonadal investment should be low between the two morphs (Parker 1990; Simmons 2001). A previous study with several populations of *S. proximum* (including the one we studied here) indeed showed that gonadal investment did not differ between morphs (Munguía-Steyer et al. 2012). Moreover, gonadal investment was not influenced by the relative frequency of sneakers, suggesting that population-level measures may not be appropriate proxies of SCI in *S. proximum*, and perhaps other species with similar mating systems (Munguía-Steyer et al. 2012).

Contrary to Munguía-Steyer et al. (2012), our results do not indicate that larger harems are more likely to be invaded. This difference probably occurs because our definition of invasion includes only cases in which an invader male was observed less than 20 cm from a female, whereas Munguía-Steyer et al. (2012) defined invasion as any male found less than 1 m from a female. Although our definition is more rigorous, our results still support the prediction made by Munguía-Steyer et al. (2012) that SCI should be more heterogeneous among territorials (Figure 3). Territorials should, therefore, respond strategically adjusting their ejaculate expenditure in response to perceived SCI (Simmons 2001). Based on the results presented here (Figure 4), and also on the theory of strategic ejaculation (Parker et al. 1996), the increased intensity of sperm competition faced by males holding small harems should result in decreased sperm numbers transferred to newcomer females. However, although the pattern predicted by the theoretical model has already been observed in some species (e.g., Simmons and Kvarnemo 1997), a recent meta-analysis of strategic ejaculation indicates that ejaculate expenditure responds positively to sperm competition risk, but not to SCI (Kelly and Jennions 2011). The contrast between theory and empirical results is probably related to the fact that the underlying assumptions of the models are rarely checked in the studied species (Kelly and Jennions 2011). As we stressed before, sperm precedence and the forms of sperm use by females in harvestmen are unknown, and without this information it is not possible to robustly test hypotheses of strategic ejaculation in the group.

The use of a sexual network approach circumvents the assumption that sperm competition is evenly distributed among all males in the population, allowing us to investigate the factors that influence inter-individual variation in SCI. We showed that territorials owning large harems experienced less SCI (Figure 4). The simpler explanation to this unexpected finding is that not all females of a large harem copulate with invaders (Figure 2). Thus, SCI seems to be diluted in large harems, which leads to the prediction that mean paternity per clutch should increase as harem size increases. This prediction contrasts with the pattern reported for some mammals, in which males holding the largest harems do not have the highest reproductive success (e.g., Kaseda and Khalil 1996; Heckel et al. 1999). Since mammals usually give birth to one or few offspring at a time, whereas arthropods lay a clutch containing many eggs, the negative impact of multiple paternity on the reproductive success of harem holding males should be different in these two groups. In mammals, each individual offspring sired by an invader promotes great reduction in the reproductive success of the harem holding male. In arthropods, invaders may sire only a small fraction of the eggs laid by each female. Assuming a scenario where (i) invaders are unable to copulate

with all females in a harem, (ii) an increase in harem size does not lead to an increase in the number of invaders, and (iii) successful males do not suffer from sperm depletion, we expect that the mean proportion of eggs sired by territorial males should increase as harem size increases. In a harem containing one female, for instance, an invader that sires 15% of the eggs promotes a decrease of 15% in the territorial reproductive success. In a harem containing three females, an invader that sires 15% of the eggs of one female promotes a decrease of only 5% in the territorial reproductive success (considering that all females lay the same number of eggs).

The presence of alternative mating tactics allegedly decreases the total intensity of sexual selection among males, since the number of males without any copulation decreases (Webster et al. 1995; Jones et al. 2001; Shuster and Wade 2003). However, the influence of the sneaker tactic on the total intensity of sexual selection can follow different directions depending on which females the sneakers copulate with (McDonald et al. 2013). Based on a previous study of *S. proximum* (Munguía-Steyer et al. 2012), we expected that harem invasion would be positively correlated with harem size, so that the most successful territorials would also be the ones facing the strongest sperm competition, which weakens the total intensity of sexual selection (Sih et al. 2009; McDonald et al. 2013). However, as explained above, territorials with larger harems faced lower SCI (Figure 4). In this sense, if territorials with smaller harems are also the ones losing more fertilizations to sneakers, the presence of the alternative tactic in *S. proximum* may indeed increase the total intensity of sexual selection among territorials. Further quantifications based on the number of offspring sired by each male in the population are necessary to test this hypothesis.

Contrary to our expectations, both the SCI faced by individual territorials and the intensity of harem invasion were not influenced by the distance to other harems. Even the most isolated harems were found and eventually invaded at least once (Figure 2), suggesting that sneakers are highly effective in detecting harems. However, the number of harems invaded by each sneaker was limited by the spatial distribution of harems, so that only neighbouring harems were invaded by the same individual sneakers (Figure 2). The result was a poorly connected sexual network, divided in more components than would be expected if there was no spatial constraint on sneaker movements. Most of the invader males were sneakers, but seven harems were invaded by males that unequivocally belonged to the territorial morph (Figure 2). These territorials could be defending territories without females close to the invaded harems and, at some point of the breeding season, they adopted the

sneaker mating tactic after failing to attract females. This implies behavioral plasticity that allows territorials to switch between territorial defense and invasion. Similarly, in damselflies and dragonflies, the costs of territoriality may only be affordable to attractive males that are large and/or in good condition (e.g., Forsyth and Montgomerie 1987; Tsubaki and Ono 1986, 1987).

At the population level, we found a negative correlation between the distance between any two harems and the probability of these harems being sexually connected by sharing at least one invader male (Table 2, Network 2). Given that some harems are spatially aggregated along the river margins (Figure 2), once a sneaker finds a harem, he probably invests some time searching for additional mates in the nearby areas. Therefore, the spatial distribution of harems in the population plays an important role influencing the topology of the sexual network, as has been previously demonstrated for an insect-pollinated plant in which close mother trees shared the same pollen donors (Fortuna et al. 2008). In resource defense mating systems, the proximity to other harems may increase the number and strength of sexual connections in the network. In fact, our results supports this prediction and show that spatially aggregated harems are more strongly connected (Table 2, Network 1), which may have important implications for the spread of sexually-transmitted diseases (Proulx et al. 2005).

In conclusion, the network approach we employed allowed us to investigate SCI at multiple scales, ranging from individual males to the whole population. To the best of our knowledge, this is the first empirical study to use a spatially-structured sexual network to make inferences on sperm competition. By incorporating the spatial structure of the population in studies of sperm competition, we can start thinking about a sperm competition landscape in which harems are analogous to habitat patches. In landscape ecology, theory predicts that the closer the patches, the more intense the transit of individuals among them if the matrix is permeable (Forman 1995). In the same way, the more aggregated the harems, the more likely invasions by sneakers should be. Permeability in this case could be viewed as the mate-searching costs (especially mortality) paid by sneakers while moving among harems. Moreover, in the same way that large patches attract more immigrants and bear populations with higher genetic variability (Forman 1995), large harems may attract more invaders and increase the opportunity for females to mate polyandrously, which could bring benefits such as increased genetic variability in the offspring (Arnqvist and Nilsson 2000; Kvarnemo and Simmons 2013). In this sense, it would be advantageous for females to join

large harems, which could explain some cases of female copying in species with resource defense polygyny (e.g., Shuster and Wade 1991). These are exciting new ideas that could be tested and expanded to other species exhibiting alternative mating tactics.

### **Acknowledgements**

We thank Billy Requena for helping in the field work and the staff of Intervales State Park for logistical support. We also thank Paulo I. Prado, Eduardo S. A. Santos, Roberto Munguía-Steyer, Rodolfo Jaffé, Paco Garcia-Gonzalez, and two anonymous reviewers for helpful comments on early versions of the manuscript. All authors were sponsored by fellowships from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP: 2002/00381-0, 2003/05427-0, 2011/12795-2, and 2012/50229-1). GM and PRG have research grants from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

### **References**

- Arnqvist G, Nilsson T. 2000. The evolution of polyandry: multiple mating and female fitness in insects. *Anim Behav.* 60: 145–164.
- Buzatto BA, Machado G. 2008. Resource defense polygyny shifts to female defense polygyny over the course of the reproductive season of a Neotropical harvestman. *Behav Ecol Sociobiol.* 63: 85–94.
- Buzatto BA, Requena GS, Lourenço RS, Munguía-Steyer R, Machado G. 2011. Conditional male dimorphism and alternative reproductive tactics in a Neotropical arachnid (Opiliones). *Evol Ecol.* 25: 331–349.
- Buzatto BA, Requena GS, Martins EG, Machado G. 2007. Effects of maternal care on the lifetime reproductive success of females in a neotropical harvestman. *J Anim Ecol.* 76: 937–45.
- Byers J, Kitchen D. 1988. Mating system shift in a pronghorn population. *Behav Ecol Sociobiol.* 22: 355–360.
- Costa LDF, Rodrigues FA, Travieso G, Villas Boas PR. 2007. Characterization of complex networks: a survey of measurements. *Adv Phys.* 56: 167–242.

- Dominey WJ. 1984. Alternative mating tactics and evolutionarily stable strategies. *Am Zool.* 24: 385–396.
- Forman, RTT. 1995. *Land mosaics: the ecology of landscapes and regions*. Cambridge (UK): Cambridge University Press.
- Forsyth A, Montgomerie RD. 1987. Alternative reproductive tactics in the territorial damselfly *Calopteryx maculata* – sneaking by older males. *Behav Ecol Sociobiol.* 21: 73–81.
- Fortuna MA, García C., Guimarães PR, Bascompte J. 2008. Spatial mating networks in insect-pollinated plants. *Ecol Lett.* 11: 490–498.
- Gross MR. 1991. Evolution of alternative reproductive strategies: frequency-dependent sexual selection in male bluegill sunfish. *Phil Trans R Soc B.* 332: 59–66.
- Gross MR. 1996. Alternative reproductive strategies and tactics: diversity within sexes. *Trends Ecol Evol.* 11: 92–98.
- Heckel G, Voigt CC, Mayer F, Helversen O. 1999. Extra-harem paternity in the white-lined bat *Saccopteryx bilineata* (Emballonuridae). *Behaviour.* 136: 1173–1185.
- Jones AG, Walker D, Kvarnemo C, Lindström K, Avise JC. 2001. How cuckoldry can decrease the opportunity for sexual selection: data and theory from a genetic parentage analysis of the sand goby, *Pomatoschistus minutus*. *Proc Natl Acad Sci USA.* 98: 9151–9156.
- Kaseda Y, Khalil AM. 1996. Harem size and reproductive success of stallions in Misaki feral horses. *Appl Anim Behav Sci.* 47: 163–173.
- Kelly CD, Jennions MD. 2011. Sexual selection and sperm quantity: meta-analyses of strategic ejaculation. *Biol Rev.* 86: 863–884.
- Knell RJ. 2009. On the analysis of non-linear allometries. *Ecol Entomol.* 34:1–11.
- Kvarnemo C, Simmons LW. 2013. Polyandry as a mediator of sexual selection before and after mating. *Phil Trans R Soc B.* 368: 20120042.
- Manly BFJ. 2006. *Randomization, bootstrap and Monte Carlo methods in biology*. 3rd ed. Chicago (IL): Chapman & Hall.
- McDonald GC, James R, Krause J, Pizzari T. 2013. Sexual networks: measuring sexual selection in structured, polyandrous populations. *Phil Trans R Soc B.* 368: 20120356.

- Munguía-Steyer R, Buzatto BA, Machado G. 2012. Male dimorphism of a neotropical arachnid: harem size, sneaker opportunities, and gonadal investment. *Behav Ecol.* 23: 827–835.
- Newman MEJ. 2003. The Structure and function of complex networks. *SIAM Review* 45:167–256.
- Oliveira RF, Taborsky M, Brockmann HJ. 1998. *Alternative reproductive tactics: an integrative approach.* Cambridge (UK): Cambridge University Press.
- Parker GA. 1990. Sperm competition games: sneaks and extra-pair copulations. *Proc R Soc Lond B.* 242:127–133.
- Parker G, Ball M, Stockley P, Gage MJG. 1996. Sperm competition games: individual assessment of sperm competition intensity by group spawners. *Proc R Soc Lond B.* 263: 1291–1297.
- Parker GA, Lessells CM, Simmons LW. 2013. Sperm competition games: a general model for precopulatory male-male competition. *Evolution* 67: 95–109.
- Pinter-Wollman N, Hobson EA, Smith JE, Edelman AJ, Shizuka D, Silva S, Waters JS, Prager SD, Sasaki T, Wittemyer G, Fewell J, McDonald DB. 2014. The dynamics of animal social networks: analytical, conceptual, and theoretical advances. *Behav Ecol.* 25: 242–255
- Proulx SR, Promislow DEL, Phillips PC. 2005. Network thinking in ecology and evolution. *Trends Ecol Evol.* 20:345–53.
- R Development Core Team. 2013. R: A language and environment for statistical computing. R foundation for statistical computing: Vienna. Available from: <http://www.r-project.org>.
- Schielzeth H. 2010. Simple means to improve the interpretability of regression coefficients. *Methods Ecol Evol.* 1: 103–113.
- Shuster SM. 1987. Alternative reproductive behaviors: three discrete male morphs in *Paracerceis sculpta*, an intertidal isopod from the northern Gulf of California. *J. Crust Biol.* 7: 318–327.
- Shuster SM, Wade MJ. 1991. Female copying and sexual selection in a marine isopod crustacean, *Paracerceis sculpta*. *Anim Behav.* 41: 1071–1078.

- Shuster SM, Wade MJ. 2003. Mating systems and strategies. Princeton (NJ): Princeton University.
- Sih A, Hanser SF, McHugh KA. 2009. Social network theory: new insights and issues for behavioral ecologists. *Behav Ecol Sociobiol.* 63: 975–988.
- Simmons LW. 2001. Sperm competition and its evolutionary consequences in the insects. Princeton (NJ): Princeton University Press.
- Simmons LW, Kvarnemo K. 1997. Ejaculate expenditure by male bushcrickets decreases with sperm competition intensity. *Proc R Soc Lond B.* 264: 1203-1208.
- Tsubaki Y, Ono T. 1986. Competition for territorial sites and alternative mating tactics in the dragonfly, *Nannophya pygmaea* Rambur (Odonata, Libellulidae). *Behaviour.* 97: 234-252.
- Tsubaki Y, Ono T. 1987. Effects of age and body size on the male territorial system of the dragonfly, *Nannophya pygmaea* Rambur (Odonata, Libellulidae). *Anim Behav.* 35: 518-525.
- Webster M, Pruett-Jones S, Westneat D, Arnold S. 1995. Measuring the effects of pairing success, extra-pair copulations and mate quality on the opportunity for sexual selection. *Evolution* 49: 1147–1157.
- Willemart RH, Farine JP, Gnaspini P. 2009. Sensory biology of Phalangida harvestmen (Arachnida, Opiliones): a review, with new morphological data on 18 species. *Acta Zool.* 90: 209–227.
- Zar JH. 2010. Biostatistical analysis. 5th ed. Upper Saddle River NJ: Prentice-Hall.

## Supplementary Material

**Table S1.** Summary of the linear models relating sperm competition intensity (SCI) faced by territorial males of the harvestman *Serracutisoma proximum* and harem invasion to the mean distance to the closest harems (D) and to the harem size (H). D x H is the interaction between distance and harem size. In this table we present the results when we vary the number of harems considered as closest harems between 3 and 10. Predictor variable coefficients are presented as estimate  $\pm$  standard error.

Number of closest harems: 3				
Response Variable	Predictor	Coefficient	t value	p
Sperm Competition	Intercept	1.63 $\pm$ 0.07	22.26	< 0.01
	H	-0.18 $\pm$ 0.08	-2.22	0.04
	D	-0.11 $\pm$ 0.07	-1.49	0.15
	H:D	-0.08 $\pm$ 0.11	-0.73	0.47
Response Variable	Predictor	Coefficient	t value	p
Harem Invasion	Intercept	1.5 $\pm$ 0.18	8.16	< 0.01
	H	0.25 $\pm$ 0.19	1.31	0.2
	D	-0.16 $\pm$ 0.19	-0.86	0.4
	H:D	-0.4 $\pm$ 0.2	-2.01	0.06
Number of closest harems: 4				
Response Variable	Predictor	Coefficient	t value	p
Sperm Competition	Intercept	1.63 $\pm$ 0.07	22.48	< 0.01
	H	-0.19 $\pm$ 0.08	-2.23	0.04
	D	-0.12 $\pm$ 0.07	-1.6	0.12
	H:D	-0.09 $\pm$ 0.11	-0.77	0.45
Response Variable	Predictor	Coefficient	t value	p
Harem Invasion	Intercept	1.49 $\pm$ 0.18	8.17	< 0.01
	H	0.24 $\pm$ 0.19	1.25	0.22
	D	-0.15 $\pm$ 0.19	-0.78	0.45
	H:D	-0.39 $\pm$ 0.19	-2.03	0.06

Number of closest harems: 6				
Response Variable	Predictor	Coefficient	t value	p
Sperm Competition	Intercept	1.63 ± 0.07	22.6	< 0.01
	H	-0.2 ± 0.09	-2.27	0.03
	D	-0.11 ± 0.07	-1.54	0.14
	H:D	-0.1 ± 0.12	-0.84	0.41
Response Variable	Predictor	Coefficient	t value	p
Harem Invasion	Intercept	1.49 ± 0.18	8.16	< 0.01
	H	0.23 ± 0.19	1.19	0.25
	D	-0.15 ± 0.19	-0.76	0.45
	H:D	-0.37 ± 0.19	-1.97	0.06
Number of closest harems: 7				
Response Variable	Predictor	Coefficient	t value	p
Sperm Competition	Intercept	1.63 ± 0.07	22.68	< 0.01
	H	-0.2 ± 0.09	-2.33	0.03
	D	-0.11 ± 0.07	-1.52	0.14
	H:D	-0.11 ± 0.12	-0.92	0.37
Response Variable	Predictor	Coefficient	t value	p
Harem Invasion	Intercept	1.49 ± 0.18	8.19	< 0.01
	H	0.22 ± 0.19	1.19	0.25
	D	-0.14 ± 0.19	-0.74	0.47
	H:D	-0.38 ± 0.19	-2.02	0.054
Number of closest harems: 8				
Response Variable	Predictor	Coefficient	t value	p
Sperm Competition	Intercept	1.63 ± 0.07	22.77	< 0.01
	H	-0.21 ± 0.09	-2.36	0.03
	D	-0.11 ± 0.07	-1.55	0.14
	H:D	-0.11 ± 0.12	-0.97	0.34
Response Variable	Predictor	Coefficient	t value	p
Harem Invasion	Intercept	1.49 ± 0.18	8.21	< 0.01
	H	0.22 ± 0.19	1.19	0.25
	D	-0.14 ± 0.19	-0.73	0.47
	H:D	-0.39 ± 0.19	-2.05	0.052

Number of closest harems: 9				
Response Variable	Predictor	Coefficient	t value	p
Sperm Competition	Intercept	1.63 ± 0.07	22.81	< 0.01
	H	-0.21 ± 0.09	-2.4	0.03
	D	-0.11 ± 0.07	-1.55	0.13
	H:D	-0.12 ± 0.12	-1.01	0.32
Response Variable	Predictor	Coefficient	t value	p
Harem Invasion	Intercept	1.49 ± 0.18	8.24	< 0.01
	H	0.22 ± 0.19	1.19	0.25
	D	-0.13 ± 0.19	-0.68	0.5
	H:D	-0.4 ± 0.19	-2.09	0.05
Number of closest harems: 10				
Response Variable	Predictor	Coefficient	t value	p
Sperm Competition	Intercept	1.63 ± 0.07	22.81	< 0.01
	H	-0.21 ± 0.09	-2.41	0.02
	D	-0.11 ± 0.07	-1.53	0.14
	H:D	-0.12 ± 0.12	-1.02	0.32
Response Variable	Predictor	Coefficient	t value	p
Harem Invasion	Intercept	1.5 ± 0.18	8.25	< 0.01
	H	0.22 ± 0.19	1.2	0.24
	D	-0.13 ± 0.19	-0.66	0.51
	H:D	-0.4 ± 0.19	-2.1	0.05

**Table S2.** Results of the sensitivity analysis to the value of the copulation probability  $c$ , which represents the probability of an invader male copulating with a female once he was observed less than 20 cm from that female. *Part 1:* Summary of tests for difference on the mean and variance in sperm competition intensity between territorial and sneaker males of the harvestman *Serracustoma proximum*. The table presents test statistics with significance values between parentheses.

<b>Copulation probability (<math>c</math>)</b>	<b>Difference on the mean (t value)</b>	<b>Difference on the variance (F value)</b>
0.0	4.394 (<0.001)	0.137 (0.002)
0.1	1.067 (0.148)	0.658 (0.223)
0.2	1.252 (0.11)	0.547 (0.133)
0.3	1.455 (0.078)	0.453 (0.072)
0.4	1.652 (0.054)	0.375 (0.037)
0.5	1.831 (0.037)	0.313 (0.018)
0.6	1.985 (0.027)	0.266 (0.009)
0.7	2.112 (0.02)	0.231 (0.005)
0.8	2.213 (0.016)	0.206 (0.003)
0.9	2.288 (0.013)	0.188 (0.002)
1.0	2.341 (0.012)	0.178 (0.001)

**Table S3.** Results of the sensitivity analysis to the value of the copulation probability  $c$ , which represents the probability of an invader male copulating with a female once he was observed less than 20 cm from that female. *Part 2:* Summary of the linear models relating sperm competition intensity (SCI) faced by territorial males of the harvestman *Serracutisoma proximum* to the mean distance to the closest harems (D) and to the harem size (H). D x H is the interaction between distance and harem size. In this table we present the regression slopes with significance values between parentheses.

<b>Copulation</b>			
<b>probability (<math>c</math>)</b>	<b>H</b>	<b>D</b>	<b>DxH</b>
0.0	-0.158 (0.055)	-0.206 (0.004)	-0.028 (0.786)
0.1	-0.158 (0.053)	-0.195 (0.006)	-0.036 (0.724)
0.2	-0.162 (0.052)	-0.182 (0.011)	-0.044 (0.675)
0.3	-0.167 (0.049)	-0.168 (0.02)	-0.052 (0.628)
0.4	-0.173 (0.045)	-0.154 (0.034)	-0.06 (0.581)
0.5	-0.18 (0.041)	-0.141 (0.056)	-0.069 (0.533)
0.6	-0.188 (0.038)	-0.128 (0.086)	-0.079 (0.485)
0.7	-0.196 (0.034)	-0.116 (0.126)	-0.09 (0.438)
0.8	-0.205 (0.031)	-0.104 (0.176)	-0.102 (0.393)
0.9	-0.214 (0.029)	-0.094 (0.235)	-0.114 (0.35)
1.0	-0.224 (0.026)	-0.084 (0.299)	-0.128 (0.31)

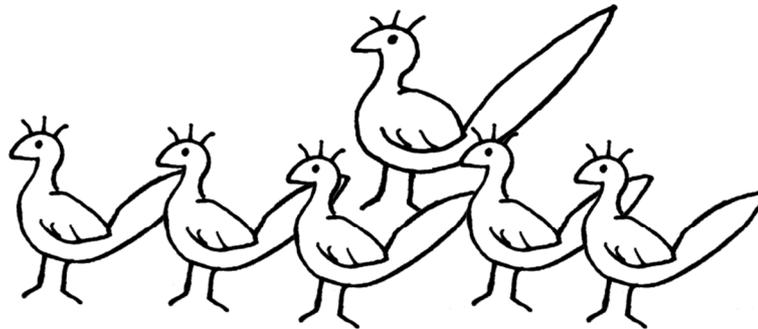
**Table S4.** Results of the sensitivity analysis to the value of the copulation probability  $c$ , which represents the probability of an invader male copulating with a female once he was observed less than 20 cm from that female. *Part 3:* Summary of the linear models relating harem invasion on the harvestman *Serracutisoma proximum* to the mean distance to the closest harems (D) and to the harem size (H). D x H is the interaction between distance and harem size. In this table we present the regression slopes with significance values between parentheses.

<b>Copulation</b>			
<b>probability (<math>c</math>)</b>	<b>H</b>	<b>D</b>	<b>DxH</b>
0.0	-0.057 (0.772)	-0.481 (0.007)	-0.374 (0.153)
0.1	-0.04 (0.841)	-0.456 (0.01)	-0.385 (0.146)
0.2	-0.032 (0.875)	-0.423 (0.02)	-0.393 (0.151)
0.3	-0.03 (0.888)	-0.388 (0.035)	-0.401 (0.152)
0.4	-0.03 (0.889)	-0.354 (0.058)	-0.412 (0.151)
0.5	-0.032 (0.882)	-0.321 (0.09)	-0.425 (0.147)
0.6	-0.037 (0.869)	-0.29 (0.131)	-0.44 (0.142)
0.7	-0.043 (0.849)	-0.261 (0.181)	-0.458 (0.135)
0.8	-0.052 (0.825)	-0.234 (0.238)	-0.479 (0.126)
0.9	-0.062 (0.797)	-0.209 (0.3)	-0.503 (0.117)
1.0	-0.073 (0.765)	-0.187 (0.364)	-0.529 (0.108)

## Capítulo 2

# A spatially explicit model of mate choice and assortative mating<sup>3</sup>

Danilo G. Muniz<sup>1,2</sup>, Eduardo S. A. Santos<sup>2,3</sup>, Paulo R. Guimarães Jr<sup>2</sup>, Shinichi Nakagawa<sup>4</sup>  
and Glauco Machado<sup>2</sup>



1. Programa de Pós-Graduação em Ecologia, Departamento de Ecologia, Instituto de Biociências, Universidade de São Paulo
2. LAGE do Departamento de Ecologia, Instituto de Biociências, Universidade de São Paulo
3. Departamento de Zoologia, Instituto de Biociências, Universidade de São Paulo
4. Inter-disciplinary Ecology and Evolution Lab, School of Biological, Earth and Environmental Sciences, University New South Wales

---

<sup>3</sup> Capítulo formatado de acordo com as regras da revista *Methods in Ecology and Evolution*.

## Abstract

1. Mate sampling, the fact that individuals cannot access all potential mating partners in the population, is a ubiquitous yet poorly explored process in nature. Ignoring mate sampling may greatly underestimate measures of choosiness. One of the main factors promoting mate sampling is the spatial distribution of individuals in the population. Thus, including distances in the models of mate choice should increase the accuracy of choosiness estimates. Spatial distances between individuals, however, are pairwise variables and cannot be included in the models commonly used to investigate mate choice.
2. We propose a multinomial model of mate choice and review a previously published binomial model. Both models allow for the inclusion of pairwise and non-pairwise predictor variables. Thus they can be used to include the spatial distances between individuals in the analysis of mate choice and also to investigate assortative mate choice. We evaluated the performance of the binomial and multinomial models in detecting directional and assortative mate choice using datasets generated by spatially explicit individual based simulations. We also tested the effect of spatial autocorrelation of phenotypic traits on the performance of both models.
3. Our results show that spatially explicit models of mate choice are more accurate and powerful than models ignoring space. In the analysis of directional mate choice, the multinomial model was more powerful and precise than the binomial, especially when there was spatial autocorrelation of male traits. In the analysis of assortative mate choice, both models were highly effective and were not affected by spatial autocorrelation of phenotypic traits.
4. We conclude that the inclusion of space in the analysis of mate choice greatly increases our ability to detect mate choice and accurately estimate its intensity using observational data. The binomial and multinomial models are powerful and flexible tools that should be used in the investigation of mate choice in spatially structured populations and in the study of assortative mate choice. In addition, these models can be employed in the investigation of choice in other contexts, such as social interactions and interspecific interactions.

**Key-words:** information filtering, mate sampling, sexual selection, social networks, spatial distribution, spatial structure

## Introduction

Mate choice is an important process promoting sexual selection and influencing the evolution of mating systems (Andersson 1994, Jones & Ratterman 2009). Females usually exert mate choice, and the criteria used in mate choice vary widely among taxonomic groups. For instance, males can be selected based on the conspicuousness of their coloration (*e.g.*, Houde 1987), size-related traits (*e.g.*, Côte & Hunte 1989, Andersson 1982), acoustic features of their calls (*e.g.*, Gerhardt 1991), the chemical composition of their cuticular hydrocarbons (*e.g.*, Thomas & Simmons 2009) or multiple suites of traits (*e.g.*, Zuk et al. 1992). Irrespective of the mate choice criterion, females cannot access all males in the population and must choose their mates from a limited sample of males (Janetos 1980, Real 1990). Mate search can be both costly and risky, regardless of the sex performing the search (*e.g.*, Polis et al. 1998; Byers et al. 2005; Lane et al. 2010). Thus, time spent performing mate search is frequently limited and the average number of potential mates sampled by a female can be very low (Roff & Fairbairn 2014). In natural populations, one of the main factors determining the number of males available as potential mates for a female is the spatial distribution of individuals, since a female can only choose to mate with a male once she encounters him (Byers et al. 2005). This limitation in female sampling establishes an information filtering process (Mossa et al. 2002), in which females must make mating decisions based on limited information.

Information filtering makes the detection of mate choice in observational studies more difficult: the smaller the sample of males taken by females, the weaker the correlation between a sexually selected male trait and a male's mating success (Benton & Evans 1998). This weakening of the correlation between the sexually selected trait and a male's mating success occurs because in some situations a female will be restricted to a sample of suboptimal males in respect to the choice criterion, and yet she will mate with a male. Although information filtering is probably very common in natural populations, it is regularly ignored in the statistical analyses of studies about mate choice (but see Schlicht et al. 2015). Therefore, the inclusion of spatial distances between males and females in the analysis of mating data would account for a major source of information filtering. However, it is not trivial to include spatial distances between individuals in the population as a predictor variable in statistical models because distance is a pairwise variable.

Pairwise variables are difficult to include in regular statistical models (such as generalized linear models, GLM) because they only make sense when considering two specific individuals. Since the sampling unit used in the analysis of mating data is typically the male, there is no way of including distances between each male and individual female in the model. This is a problem not only when trying to include spatial distance in a model of mate choice, but also when testing hypotheses about assortative mating, since the similarity between two individuals is also a pairwise variable. Assortative mating is considered to be a common pattern in nature (Jiang et al. 2013), and it can be generated by multiple processes (Burley 1983). Commonly, it is hypothesized or assumed that the pattern of assortative mating is generated by the process of individuals choosing mates similar to them, i.e., assortative mate choice. However, there is no standardized or generally accepted method for detecting or estimating assortative mate choice. One possible solution is to use randomization techniques (e.g., Freeman-Gallant et al. 2003), which are powerful, but do not allow the inclusion of multiple variables within the same test, and also may not provide comparable estimates of choosiness.

In this paper, we propose a multinomial model of mate choice that can accommodate multiple pairwise and non-pairwise predictor variables, and we also review a binomial model recently proposed by Schlicht et al. (2015). We tested these models by analyzing datasets generated by spatially explicit individual based simulations. To test the performance of the models under different preference criteria, we ran simulations in which females had directional preference for large male traits, analyzed these data using the multinomial and binomial models, and also a regular GLM in which spatial distances were not included. We also ran simulations in which females performed assortative mate choice, and analyzed these data using only the multinomial and binomial models, since there is no standardized method to investigate assortative mate choice. In addition, to test the models under the confounding effect of spatial autocorrelation, we ran simulations with both preference criteria in a scenario of spatial autocorrelation of phenotypic traits to investigate the effect of this confounding factor in the performance of the models.

We use the term *model of mate choice* to describe a probabilistic model that calculates the probability of an individual mating with another individual. In addition, we consider as *spatially explicit* any model of mate choice that includes the spatial distribution of the individuals in the calculation of these probabilities. The multinomial and binomial models discussed in this paper fulfill both of these definitions.

## Multinomial model of mate choice

We developed a multinomial model of mate choice based on the concepts of network theory (Newman 2010). Under this conceptual framework, a mating system can be seen as a sexual network in which the nodes are individuals and the links are copulations (*e.g.*, McDonald et al. 2013; Muniz et al. 2015). In this model, the sampling units are the observed copulations, which are also the links in the sexual network. We assume that in each copulation one individual acts as the choosing agent, while the other individual was available in the mating pool to be chosen. For simplicity, hereafter we will refer to the choosing individual as the female and to the individuals available to be chosen as males, but these roles can be easily switched according to the study system.

The model considers that a female can mate with any male available in the mating pool and generates a probability for each male based on the set of  $n$  predictor variables. The probability  $P_{ij}$  of a female  $i$  mating with male  $j$  is calculated as:

$$P_{ij} = \frac{\prod_{h=1}^n a_{jh}^{A_h}}{\sum_{k=1}^m \left( \prod_{h=1}^n a_{kh}^{A_h} \right)} \text{ (Eq. 1),}$$

in which the probability of a male  $j$  being chosen is calculated as a multiplication of terms divided by the sum of the same multiplication for all  $m$  males available in the mating pool. In equation 1,  $a_{jh}$  represents one of  $n$  predictor variables that can be a descriptor of the male  $j$ , such as body size or age, or a pairwise variable, such as spatial distance, phenotypic similarity or genetic relatedness between female  $i$  and male  $j$ . The exponents  $A_h$  are the parameters of the model that estimate the importance of each variable in determining copulation probabilities. The formula ensures that the sum of the probabilities calculated for each copulation adds to one, following the multinomial distribution. Thus, the probability of a male being chosen is dependent upon its own characteristics and also on the characteristics of the other males in the mating pool.

To fit the model, one must calculate the probabilities  $P_{ij}$  for all observed copulations, so that the identity of the male chosen in each copulation is the multinomial response variable of the model. Then, considering that the observed copulations are independent, and following the likelihood principle, we can consider these probabilities as likelihoods (Royall 1997). Once the probabilities  $P_{ij}$  are calculated, the negative log likelihood  $L$  of the model is calculated as:

$$L = - \sum_{i=1}^N \log(P_{ij}) \text{ (Eq. 2),}$$

in which  $N$  is the number of observed copulations. Then, the model can be fitted using maximum likelihood procedures. In the Supplemental Material, we provide R codes and worked examples of how to implement the multinomial model using R package *bbmle* (Bolker & R Development Core Team 2014).

The formula presented above does not accept predictor variables with values equal to or below zero because any such value will generate a negative probability (which does not make sense) or a zero probability (and the logarithm of zero is undefined). A solution to this problem is to exponentiate both the numerator and denominator of the formula, so that it becomes:

$$P_{ij} = \frac{\sum_{h=1}^n \exp(A_h \cdot a_{jh})}{\sum_{k=1}^m [\sum_{h=1}^n \exp(A_h \cdot a_{kh})]} \text{ (Eq. 3).}$$

The exponential transformation is adequate in this case because it preserves the ordering of the data. In this form, interpretation of the model coefficients becomes less intuitive. However, a model that does not accept negative predictor values can be problematic, especially if one wants to use centered predictor variables (Schielzeth 2010). In addition, only some of the variables can be exponentiated (see below).

#### *Categorical data*

It is possible to include categorical predictor variables in the model, but they must be binary variables. If there is a categorical variable with three states in the dataset, this variable must be included in the model as two binary variables (*i.e.*, the so-called dummy variables), as in most statistical models. We advise that the element representing each binary variable be written as  $\exp(D \cdot d_j)$ , where the variable  $d_j$  is a binary state and  $D$  is the parameter measuring its importance.

#### *Multiple measures and random factors*

In the multinomial model, we assume that all copulations are independent from one another. However, if there are multiple copulations by the same female, or if data is in some way divided in groups, a hierarchical structure must be included in the model to account for

this interdependence (Gelman & Hill 2006). Hierarchical models incorporate multiple measures or the presence of groups within the data via random factors, which are usually varying intercepts for each individual (or group) and sometimes varying slopes (Gelman & Hill 2006). The multinomial model does not have an intercept, and thus random factors will necessarily be present in the form of varying coefficients for some of the predictor variables. Unless there is reason to do otherwise, we suggest that the random factor is implemented as varying distance coefficients between individuals. As our results indicate (see below), ignoring the spatial distribution of individuals underestimates mate choosiness, thus a null model of mate choice should always include the spatial distance between individuals. If one has multiple hypotheses about mate choice and wants to use a model selection approach, all models should include the effect of space. This way, all competing models will have varying spatial coefficients and the coefficients of the variables that are not present in all models can be fixed. Because there is no analytical method to estimate varying coefficients in the multinomial model, the individual values can only be obtained via optimization. In the Supplemental Material we provide an example of linear optimization (Burnham & Andersson 2002) of the distance parameter using the dataset of blue tit extra pair copulations of Schlicht et al. (2015). For more complex hierarchical structures, linear optimization becomes impractical, and Bayesian MCMC will be the most adequate tool to obtain model parameters (Gelman et al. 2014).

### **Binomial model of mate choice**

Schlicht et al. (2015) proposed a model of mate choice in the context of extra pair copulations, but the model can be easily applied to other mating systems. This model (hereafter called *binomial model*) is a generalized linear mixed model (GLMM) in which each possible pair of male and female in the population is seen as a sampling unit. The binary response variable is the occurrence of copulation between a pair of individuals. Since each possible pair of individuals is a different sampling unit, pairwise variables can be easily included in the model along with non-pairwise ones. The identity of individuals is included as a random factor, so that the interdependence between multiple observations of the same individuals is accounted for in the model. One of the advantages of this model is that it is a GLMM, so that the analysis can be performed using any software that runs GLMMs.

## Individual based simulation

In order to test the efficiency of the multinomial and binomial models in the estimation of mate choosiness, we generated datasets using an individual based simulation (Grimm & Railsback 2005). The simulation is spatially explicit and mimics a simplified mating system with conventional sex roles. Each simulation represents one independent mating season of an entire population with an even sex ratio comprising  $N$  males and  $N$  females that occupy a two dimensional, squared landscape. We ran simulations with  $N$  values of 30, 60 and 100. The size of the landscape varies according to population size, maintaining a density  $d$  of 200 individuals per arbitrary square unit. Each male can mate with an unlimited number of females, and each female copulates with a single male, which she must choose among the males available within a radius  $r = 0.1$  from her own spatial coordinate. The  $r$  value was calculated so that, on average, a female will sample only six males, imposing strong information filtering. However, if there is no male within the radius  $r$  of a female, she simply mates with the closest male.

We ran simulations with two different mate choice criteria: 1) directional and 2) assortative. In the directional mate choice simulations, males have a phenotypic trait  $z$  and females prefer males with large  $z$  values. The probability of a male  $j$  being chosen by any female is proportional to  $z_j^B$ , in which  $z_j$  is the male trait and  $B$  measures the intensity of female choosiness. In the assortative preference simulation, both males and females have a phenotypic trait  $z$ , and females prefer males similar to themselves. In this simulation, the probability of a male  $j$  being chosen by a female  $i$  is proportional to  $\exp\left(B \cdot \frac{|z_i - z_j|}{z_i}\right)$ , in which  $z_j$  is the male trait,  $z_i$  is the female trait, and  $B$  measures the intensity of female choosiness. In the directional simulation, we set  $B$  to 2, while in the assortative simulation  $B$  was -1, so that in both cases female preference was moderate and non-linear. In both simulations,  $z$  follows a truncated normal distribution with  $z_{mean} = 4$ , standard deviation  $z_{sd} = 2$ , and minimum value  $z_{min} = 1$ .

One factor that can confound statistical tests is spatial autocorrelation of the data (Dale & Fortin 2002). To evaluate the impact spatial autocorrelation of phenotypic traits in the performance of the multinomial and binomial models, we ran simulations in which male and female phenotypic traits were spatially autocorrelated. We generated spatial autocorrelation by determining the mean from which  $z$  values were sampled based on the spatial location of individuals using the formula:

$$z_{mean} = 1 + \left( \frac{s_{io}}{s_o} \cdot 4 \right) \text{ (Eq. 4)},$$

in which  $s_{io}$  is the Euclidean distance of individual  $i$  from the origin of the squared landscape (coordinates  $x = 0, y = 0$ ) and  $s_o$  is the maximum distance to the origin. Therefore, individuals farther from the origin have higher phenotypic values. We considered a value of  $z_{min} = z_{mean}/4$  and a  $z_{sd} = 1$ . We implemented all simulations in the language R (R Development Core Team 2015).

## Analysis of the simulation data

### *Directional mate choice*

We used relatively simple versions of the multinomial and binomial models, in which we included only two predictor variables: the spatial distance between male and female and a male trait  $z$ , which are all variables that influence mating in the simulations. The probabilities  $P_{ij}$  in the multinomial model were calculated as:

$$P_{ij} = \frac{\exp(-A \cdot s_{ij}) \cdot z_j^B}{\sum_{k=1}^m [\exp(-A \cdot s_{ik}) \cdot z_k^B]} \text{ (Eq. 5)},$$

in which  $s_{ij}$  is the spatial distance between female  $i$  and male  $j$ ,  $z_j$  is the trait of male  $j$ , and  $m$  is the total number of males in the population. The parameter  $A$  measures the importance of spatial limitation on mate choice, while the parameter  $B$  measures female choosiness for values of  $z$ . We also analyzed the data using a generalized linear model (GLM) in which the spatial structure of the population was ignored. In the GLM, we used the male trait  $z$  as predictor of the number of copulations and assumed a Poisson distribution for the number of copulations. In all these models, we considered the coefficient  $B$  in the equation of  $P_{ij}$ , and the coefficients associated with the male trait in the binomial and Poisson models as estimates of female choosiness.

### *Assortative mate choice*

We analyzed the data from the assortative mating simulation using only the multinomial and the binomial models because there is no standardized method to investigate assortative mate choice. In both cases, we included as predictor variables the

spatial distance between individuals and the relative difference  $z_{ij}$  in the phenotypic trait  $z$  between pairs. We calculated the relative difference as:

$$z_{ij} = \frac{|z_i - z_j|}{z_i} \text{ (Eq. 6),}$$

in which  $z_i$  is the female trait and  $z_j$  is the male trait. In the multinomial model, we calculated the probability  $P_{ij}$  of a female  $i$  choosing to mate with a male  $j$  as:

$$P_{ij} = \frac{\exp(-A \cdot s_{ij} + B \cdot z_{ij})}{\sum_{k=1}^n [\exp(-A \cdot s_{ik} + B \cdot z_{ik})]} \text{ (Eq. 7),}$$

in which  $s_{ij}$  is the spatial distance between female  $i$  and male  $j$ . We used the coefficient  $B$  in the equation of  $P_{ij}$ , and the coefficient associated with the relative trait difference in the binomial model as estimates of female assortative choosiness.

#### *Detection of female choice and power*

We adopted a significance level of 5%, thus we considered that a fitted model successfully detected female choice when the 95% confidence interval of the coefficient of female choosiness did not include zero. Additionally, we measured the power of each model as the proportion of significant model fits.

#### *Coefficient calibration and confidence interval coverage*

The expected values for the coefficient of female choosiness in the multinomial model are simply the values of the parameter  $B$  used in the simulations. To estimate the expected values for the Poisson GLM and the binomial model we ran 100 calibration simulations for each mate choice criterion. These simulations followed the same parameter values presented above, with no spatial autocorrelation of traits, no information filtering (*i.e.*, females had access to all males in the population), and large population size ( $N = 500$ ), *i.e.*, ideal conditions for data analysis. We analyzed these simulations as described above and used the mean coefficient values obtained as the expected coefficient values for the Poisson GLM and binomial model. We used these expected values to estimate the confidence interval coverage (CIC) of the models, which we calculated as the proportion of fitted models in which the 95% confidence interval of the coefficient of female choosiness included the expected value.

### *Coefficient standardization and prediction error*

To make coefficients comparable between models, we standardized the coefficients of all models dividing them by their expected values. In addition, we estimated predictive accuracy of the models using the metric of mean squared prediction error (MSPE, Wallach & Goffinet 1989). To produce comparable prediction errors, we calculated MSPEs based on the standardized coefficients. Therefore, the expected value was always one and we calculated MSPE as:

$$MSPE = \frac{1}{N} \sum_{i=1}^N [(B_i - 1)^2] \text{ (Eq. 8),}$$

in which  $N$  is the total number of fitted models analyzed and  $B_i$  is the standardized coefficient obtained from each fitted model.

## **Results**

### *Directional mate choice*

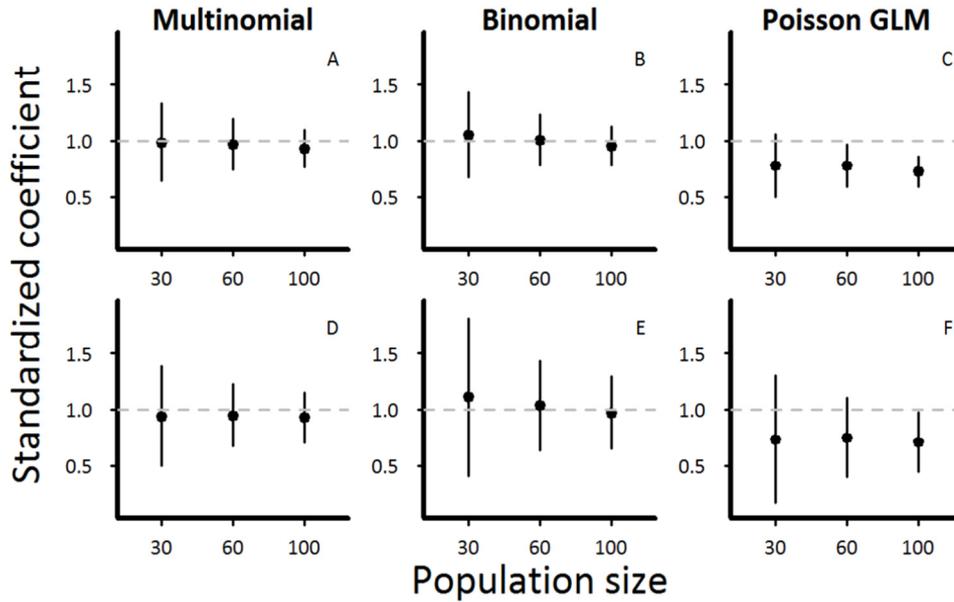
The mean ( $\pm$  SD) coefficient of the male trait in the calibration simulations was  $0.39 \pm 0.02$  for the binomial model for the Poisson GLM. Both the multinomial and binomial model always detected a significant spatial effect in the mating process. The Poisson GLM consistently underestimated female choosiness (Fig. 1), while the multinomial and binomial models accurately estimated female choosiness (Fig. 1). In all simulations, within each population size analysed, the multinomial model consistently presented lower prediction error, higher power, and higher confidence interval coverage than the binomial or the Poisson GLM (Table 1). Spatial autocorrelation decreased power and increased prediction error of all models, but the multinomial model was less affected (Table 1). Across all population sizes, the prediction error of the multinomial model was about half that of the binomial model, when spatial autocorrelation was present (Table 1).

**Table 1.** Summary of the analysis of datasets generated by simulations of directional mate choice. We analysed the data using the spatially explicit multinomial and binomial models, and a Poisson generalized linear model that ignored the spatial structure of the population.  $N$ : size of the analysed population. MSPE: the mean squared prediction error, which was calculated using standardized coefficients so that values are comparable between models. Power represents the proportion of significant results in 150 models. CIC: confidence interval coverage is the proportion of models in which the 95% confidence interval included the expected value.

<b>Directional mate choice</b>									
	<b>MSPE</b>			<b>Power</b>			<b>CIC</b>		
$N$	Multinomial	Binomial	Poisson	Multinomial	Binomial	Poisson	Multinomial	Binomial	Poisson
30	0.121	0.148	0.128	0.913	0.887	0.82	0.933	0.907	0.813
60	0.053	0.052	0.085	1	0.993	0.973	0.92	0.907	0.767
100	0.033	0.034	0.093	1	1	1	0.933	0.907	0.493

<b>Directional mate choice with spatial autocorrelation of phenotypes</b>									
	<b>MSPE</b>			<b>Power</b>			<b>CIC</b>		
$N$	Multinomial	Binomial	Poisson	Multinomial	Binomial	Poisson	Multinomial	Binomial	Poisson
30	0.203	0.502	0.391	0.76	0.58	0.46	0.907	0.86	0.773
60	0.079	0.16	0.184	1	0.9	0.753	0.933	0.927	0.76
100	0.055	0.104	0.156	1	0.94	0.853	0.92	0.86	0.693



**Figure 1.** Standardized coefficients of the analysis of datasets generated by simulations of directional mate choice. We analysed the data using the spatially explicit multinomial model (A and D), the spatially explicit binomial model (B and E) and a Poisson generalized linear model that ignores the spatial structure of the population (C and F). The standardized coefficients are presented as mean ( $\pm$  SD) of 150 simulations, and represent the intensity of female choosiness for males with large phenotypic traits. The grey horizontal line highlights the expected value. A, B and C: simulations without spatial autocorrelation of male traits. D, E and F: simulations with spatial autocorrelation of male traits.

### *Assortative mate choice*

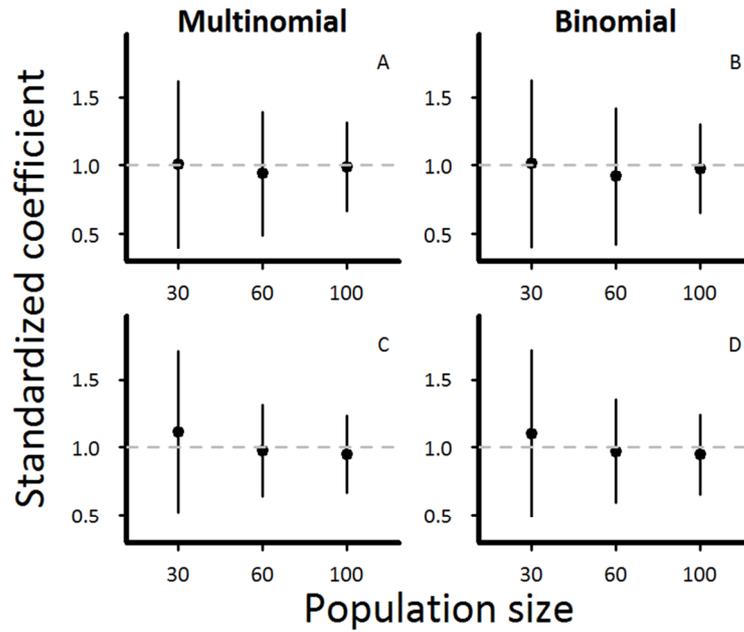
The mean ( $\pm$  SD) coefficient of the relative trait difference in the calibration simulations was  $-0.86 \pm 0.09$  for the binomial model. Both the multinomial and binomial model always detected the spatial effect in the mating process (Fig. 2). In the simulations without spatial autocorrelation, the multinomial model had higher power and lower prediction error than the binomial, but there was no consistent difference in the confidence interval coverage (Table 2). In the scenario with spatial autocorrelation there was no consistent difference in confidence interval coverage between the models. While the multinomial model presented slightly lower prediction error, the binomial model was more powerful in most population sizes (Table 2).

**Table 2.** Summary of the analysis of datasets generated by simulations of assortative mate choice. We analysed the data using the spatially explicit multinomial and binomial models  $N$ : size of the analysed population. MSPE: the mean squared prediction error, which was calculated using standardized coefficients so that values are comparable between models. Power represents the proportion of significant results in 150 models. CIC: confidence interval coverage, which is the proportion of models in which the 95% confidence interval included the expected value.

<b>Assortative mate choice</b>						
	<b>MSPE</b>		<b>Power</b>		<b>CIC</b>	
$N$	Multinomial	Binomial	Multinomial	Binomial	Multinomial	Binomial
30	0.369	0.37	0.507	0.453	0.94	0.96
60	0.209	0.251	0.7	0.66	0.9	0.873
100	0.102	0.104	0.953	0.927	0.947	0.953

<b>Assortative mate choice with spatial autocorrelation of phenotypes</b>						
	<b>MSPE</b>		<b>Power</b>		<b>CIC</b>	
$N$	Multinomial	Binomial	Multinomial	Binomial	Multinomial	Binomial
30	0.366	0.387	0.687	0.62	0.92	0.933
60	0.114	0.144	0.7	0.853	0.973	0.96
100	0.082	0.088	0.953	0.987	0.947	0.94



**Figure 2.** Standardized coefficients of the analysis of datasets generated by simulations of assortative mate choice. We analysed the data using the spatially explicit multinomial model (A and C) and the spatially explicit binomial model (B and D). The standardized coefficients are presented as mean ( $\pm$  SD) of 150 simulations and represent the intensity of female choosiness for males phenotypically similar to themselves. The grey horizontal line highlights the expected value. A and B: simulations without spatial autocorrelation of phenotypic traits. C and D: simulations with spatial autocorrelation of phenotypic traits.

## Discussion

The inclusion of distance as a predictor variable in the binomial and multinomial models greatly increased statistical power and produced unbiased estimates of female choosiness (Fig. 1). As expected, the GLM that ignored spatial structure consistently underestimated female choosiness, generating the pattern that confidence interval coverage decreased with increasing population size (Table 1). It has been known for a long time that information filtering could decrease the detectability of female choice (Benton & Evans 1998), but no general solution has been proposed. To the best of our knowledge, the binomial and multinomial models are the first proposed solutions to this problem. In the analysis of directional mate choice, the multinomial model performed generally better than the binomial model (Table 1). Perhaps this was caused by the fact that the multinomial model explicitly considers the characteristics of all males in the mating pool in the computation of probabilities. Spatial autocorrelation increased prediction error, especially in the binomial model, but did not promote consistent bias in the estimate of female choosiness (Fig 1., Table

1). Therefore, the spatially explicit multinomial and binomial models proved to be robust and flexible tools that allowed accurate estimates of female choosiness even with strong information filtering.

Both spatially explicit models were effective in detecting assortative mate choice, and both were little influenced by spatial autocorrelation of phenotypic traits (Fig. 2, Table 2). Even though assortative mating can be generated by multiple processes, assortative mate choice is an especially interesting mechanism for many reasons. For instance, assortative mate choice is an important component of many speciation models (*e.g.*, Kondrashov & Shpak 1988, Aguiar et al. 2009). To the best of our knowledge, there is no generally accepted method to estimate assortative mate choosiness based on observational data. In this sense, the binomial and multinomial models can be very useful in the study of assortative mate choice because they can separately estimate the effect of spatial proximity from the effect of actual assortative mate choice in observational studies. In addition, the coefficients of choosiness produced by these models can be used as comparative estimates of the intensity of assortative mate choice. Finally, the flexibility of the binomial and multinomial models allows one to include in the same model assortative mate choice for some traits and directional mate choice for other traits, so that mate choice based on multiple traits can be studied even when the criteria of mate choice vary between traits.

Our results indicate that the spatially explicit models of mate choice are more powerful and accurate than the alternative model that ignored space, but they also require more detailed data about the study populations. This detailed data, however, is already available for analysis for many populations. For instance, in the study of nesting birds, nest locations are usually registered, so that spatially explicit models can be used to investigate extra pair copulations in more detail using already available data (Brouwer et al. 2011, Stewart et al. 2010). The same is valid for other animals that live in burrows such as the wild cricket *Gryllus campestris*, for which detailed spatial data has been recorded (Bretman et al. 2011). When individuals do not build nests or burrows, home range centroids can be adopted as individual coordinates, or more detailed estimates of encounter probability can be used as predictor variables. In the case of animals that do not have stable home ranges, the use of spatially explicit models requires more detailed information, since each individual will have a different spatial coordinate in each sampling event. Data gathering in these scenarios may seem too laborious, but the result is a more detailed portrayal of the mating system under study, since the spatially explicit models not only provide better estimates of

female choosiness, but also a quantitative estimate of the spatial limitation involved in the mating process.

In conclusion, we presented here the multinomial and binomial models as spatially explicit models of mate choice, but they are actually spatially explicit models of choice or association in a broader sense. The binomial model does not presuppose a choosing agent, and can be seen as an association model, *i.e.*, a model that could be used to investigate associations others than copulations, such as predator-prey relationships or social associations like friendship. In this sense, the binomial model is conceptually similar to the  $p^*$  models, which are proposed as general network models (Anderson et al. 1999), but within a more familiar GLMM framework. The multinomial model presupposes a choosing agent and could be used to study choice in other contexts, such as oviposition site choice by females or plant choice by individual pollinators. In all of these examples, both models can be used either with or without explicit spatial information. If spatial coordinates for the individuals involved in the process of interest are not available, the models can still be used as flexible tools that can accommodate both pairwise and non-pairwise variables.

### **Acknowledgements**

DGM, GM and ESAS, were funded by the São Paulo Research Foundation (FAPESP grants 2013/13632-5, 2012/50229-1 and 2012/20468-4.. DGM is thankful to Nat Lim for logistic support and to Benedikt Holtmann, Carlos Esteban Lara, and Yu-Hsun Hsu for suggestions and insights in the early stages of this work. Early drafts of the manuscript received comments from Diogo Melo, to whom we are very grateful.

### **References**

- Aguiar, M.A.M., Baranger, M., Baptestini, E.M., Kaufman, L. & Bar-Yam, Y. (2009). Global patterns of speciation and diversity. *Nature*, **460**, 384–387.
- Anderson, C.J., Wasserman, S. & Crouch, B. (1999). A  $p^*$  primer: logit models for social networks. *Social Networks*, **21**, 37–66.
- Andersson, M. (1982). Female choice selects for extreme tail length in a widowbird. *Nature*, **299**, 818–820.
- Andersson, M. (1994). *Sexual Selection*. Princeton University Press, Princeton.

- Benton, T.G. & Evans, M.R. (1998). Measuring mate choice using correlation: the effect of female sampling behaviour. *Behavioral Ecology and Sociobiology*, **44**, 91–98.
- Bolker, B. & R Development Core Team (2014). *bbmle: Tools for general maximum likelihood estimation*. R package version 1.0.17. Available from: <http://CRAN.R-project.org/package=bbmle>.
- Bretman, A., Rodríguez-Muñoz, R., Walling, C., Slate, J. & Tregenza, T. (2011). Fine-scale population structure, inbreeding risk and avoidance in a wild insect population. *Molecular Ecology*, **20**, 3045–3055.
- Brouwer, L., Van De Pol, M., Atema, E. & Cockburn, A. (2011). Strategic promiscuity helps avoid inbreeding at multiple levels in a cooperative breeder where both sexes are philopatric. *Molecular Ecology*, **20**, 4796–807.
- Burley, N. (1983). The meaning of assortative. *Ethology & Sociobiology*, **203**, 191–203.
- Burnham, K.P. & Anderson, D.R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd edition. Springer, New York, USA.
- Byers, J., Wiseman, P., Jones, L. & Roffe, T. (2005). A large cost of female mate sampling in pronghorn. *The American Naturalist*, **166**, 661–668.
- Côte, I.M. & Hunte, W. (1989). Male and female mate choice in the redlip blenny : why bigger is better. *Animal Behaviour*, **38**, 78–88.
- Dale, M. & Fortin, M. (2002). Spatial autocorrelation and statistical tests in ecology. *Ecoscience*, **9**, 162–167.
- Freeman-Gallant, C., Megherdichian, M., Wheelright, N. & Sollecito, S. (2003). Social pairing and female mating fidelity predicted by restriction fragment length polymorphism similarity at the major histocompatibility complex in a songbird. *Molecular Ecology*, **12**, 3077–3083.
- Gelman, A., Carlin, J.B., Stern, H.S. & Rubin, D.B. (2014). *Bayesian Data Analysis*, 3rd edition. Chapman & Hall, London.
- Gelman, A. & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, Cambridge.
- Gerhardt, H.C. (1991). Female mate choice in treefrogs: static and dynamic acoustic criteria. *Animal Behaviour*, **42**, 615–635.

- Grimm, V. & Railsback, S.F. (2005). *Individual Based Modeling and Ecology*. Princeton University Press, Princeton.
- Houde, A.E. (1987). Mate choice based upon naturally occurring color-pattern variation in a guppy population. *Evolution*, **41**, 1–10.
- Janetos, A.C. (1980). Strategies of female mate choice: a theoretical analysis. *Behavioral Ecology and Sociobiology*, **12**, 107–112.
- Jones, A.G. & Ratterman, N.L. (2009). Mate choice and sexual selection: What have we learned since Darwin? *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 10001–10008.
- Kondrashov, A. & Shpak, M. (1998). On the origin of species by means of assortative mating. *Proceedings of the Royal Society B: Biological Sciences*, **265**, 2273–2298.
- Lane, J.E., Boutin, S., Speakman, J.R. & Humphries, M.M. (2010). Energetic costs of male reproduction in a scramble competition mating system. *Journal of Animal Ecology*, **79**, 27–34.
- McDonald, G.C.; James, R.; Krause, J.; Pizzari, T. (2013). Sexual networks: measuring sexual selection in structured, polyandrous populations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **368**, 20120356.
- Mossa, S., Barthélémy, M., Eugene Stanley, H.E. & Amaral, L.A.N. (2002). Truncation of power law behavior in “scale-free” network models due to information filtering. *Physical Review Letters*, **88**, 138701.
- Muniz, D.G., Guimarães Jr., P.R., Buzatto, B.A. & Machado, G. (2015). A sexual network approach to sperm competition in a species with alternative mating tactics. *Behavioral Ecology*, **26**, 121–129.
- Newman, M. (2010). *Networks: An Introduction*. Oxford University Press, Oxford.
- Polis, G., Barnes, J. & Seely, M. (1998). Predation as a major cost of reproduction in Namib Desert tenebrionid beetles. *Ecology*, **79**, 2560–2566.
- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available from: <http://www.R-project.org/>.
- Real, L. (1990). Search theory and mate choice. I. Models of single-sex discrimination. *The American Naturalist*, **136**, 376–405.

- Roff, D.A. & Fairbairn, D.J. (2014). The evolution of phenotypes and genetic parameters under preferential mating. *Ecology and Evolution*, **4**, 2759–2776.
- Royall, R. (1997). *Statistical Evidence: A Likelihood Paradigm*. Chapman & Hall, London.
- Schlicht, L., Valcu, M. & Kempenaers, B. (2015). Spatial patterns of extra-pair paternity: beyond paternity gains and losses. *The Journal of Animal Ecology*, **84**, 518–531.
- Stewart, S.L.M., Westneat, D.F. & Ritchison, G. (2009). Extra-pair paternity in eastern bluebirds: effects of manipulated density and natural patterns of breeding synchrony. *Behavioral Ecology and Sociobiology*, **64**, 463–473.
- Thomas, M.L. & Simmons, L.W. (2009). Sexual selection on cuticular hydrocarbons in the Australian field cricket, *Teleogryllus oceanicus*. *BMC Evolutionary Biology*, **9**, 162.
- Wallach, D. & Goffinet, B. (1989). Mean squared error of prediction as a criterion for evaluating and comparing system models. *Ecological Modelling*, **44**, 299–306.
- Zuk, M., Ligon, J.D. & Thornhill, R. (1992). Effects of experimental manipulation of male secondary sex characters on female mate preference in red jungle fowl. *Animal Behaviour*, **44**, 999–1006.

# Supplemental Material

## Tutorial for the multinomial model of mate choice in R

This is the Supplemental Material of the manuscript *Spatially explicit models of mate choice and assortative mating*. In this tutorial you will learn how to calculate copulation probabilities using the multinomial model and how to optimize the model parameters using the R package *bbmle*. At first, we will see how to fit the multinomial model using simulated data. Then, we will use the blue tit data available in the package *expp* to see how to use linear optimization to add a random factor to the model. This tutorial was elaborated using R package *rmarkdown*.

### Part 1. Simulated data with directional mate choice

In this first part we will use a simulated dataset. In this dataset females have directional preference for males with high trait values. But let us pretend we do not know that and use the multinomial model to estimate female preference for large male traits. The first step is to load the data. It is saved in three files:

1. `tmales.txt` - a file with the information on the males.

```
males = read.table("tmales.txt", sep="\t", dec=".", header=TRUE)
head(males)
##      wid id      x      y ms  trait
## 1 2647  1 0.61330978 0.10805031  2 4.321517
## 2 2647  2 0.03135099 0.25517120  1 2.893953
## 3 2647  3 0.54606595 0.49578497  4 5.420178
## 4 2647  4 0.11690116 0.17813385  1 6.542484
## 5 2647  5 0.12727202 0.09846337  2 4.691839
## 6 2647  6 0.03660171 0.12933401  2 4.843570
```

In this file, we have `x` and `y` spatial positions for each male, and its trait value. There is also a column `ms`, which records the mating success of each male.

## 2. tfemales.txt - a file with the information on the females.

```
females = read.table("tfemales.txt", sep="\t", dec=".", header=TRUE)
head(females)
##      wid id      x      y neighbors
## 1 2647  1 0.6916669 0.41298186      4
## 2 2647  2 0.3388266 0.12437187      3
## 3 2647  3 0.4387549 0.19121923      3
## 4 2647  4 0.5953950 0.22947917      2
## 5 2647  5 0.1155520 0.15945149      8
## 6 2647  6 0.3456921 0.03174804      2
```

Females also have  $x$  and  $y$  positions, but they do not have trait values. The column `neighbors` records how many males each female could sample during mate choice.

## 3. tcouples.txt - a file showing which males mated with which females.

```
couples = read.table("tcouples.txt", sep="\t", dec=".", header=TRUE)
head(couples)
##      wid female male      z
## 1 2647      1     7 4.417459
## 2 2647      2    59 4.575273
## 3 2647      3    29 6.281996
## 4 2647      4    48 4.821807
## 5 2647      5    66 9.865502
## 6 2647      6    21 4.214258
```

This is the edgelist of the sexual network, i.e., a list of who mated with whom. In addition to the male and female IDs, the column `z` records the trait value of the male.

The model we are going to fit will include as predictor variables the spatial distance and male trait. So that the probability of female  $i$  mating with male  $j$  can be written as:

$$P_{ij} = \frac{\exp(-A \cdot s_{ij}) \cdot z_j^B}{\sum_{k=1}^m [\exp(-A \cdot s_{ik}) \cdot z_k^B]}$$

In the equation above,  $s_{ij}$  is the spatial distance between female  $i$  and male  $j$ ,  $z_j$  is the value of male's  $j$  trait,  $s_{ik}$  are the distances between the female  $i$  and all  $m$  males in the mating pool, and  $z_k$  are the values of the traits of all males in the mating pool.  $A$  and  $B$  are the model parameters.

The easiest way to calculate the probabilities for all copulations is by using vectorized operations and matrix manipulations. However, it may be easier to understand the whole

process if we start calculating the probability of a single copulation. So let us use the first copulation in the couples data.frame and let us assume parameter values of  $A = -10$  and  $B = 2$ . So first, let us take a look at the data:

```
head(couples) #so we have female 1 and male 7
##      wid female male      z
## 1 2647      1     7 4.417459
## 2 2647      2    59 4.575273
## 3 2647      3    29 6.281996
## 4 2647      4    48 4.821807
## 5 2647      5    66 9.865502
## 6 2647      6    21 4.214258
```

As you can see, we are dealing with female 1 and male 7.

```
females[1,]
##      wid id      x      y neighbors
## 1 2647  1 0.6916669 0.4129819      4
males[7,]
##      wid id      x      y ms  trait
## 7 2647  7 0.6208285 0.3538615  4 4.417459
```

In order to continue this example, we will calculate the spatial distances between female 1 and all males in the mating pool. So it is time to introduce a function that quickly calculates Euclidean distances between matrices of cartesian (x and y) coordinates:

```
euclid = function(m1, m2=m1)
{
  return(sqrt((outer(m1[,1],m2[,1], "-")^2)+(outer(m1[,2],m2[,2], "-")^2)))
}
```

With this function we can calculate all the  $s_{ik}$  distances:

```
sik = euclid(females[1, c("x","y")], males[, c("x","y")])
round(sik,3)
##      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11]
## [1,] 0.315 0.679 0.167 0.621 0.646 0.714 0.092 0.369 0.31 0.527 0.441
##      [,12] [,13] [,14] [,15] [,16] [,17] [,18] [,19] [,20] [,21] [,22]
## [1,] 0.569 0.316 0.605 0.494 0.598 0.327 0.591 0.532 0.646 0.492 0.256
##      [,23] [,24] [,25] [,26] [,27] [,28] [,29] [,30] [,31] [,32] [,33]
## [1,] 0.77 0.715 0.289 0.354 0.362 0.287 0.321 0.295 0.63 0.387 0.19
```

```
##      [,34] [,35] [,36] [,37] [,38] [,39] [,40] [,41] [,42] [,43] [,44]
## [1,] 0.608 0.763 0.203 0.678 0.361 0.717 0.599 0.133 0.385 0.079 0.104
##      [,45] [,46] [,47] [,48] [,49] [,50] [,51] [,52] [,53] [,54] [,55]
## [1,] 0.074 0.412 0.428 0.113 0.309 0.189 0.418 0.234 0.583 0.577 0.392
##      [,56] [,57] [,58] [,59] [,60] [,61] [,62] [,63] [,64] [,65] [,66]
## [1,] 0.539 0.37 0.235 0.41 0.743 0.376 0.567 0.687 0.334 0.526 0.619
##      [,67] [,68] [,69] [,70] [,71] [,72] [,73] [,74] [,75] [,76] [,77]
## [1,] 0.232 0.328 0.358 0.655 0.479 0.65 0.606 0.584 0.449 0.564 0.752
##      [,78] [,79] [,80] [,81] [,82] [,83] [,84] [,85] [,86] [,87] [,88]
## [1,] 0.687 0.198 0.169 0.255 0.129 0.476 0.388 0.614 0.394 0.453 0.651
##      [,89] [,90] [,91] [,92] [,93] [,94] [,95] [,96] [,97] [,98] [,99]
## [1,] 0.436 0.692 0.445 0.493 0.585 0.679 0.086 0.575 0.494 0.416 0.683
##      [,100]
## [1,] 0.694
```

Now that we have the distances, the  $z_k$  values are easy to recover:

```
zk = males$trait
round(zk,3)
## [1] 4.322 2.894 5.420 6.542 4.692 4.844 4.417 3.143 3.077 2.131 3.450
## [12] 3.709 6.255 5.762 4.684 2.429 2.538 6.228 4.641 2.572 4.214 4.294
## [23] 3.773 3.213 3.676 4.440 1.810 7.740 6.282 4.601 5.021 3.536 2.754
## [34] 2.008 2.362 2.086 5.510 5.634 3.717 3.519 2.073 4.670 2.224 3.587
## [45] 4.448 7.646 4.058 4.822 6.650 3.738 6.313 3.509 4.531 6.159 4.730
## [56] 6.332 2.851 4.483 4.575 2.301 3.988 1.245 7.625 4.001 7.085 9.866
## [67] 3.010 7.288 6.960 2.142 6.960 2.371 3.665 1.876 5.263 4.851 3.702
## [78] 3.634 1.549 2.121 3.186 2.706 3.973 5.202 5.106 4.658 3.637 6.087
## [89] 5.995 2.384 3.320 1.219 3.639 4.116 2.686 3.722 1.894 5.309 5.302
## [100] 5.582
```

With these values we can calculate  $\exp(-A \cdot s_{ij}) \cdot z_j^B$ :

```
A = -10
B = 2
exp(A*sik[7]) * zk[7]^B
## [1] 7.755902
```

Since all mathematical operators in R are vectorized, we can calculate  $\sum_{k=1}^m [\exp(-A \cdot s_{ik}) \cdot z_k^B]$  with a very similar command line:

```
sum(exp(A*sik) * zk^B)
```

```
## [1] 81.4379
```

Now it is easy to calculate the value  $P_{ij}$ :

```
Pij = exp(A*sik[7])*zk[7]^B/sum(exp(A*sik)*zk^B)
## [1] 0.09523701
```

Another way to calculate  $P_{ij}$  would be to build a vector with all values of  $\exp(-A \cdot s_{ik}) \cdot z_k^B$  and then calculate  $P_{ij}$  using:

```
vectorik = exp(A*sik)*zk^B
(Pij = vectorik[7]/sum(vectorik))
## [1] 0.09523701
```

This approach leads us to the same value.

Now we need to do the same for all 100 copulations in our couples data.frame. If for one female we had to build a vector, now we will need a matrix, in which each row will work just like the vector we build above. So our first step is to build the spatial matrix using the function `euclid()`:

```
sikMatrix = euclid(females[, c("x","y")], males[, c("x","y")])
round(sikMatrix[1:10,1:10],2)
##      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
## [1,] 0.31 0.68 0.17 0.62 0.65 0.71 0.09 0.37 0.31 0.53
## [2,] 0.27 0.33 0.43 0.23 0.21 0.30 0.36 0.39 0.15 0.34
## [3,] 0.19 0.41 0.32 0.32 0.33 0.41 0.24 0.34 0.04 0.35
## [4,] 0.12 0.56 0.27 0.48 0.49 0.57 0.13 0.39 0.15 0.47
## [5,] 0.50 0.13 0.55 0.02 0.06 0.08 0.54 0.42 0.33 0.26
## [6,] 0.28 0.39 0.51 0.27 0.23 0.32 0.42 0.49 0.22 0.42
## [7,] 0.57 0.14 0.63 0.10 0.08 0.02 0.63 0.50 0.42 0.32
## [8,] 0.65 0.30 0.42 0.36 0.44 0.42 0.52 0.21 0.44 0.13
## [9,] 0.33 0.30 0.48 0.19 0.15 0.25 0.43 0.43 0.21 0.34
## [10,] 0.61 0.65 0.22 0.65 0.71 0.74 0.37 0.25 0.48 0.44
```

The traits of the males are still only a vector of values. However, we must arrange the values in form of a matrix. This matrix will contain, in each row, the trait values of all males.

```
zkMatrix = traitMatrix = matrix(males$trait, nrow=nrow(couples), ncol=nrow(males),byrow=TRUE)
round(zkMatrix[1:10,1:10],2)
##      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
## [1,] 4.32 2.89 5.42 6.54 4.69 4.84 4.42 3.14 3.08 2.13
```

```
## [2,] 4.32 2.89 5.42 6.54 4.69 4.84 4.42 3.14 3.08 2.13
## [3,] 4.32 2.89 5.42 6.54 4.69 4.84 4.42 3.14 3.08 2.13
## [4,] 4.32 2.89 5.42 6.54 4.69 4.84 4.42 3.14 3.08 2.13
## [5,] 4.32 2.89 5.42 6.54 4.69 4.84 4.42 3.14 3.08 2.13
## [6,] 4.32 2.89 5.42 6.54 4.69 4.84 4.42 3.14 3.08 2.13
## [7,] 4.32 2.89 5.42 6.54 4.69 4.84 4.42 3.14 3.08 2.13
## [8,] 4.32 2.89 5.42 6.54 4.69 4.84 4.42 3.14 3.08 2.13
## [9,] 4.32 2.89 5.42 6.54 4.69 4.84 4.42 3.14 3.08 2.13
## [10,] 4.32 2.89 5.42 6.54 4.69 4.84 4.42 3.14 3.08 2.13
```

Thus, calculating all  $\exp(-A \cdot s_{ik}) \cdot z_k^B$  values is straightforward:

```
pikMatrix = exp(A*sikMatrix)*(zkMatrix^B)
round(pikMatrix[1:10,1:10],2)
##      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
## [1,] 0.80 0.01 5.50 0.09 0.03 0.02 7.76 0.25 0.43 0.02
## [2,] 1.19 0.30 0.42 4.36 2.61 1.14 0.51 0.19 2.17 0.16
## [3,] 2.70 0.14 1.16 1.71 0.85 0.40 1.70 0.32 6.51 0.13
## [4,] 5.47 0.03 1.96 0.35 0.17 0.08 5.48 0.21 2.06 0.04
## [5,] 0.13 2.34 0.12 35.49 11.83 10.08 0.09 0.14 0.33 0.34
## [6,] 1.16 0.18 0.19 2.83 2.24 0.92 0.28 0.08 1.05 0.07
## [7,] 0.06 1.98 0.05 15.99 9.53 19.27 0.04 0.07 0.15 0.18
## [8,] 0.03 0.41 0.46 1.16 0.27 0.35 0.11 1.26 0.12 1.25
## [9,] 0.67 0.42 0.23 6.69 4.73 1.98 0.27 0.13 1.13 0.15
## [10,] 0.04 0.01 3.40 0.06 0.02 0.01 0.47 0.83 0.08 0.05
```

Now, using indexation, we can retrieve from this matrix the values that represent the copulations that actually happened. To do this, we can use the *couples* data.frame as our index:

```
head(couples)
##      wid female male      z
## 1 2647      1     7 4.417459
## 2 2647      2    59 4.575273
## 3 2647      3    29 6.281996
## 4 2647      4    48 4.821807
## 5 2647      5    66 9.865502
## 6 2647      6    21 4.214258

positions = cbind(couples[, "female"], couples[, "male"])
head(positions)
##      [,1] [,2]
## [1,]    1    7
```

```
## [2,] 2 59
## [3,] 3 29
## [4,] 4 48
## [5,] 5 66
## [6,] 6 21

(almostPij = pikMatrix[positions])
## [1] 7.755902 10.637957 17.624599 9.056371 36.138224 8.232941 19.267297
## [8] 13.681793 25.159068 5.058646 9.378082 12.983386 20.055675 20.694936
## [15] 14.008325 10.218066 15.331798 10.914285 14.359189 22.359122 7.036750
## [22] 10.717382 6.395268 11.917992 5.467871 6.818757 11.472977 4.207353
## [29] 12.234491 3.900088 18.970059 9.808395 7.950080 13.621967 23.135002
## [36] 15.488633 8.586017 51.229911 12.502006 8.084977 14.112337 10.052263
## [43] 5.351341 9.319521 22.996286 6.907052 8.843169 26.915263 6.655741
## [50] 15.550202 18.210203 4.235583 18.111156 9.596967 5.922906 10.009126
## [57] 16.132469 6.158778 5.225619 12.837323 11.961637 9.766017 25.472170
## [64] 9.447477 21.656797 15.897218 22.877323 4.128583 22.555742 32.929108
## [71] 3.808382 9.821055 2.568802 3.778696 7.288195 4.108277 7.741786
## [78] 21.143831 8.792542 6.274846 9.791517 32.045151 12.853835 7.197299
## [85] 7.497651 6.338364 2.219462 22.249319 17.233116 42.318570 9.619406
## [92] 15.350002 8.761898 4.394707 14.347624 17.524500 5.161339 4.955011
## [99] 13.496368 28.308966
```

As you can see, the object *almostPij* is full of values above one. To convert these values into probabilities, we must divide each value by the sum of the row from which it came. These sums can be calculated using the `apply()` function:

```
pikSums = apply(pikMatrix,1,sum)
Pij = almostPij/pikSums
## [1] 0.09523701 0.06163079 0.10938532 0.08353207 0.14941254 0.07068170
## [7] 0.10250974 0.06407080 0.12836701 0.06371333 0.04875194 0.12337759
## [13] 0.19319930 0.10968135 0.07062546 0.09763362 0.06334598 0.07166604
## [19] 0.09503188 0.09958911 0.07975322 0.06419204 0.03109384 0.05848906
## [25] 0.03486127 0.14272883 0.10671720 0.04672176 0.11398729 0.04821943
## [31] 0.07659503 0.07259572 0.05177375 0.11276552 0.10580918 0.06474638
## [37] 0.06886215 0.20912611 0.18324449 0.07854086 0.06589978 0.07432014
## [43] 0.02338726 0.03219044 0.12796294 0.05120548 0.07196984 0.11104760
## [49] 0.05870822 0.13254496 0.10398937 0.05110019 0.09241593 0.08979151
## [55] 0.08042074 0.11255139 0.08485428 0.02683898 0.04739500 0.05649323
## [61] 0.09185539 0.07748109 0.10529304 0.05880385 0.07121975 0.12748862
## [67] 0.12854648 0.02249577 0.07274885 0.27450163 0.04617613 0.04962291
## [73] 0.02362365 0.02332507 0.09858932 0.03081318 0.10150461 0.15361050
## [79] 0.08175732 0.04505729 0.10088179 0.12272936 0.10255482 0.03487167
```

```
## [85] 0.08038163 0.08720488 0.01172953 0.09647572 0.08450863 0.16979105
## [91] 0.09724441 0.05576114 0.07108582 0.01848819 0.12153637 0.09100097
## [97] 0.05218920 0.02046626 0.06380140 0.09722768
```

Alternatively, we can calculate the  $P_{ij}$  values with less commands:

```
pikMatrix = exp(A*sikMatrix)*(zkMatrix^B)
positions = cbind(couples[, "female"], couples[, "male"])
Pij = pikMatrix[positions]/apply(pikMatrix, 1, sum)
```

Thus, the negative log likelihood of the model (considering the parameter values of  $A = -10$  and  $B = 2$ ) is:

```
-sum(log(Pij))
## [1] 261.3027
```

Now we can calculate the negative log-likelihood (which we will call just log-likelihood from now on) of our model. Or at least, we can calculate the likelihood given certain parameter values. We can use the optimization function in the package *bbmle* to find the most likely values of our parameters. The optimization function we will use is the `mle2()`. The main parameters of this function are:

- `minuslogl` - which is a function that returns a negative likelihood value;
- `start` - which is a named list of starting parameter values for the optimizer;
- `data` - a named list with the objects containing the data.

What is the data needed by our function? Previously we stored all the data in the objects *sikMatrix*, *zkMatrix* and *positions*. With all this we can write our log-likelihood function, which we will call `LL.copula()`.

The function will have five parameters:  $A$  and  $B$ , which are the model parameters,  $s_{ik}$  and  $z_k$ , which are the matrices of spatial distances and of male traits, and *pos*, which is the indexing matrix.

```
LL.copula = function(A, B, sik, zk, pos)
{
  pik = exp(sik*A) * (zk^B)

  pij = pik[pos]/apply(pik, 1, sum)
```

```

return(-sum(log(pij)))
}

```

Finally, we will fit the model. To do so, we need to load the package *bbmle* and call the `mle2()` function. As we said, we need starting values, since we suppose *A* to be negative and *B* to be positive, we will start *A* with -1 and *B* with 1.

```

library(bbmle)
## Loading required package: stats4
fitted = mle2(minuslogl = LL.copula, start = list(A=-1, B=1),
              data = list(sik=sikMatrix, zk=zkMatrix, pos=positions))
fitted
##
## Call:
## mle2(minuslogl = LL.copula, start = list(A = -1, B = 1), data = list(sik = sikMatrix,
##      zk = zkMatrix, pos = positions))
##
## Coefficients:
##           A           B
## -34.256275   2.053015
##
## Log-likelihood: -182.26

```

Model fitted! As we expected, the value of *B* is positive, so females really prefer males with large ornaments. Now we need to calculate confidence intervals, or likelihood intervals to be able to actually say something about these values. The confidence intervals can be easily obtained using the `confint()` function:

```

(confint95 = confint(fitted))
##           2.5 %           97.5 %
## A -40.354007 -28.889100
## B  1.349098   2.805799

```

The likelihood intervals can be obtained using the `profile()` function. The two main parameters of this function are the fitted object and `zmax`. As you can see in the help file of the function `profile.mle2()`, the parameter `zmax` is the square root of the deviance difference of the likelihood interval. The deviance, as you know, is twice the log-likelihood. So if we want a log-likelihood interval with a maximum difference of two units (as is usual), we must set `zmax` to the square root of two times two... which is two!

```

(profile2 = profile(fitted, zmax=2))
## Likelihood profile:
##
## $A
##          z par.vals.A par.vals.B
## 1 -2.2315287 -41.260358  2.260653
## 2 -1.8810289 -40.093011  2.226024
## 3 -1.5225441 -38.925664  2.191415
## 4 -1.1556622 -37.758317  2.156830
## 5 -0.7799387 -36.590970  2.122245
## 6 -0.3948921 -35.423623  2.087646
## 7  0.0000000 -34.256275  2.053015
## 8  0.4053057 -33.088928  2.018337
## 9  0.8216461 -31.921581  1.983587
## 10 1.2497013 -30.754234  1.948741
## 11 1.6902192 -29.586887  1.913757
## 12 2.1440257 -28.419540  1.878608
##
## $B
##          z par.vals.A par.vals.B
## 1 -2.0697321 -33.118251  1.310916
## 2 -1.6444932 -33.310258  1.459336
## 3 -1.2249422 -33.519912  1.607756
## 4 -0.8110445 -33.747375  1.756176
## 5 -0.4027509 -33.992795  1.904595
## 6  0.0000000 -34.256275  2.053015
## 7  0.3972732 -34.537881  2.201435
## 8  0.7891474 -34.837573  2.349855
## 9  1.1757007 -35.155293  2.498275
## 10 1.5570158 -35.490870  2.646695
## 11 1.9331769 -35.844065  2.795114
## 12 2.3042696 -36.214565  2.943534

```

So, by both the “95% confidence criterion” and by the “log-likelihood < 2 criterion” we can say that the females in the population prefer males with high trait values and usually mate with nearby males.

## Part 2. Blue tit extra-pair copulation data

The next example is with the blue tit data. This is a dataset of extra-pair copulations (EPC) in which it is hypothesized that females prefer to mate with older males and with larger males. In this case, the model we are going to fit will include spatial distance between male and female nest, male age and male tarsus length (a proxy of male size). The dataset includes data from many years, but we will choose just the year with the largest sample size (2003). There are repeated EPCs from the same female. At first, we will ignore this, but later we will include a random factor of female identity.

The first step is to load the data, which is available in the R package *expp*. The data is specifically in the objects *bluetit\_breeding* and *bluetit\_epp*.

```
library(expp)
data(bluetit_breeding)
str(bluetit_breeding)
## 'data.frame': 1025 obs. of 10 variables:
## $ year_ : int 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 ...
## $ id : int 13 170 129 173 111 178 182 276 167 23 ...
## $ x : num 4417656 4417608 4417414 4417477 4417785 ...
## $ y : num 5334188 5334576 5334479 5334591 5334369 ...
## $ female : chr NA NA NA "f161" ...
## $ male : chr NA NA "m245" "m173" ...
## $ layingDate : int 100 105 106 103 104 101 100 106 104 103 ...
## $ male_age : chr "adult" "adult" "adult" "adult" ...
## $ male_tarsus: num NA NA NA NA 16.5 ...
## $ study_area : chr "Westerholz" "Westerholz" "Westerholz" "Westerholz" ...
data(bluetit_epp)
str(bluetit_epp)
## 'data.frame': 425 obs. of 3 variables:
## $ year_ : int 2001 2001 2003 2001 2001 2002 2002 2002 2001 2001 ...
## $ male : chr "m525" "m515" "m406" "m7" ...
## $ female: chr "f2" "f4" "f4" "f5" ...
```

To make things easier, we will split the data of males and females in different data.frames. In addition, we will save the EPC data in a object called simply *epp* (for extra-pair paternity) and the information about the social couples in the data.frame *social*. We will select only the columns that will be useful to us, and will use only the data from year 2003.

```
epp = bluetit_epp[bluetit_epp$year == 2003,]
males = bluetit_breeding[bluetit_breeding$year == 2003,
                        c("x", "y", "male", "male_age",
```

```

      "male_tarsus"]
males$binary_age = ifelse(males$male_age=="adult", yes=1, no=0)

females = bluetit_breeding[bluetit_breeding$year == 2003,
                          c("x", "y", "female", "male")]

social = unique(data.frame(female=bluetit_breeding$female[bluetit_breeding$year == 2003],
                          male=bluetit_breeding$male[bluetit_breeding$year == 2003]))

social = social[complete.cases(social),]

```

Some males and females are duplicated in the male and female data.frames because they had more than one nest in the same mating season. For simplicity, we will calculate mean coordinates for these individuals. We will use the function `aggregate()` to calculate these means and aggregate all the data in the same object. This is one of the reasons why we converted the male ages to a binary value (below we will need this binary variable for the likelihood calculations).

```

females = aggregate(x = females[,c("x", "y")], list(females$female), mean)
colnames(females)[1] = "id"
males = aggregate(x = males[,c("x", "y", "male_tarsus", "binary_age")],
                  list(males$male), mean)
colnames(males)[1] = "id"

```

Now, the last step before we can finally fit the model. We need to eliminate EPCs with males that have incomplete data.

```

#eliminating the males with missing data
full.males = males$id[complete.cases(males)]
epp = epp[epp$male %in% full.males,]
head(epp)
##   year_ male female
## 3   2003 m406    f4
## 12  2003 m48    f14
## 15  2003 m21    f23
## 16  2003 m44    f23
## 24  2003 m417   f32
## 27  2003 m85    f32

```

As before, we will arrange all of our data in matrices. Let us start by the spatial matrix. In this dataset, each individual is assigned to a nest, and each nest has x and y

coordinates describing its spatial position. The way we arranged the data, this information is replicated in the males and females data.frame.

As we mentioned before, some females have more than one EPC, while some females did not mate extra-pair. Thus, after we calculate the matrix of spatial distances, we need to adjust it, so that each row represents one extra-pair copulation and each cell represents the distance between the female that performed that EPC and all other males in the population. If we name the rows and columns of the distance matrix according to the female and male IDs, this adjustment is easily done with some indexing, as we show below:

```
#Spatial matrix
sik03 = euclid(females[,c("x","y")], males[,c("x","y")])
rownames(sik03) = females$id
colnames(sik03) = males$id
sik03[as.matrix(social)] = NA

sik03 = sik03[epp$female,]
rownames(sik03) = 1:nrow(epp)#adjusting rownames to avoid duplicates

sik03 = sik03/100 #just putting distance values in a scale easier to interpret
```

In addition to the adjustments we explained above, we also replaced by NA all values that refer to the social partners of the females (after all, we are working with EPCs).

With the distance matrix ready in the object *sik03*, we will create the trait matrices of males as we did in the previous example. In this example we have two male traits, one continuous, which is tarsus length, and one categorical, male age (juveniles or adults). As explained in the manuscript, categorical variables need to be transformed in binary values, so we consider that adults are ones and juveniles are zeroes.

```
#Male tarsus length matrix
tarsus03 = matrix(males$male_tarsus, nrow=nrow(epp), ncol=nrow(males), byrow=TRUE)
rownames(tarsus03) = 1:nrow(epp)
colnames(tarsus03) = males$id

#Male age matrix
age03 = matrix(males$binary_age, nrow=nrow(epp), ncol=nrow(males), byrow=TRUE)
rownames(age03) = 1:nrow(epp)
colnames(age03) = males$id
```

Before we continue, we need to assemble a matrix with the positions in the matrix that refer to the data of extra pair partners. The female positions are simply a sequence of numbers, and the male positions are the male IDs in the *epp* data.frame.

```
positions03 = cbind(1:nrow(epp), epp$male)
```

Now that the data is ready, we can write our log-likelihood function as we did before. The model we will fit now will include distance, tarsus length, and age. The hypotheses are that females will be more likely to mate with neighboring males and will prefer large, adult males. The equation of the model is:

$$P_{ij} = \frac{\exp(-A \cdot s_{ij}) \cdot t_j^B \cdot \exp(C \cdot g_j)}{\sum_{k=1}^m [\exp(-A \cdot s_{ik}) \cdot t_k^B \cdot \exp(C \cdot g_k)]},$$

in which  $s_{ik}$  represents the distances between the female and each male in the mating pool,  $t_k$  are the tarsus lengths of males, and  $g_k$  are the males binary ages.

```
LL.EPC = function(A, B, C, sik, tarsus, age, pos)
{
  pik = exp(sik*A) * (tarsus^B) * exp(age*C)
  pij = pik[pos]/apply(pik, 1, sum, na.rm=TRUE)
  return(-sum(log(pij)))
}
```

As before, we use the `mle2()` function to do the model fitting:

```
fittedEPC = mle2(minuslogl = LL.EPC, start = list(A=-1, B=0, C=0),
                 data = list(sik=sik03, tarsus=tarsus03, age=age03, pos=positions03))
fittedEPC
##
## Call:
## mle2(minuslogl = LL.EPC, start = list(A = -1, B = 0, C = 0),
##      data = list(sik = sik03, tarsus = tarsus03, age = age03,
##                  pos = positions03))
##
## Coefficients:
##           A           B           C
## -1.3575581 -0.8321577  1.3658255
##
## Log-likelihood: -121.43
```

As expected, we observe a negative value of  $A$  and a positive value of  $C$ . Curiously, we found a negative value of  $B$ , which would mean that females have preference for smaller males. As before, you can check 95% confidence intervals or likelihood intervals using functions `confint()` and `profile()` if you like.

### Part 3 - Blue tit extra-copulation data with a random factor

Now, we go for the tricky part: we will include a random factor in the parameter  $A$ , which measures the importance of distance. Thus, we will calculate one  $A$  value for each female, which we may call  $A_i$ . Each  $A_i$  value is calculated as the overall  $A$  value plus a value  $\alpha_i$  for each female. To do this, we will write a new function that calculates the likelihood of a specific  $\alpha$  value for an individual female given all the data regarding that female and certain parameter values ( $A$ ,  $B$  and  $C$ ).

The data on specific females will sometimes come in the form of small matrices and sometimes in the form of vectors (when there is only one EPC for each female). This is why the function works slightly different if it receives matrices or vectors.

```
alpha = function(alphai=0, A, B, C, siki, tarsusi, agei, posi)
{
  ms = exp(siki*A+alphai) * (tarsusi^B) * exp(agei*C)
  if(class(siki) == "matrix")
  {
    rownames(ms) = 1:nrow(ms)#renames ms rows to allow indexation
    Ps = ms[cbind(1:nrow(ms),posi[,2])] #posi[,2] contains the position of males
    sumMs = apply(ms,1,sum,na.rm=TRUE)
    return(sum(-log(Ps/sumMs))) #calculates log-likelihoods
  }
  else
  {
    #does the same thing as above, but considering its just one line of data
    Ps = ms[sirePosi[2]]
    sumMs = sum(ms, na.rm=TRUE)
    return(-log(Ps/sumMs))
  }
}
```

In order to work, the optimization will need to use the object *epp* as a global variable, and also will need an object with the IDs of all extra-pair copulating females that we will call *eppFemales*, which will also be used as a global variable.

```
eppFemales = unique(epp$female)
```

The linear optimization must impose boundaries to the alpha values, otherwise the overall optimization will not converge. We cannot allow individual alpha values to be higher than the overall *A* value, so we will impose that each alpha value must be within  $+0.9 \cdot A$  and  $-0.9 \cdot A$ .

```
LL.EPC.plus = function(A, B, C, sik, tarsus, age, pos)
{
  #vector of random coefficients for the individual females
  alphaVec = rep(0, length(eppFemales))

  #runs through the females optimizing the individual coefficients (alphas)
  if(A!=0)
    for(i in 1:length(eppFemales))
      {
        #we will pick only the lines of this individual's EPCs
        siki = sik[epp$female == eppFemales[i],]
        tarsusi = tarsus[epp$female == eppFemales[i],]
        agei = age[epp$female == eppFemales[i],]
        posi = pos[epp$female == eppFemales[i],]
        int = c(-A*0.9, A*0.9) #boundaries for the optimization

        #actual linear optimization of the individual alpha
        #notice that it receives all the data
        alphaVec[i] = optimize(f=alpha, interval = int,
                              siki=siki, tarsusi=tarsusi,
                              agei=agei, posi=posi,
                              A=A, B=B, C=C)$minimum
      }

  #organizes the A+alpha values into a matrix
  names(alphaVec) = eppFemales
  alphaVec = alphaVec[epp$female]
  aMatrix = matrix(A+alphaVec, nrow=length(alphaVec), ncol=ncol(sik),byrow=FALSE)

  #calculates the log likelihood for all EPCs
  m = exp(sik*aMatrix) * (tarsus^B) * exp(age*C)
```

```

rownames(m) = 1:nrow(epp)
colnames(m) = colnames(sik)

pij = m[pos]
sumPik = apply(m, 1, sum, na.rm=TRUE)
pij = pij/sumPik #calculates log-likelihood

return(sum(-log(pij), na.rm=TRUE))
}

```

To continue, we need to load the package *optimx* and fit the model using function `mle2()` as before:

```

library(optimx)

fittedEPC.plus = mle2(minuslogl = LL.EPC.plus, start = list(A=-1, B=0, C=0), data = list(sik=sik03, tarsus=tarsus03, age=age03, pos=positions03))

fittedEPC.plus
##
## Call:
## mle2(minuslogl = LL.EPC.plus, start = list(A = -1, B = 0, C = 0),
##      data = list(sik = sik03, tarsus = tarsus03, age = age03,
##                  pos = positions03))
##
## Coefficients:
##           A           B           C
## 23.22431  347.33020 -162.76780
##
## Log-likelihood: 0

```

The effect of space (parameter *A*) continues negative, as expected. But now parameters *B* and *C* are very close to zero. Perhaps this is a problem with sample size, since we picked only one year to do this analysis. Anyway, the goal of this tutorial was to be a walkthrough by the process of fitting the multinomial model and adding a random factor to it using linear optimization. The code may not be very simple, but running it in your machine, perhaps trying to use your own data, should make things clearer. In case of any doubt, please do not hesitate to contact us.

## Capítulo 3

### Past mating frequency and social environment lead to strategic movements in a scramble competition polygyny<sup>4</sup>

Danilo G. Muniz<sup>1,2</sup>, Martha L. Baena<sup>3</sup>, Rogelio Macías-Ordóñez<sup>4</sup> and Glauco Machado<sup>2</sup>



1. Programa de Pós-Graduação em Ecologia, Departamento de Ecologia, Instituto de Biociências, Universidade de São Paulo
2. LAGE do Departamento de Ecologia, Instituto de Biociências, Universidade de São Paulo
3. Instituto de Investigaciones Biológicas, Universidad de Veracruz
4. Red de Biología Evolutiva, Instituto de Ecología A.C. Xalapa

---

<sup>4</sup> Capítulo formatado de acordo com as normas da revista *Journal of Animal Ecology*.

## Summary

1. In scramble competition polygynies mate search has been suggested to be the main tactic employed by the individuals to acquire mates. In fact, a positive relationship between individual mobility and mating frequency has already been reported for several animal species. However, a negative relationship has been found for males and females of a few species, including the leaf beetle *Leptinotarsa undecimlineata*.

2. We hypothesize that individuals move in response to their mating frequency, and tested two predictions using *L. undecimlineata* as model system: (a) individuals should move more between host plants when past mating frequency and encounter rate with potential mates are low, and this effect should be more pronounced when they are in spatially central plants, as a way to save energy and reduce predation risk; (b) individuals should move to nearby plants with high number of potential mates.

3. Using detailed day-by-day data on individual location, copulation frequency, and social environment (i.e., presence of males and females in each host plant), we found that plant switches of both males and females are negatively influenced by past mating frequency and encounter rate with potential mates, and positively influenced by plant centrality. Moreover, individuals were more likely to switch to nearby plants with high number of conspecifics, regardless of their sex.

4. Our results suggest that males and females are moving strategically, choosing nearby plants to diminish movement costs and preferring plants with many conspecifics to increase mating opportunities.

5. We propose that strategic movement during mate search is likely to be common not only in scramble competition polygynies, but in any mating system in which mate search is an important component of reproductive success.

**Key-words:** centrality, Chrysomelidae, encounter rate, mate search, spatial distribution

## Introduction

It is often assumed that males perform most or all mate search because they would be more benefited from multiple mating while females would need only one or few copulations to fertilize all their eggs. However, recent empirical evidence suggests that mate search by females is to be expected when females benefit from multiple mating (Parker & Birkhead 2013) and when they risk mate failure (Rhainds 2010). In addition, theory does not support the notion that males will always be the searching sex. Even when benefits from multiple mating are higher for males than for females, mathematical models indicate that male and female mate search are equally likely to evolve (Kokko & Wong 2007). Thus, empirical and theoretical work suggest that mate search should be studied in males and females, and that hypotheses on the benefits of mate search should be tested for both sexes (Kokko et al. 2013).

Scramble competition polygynies are mating systems in which males do not defend mating resources, harems or individual females for prolonged periods, but rather continually search for females (Thornhill & Alcock 1983). Given that search is the main tactic employed by males to acquire mates in scramble competition polygynies, it is expected that more mobile males have higher mating success. In fact, a positive relationship between male movement and mating frequency has been observed for reptiles (e.g. Madsen 1993), mammals (e.g. Stockley 1994; Sandell 1996), and insects (e.g. Alcock 1980). However, a negative relationship has also been found for a few species (e.g. Brown & Weatherhead 1999; Baena & Macías-Ordóñez 2015). For females, the relationship between movement and mating frequency is seldom studied, but both positive (e.g. Rank 2006) and negative relationships have already been reported (e.g. Baena & Macías-Ordóñez 2015). For both males and females, it has been recently hypothesized that a negative relationship between movement and mating frequency can be the result of individuals moving when they do not obtain copulations while remaining stationary (Baena & Macías-Ordóñez 2015). This hypothesis reverses the cause-and-effect relationship according to which increased movements lead to more copulations in scramble competition polygynies, and establishes that, in turn, low copulation frequency increases movement rates.

If we hypothesize that individuals move in response to their mating frequency, two questions immediately arise: when to move and where to go? Regarding the *when* question, individuals should leave their locations when they either do not encounter potential mates or do not mate for some other reason, such as being rejected by females or because of the presence of more competitive males. Also, the spatial structure of the environment is likely

to play an important role in movement decisions. When mating sites are spatially distributed in discrete patches, such as the water pools used by many amphibians, the webs of female spiders or the host plants used by phytophagous insects, a preference to switch between nearby sites may be a way of saving energy and decreasing predation risk during displacement (e.g. Polis et al. 1998; Kasumovic et al. 2006). In this scenario, individuals in central mating sites, surrounded by many alternative sites, should be more likely to move than those in isolated sites. Regarding the *where* question, in addition to move preferentially to nearby sites, individuals should prefer moving to places with as many potential mates and as few competitors as possible. However, if individuals are unable to discriminate potential mates from competitors at distance, they should just move to crowded areas where the probability of finding a mate is likely to be higher.

Here we tested the hypothesis that copulation frequency influences the movements between plants in the phytophagous leaf beetle *Leptinotarsa undecimlineata* (Chrysomelidae). This is an insect species with scramble competition polygyny mating system in which mobility is negatively correlated with mating frequency for both males and females (Baena & Macías-Ordóñez 2015). Males only defend females during copulation and immediately after it, so that their main mating tactic is to chase nearby females on the host plant (Baena & Macías-Ordóñez 2012). Using detailed day-by-day data on individual location, copulation frequency, and social environment in each host plant, we tested the following predictions: 1) individuals are more likely to switch between host plants when mating frequency and encounter rate with potential mates are low, and this effect should be more pronounced when they are in central host plants; and 2) if individuals switch host plants according to these criteria, we expect them to move preferentially to nearby plants with high number of potential mates and low number of potential competitors (same sex individuals).

## Methods

### STUDY ORGANISM

Larvae and adults of *L. undecimlineata* feed exclusively on two host plant species, *Solanum lanceolatum* and *S. chrysothricum* (Solanaceae), which are also used as mating and oviposition sites. Both adult and larvae spend all the breeding season on the host plants and are rarely found elsewhere. Males and females mate repeatedly and with many individuals through the mating season. Females usually oviposit soon after copulation, and males usually stay mounted on females or remain near them performing mate guarding. Males may fight to

displace other males during copulation, but they do not perform prolonged territorial or female defence and present no weapon or ornament (Baena & Macías-Ordóñez 2012, 2015).

## STUDY SITE AND DATASET

In this study, we used a subset of the dataset of Baena & Macías-Ordóñez (2012, 2015). The original data were collected in a secondary forest replacing an abandoned pasture grassland, close to a fragment of cloud forest at a location known as El Riscal, Central Veracruz, Mexico (19°28'56"N, 96°59'48 "W, 1595 m asl), during the breeding season of 2004. The mean annual temperature is 20 °C (min-max = 12 - 34 °C) and annual precipitation is more than 3000 mm. The studied population lived on a patch of 75 adult host plants of *S. lanceolatum* and 8 of *S. chrysotrichumin* in an area of about 400 m<sup>2</sup>. All host plants and leaf beetles were individually marked (for more details on the marking procedures and sampling design, see Baena & Macías-Ordóñez 2012, 2015). Plants were visited once a day from July 21 to November 7 2004, which comprises the entire breeding season of *L. undecimlineata*.

In each day, the identity, location (host plant), and mating activities of all individuals present in the patch were recorded. Copulations occurred within an 80 day-period, which was divided in four 20-day periods in a previous paper (Baena & Macías-Ordóñez 2012). Each period is characterized by different beetle densities, mating frequencies, and sexual selection intensities. To test our hypotheses, we used data only from the first three periods, when most of the copulations are concentrated totalling 60 days of daily observations. This subset includes information about 306 males and 207 females. Exact spatial coordinates, measured with a measuring tape as linear distances from a standardized point, were available for 73 of the host-plants in the studied patch. We used only information of switches among these spatially located plants. We calculated distances between any two plants using Euclidean distances, and estimated spatial centrality as the inverse of the mean distance to all other plants with known coordinates.

## WHEN TO MOVE?

To test the prediction that individuals are more prone to move when mating frequency and encounter rate with potential mates are low, we used generalized linear mixed models (GLMMs) with binomial distributions (Gelman & Hill 2006). We built separate models for males and females. In both models, each recapture of an individual, which we call 'focal

recapture', was a sampling unit, and the binary response variable was the presence of the individual in the same host plant as it was found before, or if it had switched plants. The predictor variables were: (a) the spatial centrality of the host plant where the individual was previously to the focal recapture, (b) the individual's mating frequency, and (c) the encounter rate with potential mates in the day previous to the recapture. We tested models in which these variables were calculated for intervals of up to five days prior to the recapture, but the previous day was the most informative for both males and females. We excluded all observations of individuals that had not been observed at least once in the five days previous to the focal recapture. For our purposes, mating frequency was the number of observed copulations, and encounter rate with potential mates was the number of individuals of the opposite sex observed in the same host plant as the focal individual.

We built models including all three predictor variables and all possible combinations of one or two of these. We also included interaction factors between mating frequency and spatial centrality, and between encounter rate and spatial centrality. In addition to the fixed factors, we included random factors of beetle identity and plant identity in all models summing up to 13 models. We ranked the models by the Akaike Information Criterion (AIC) and considered models with a  $\Delta\text{AIC} < 2$  as equally plausible (Gelman & Hill 2006).

## WHERE TO GO?

To test the prediction that individuals move preferentially to nearby plants with high number of potential mates and low number of potential competitors, we used a spatially explicit model of choice that is based upon a multinomial distribution (see Chapter 2). In this model, each plant switch was considered a sampling unit in which the individual chose a host plant to switch to. The model considers that an individual can move to any host plant in the study area and generates a probability for each plant based on the predictor variables. Following the likelihood principle, we consider these probabilities as likelihoods and fit the model via maximum likelihood using the observed data (Royall 1997). Since it is not clear if individuals of *L. undecimlineata* can differentiate host plants with high number of potential mates, from plants simply with many individuals, we compared three alternative models of plant choice for males and females: 1) the only predictor is spatial distance between plants, 2) the predictors are distance and the total number of individuals per plant, and 3) the predictors are distance, the number of males, and the number of females per plant. Again we built separate models for males and females.

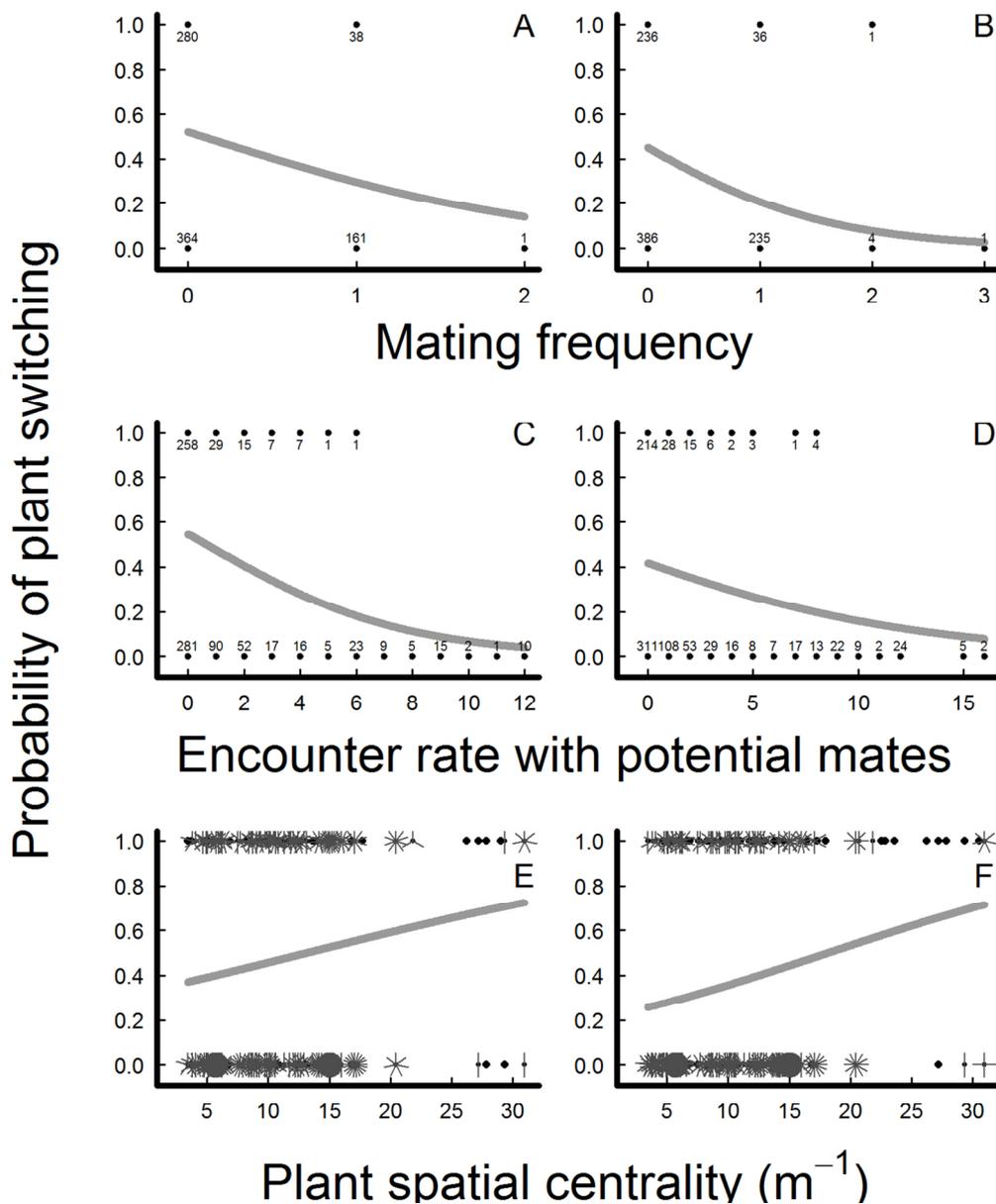
In each model, the probabilities  $P_{ij}$  of individual  $i$  moving to host plant  $j$  in each model were calculated by the following formulae 1)  $P_{ij} = \frac{\exp(A_i \cdot s_{ij})}{\sum_{k=1}^n \exp(A_i \cdot s_{ik})}$  ; 2)  $P_{ij} = \frac{\exp(A_i \cdot s_{ij}) \cdot b_j^B}{\sum_{k=1}^n [\exp(A_i \cdot s_{ik}) \cdot b_k^B]}$  ; 3)  $P_{ij} = \frac{\exp(A_i \cdot s_{ij}) \cdot f_j^B \cdot m_j^C}{\sum_{k=1}^n [\exp(A_i \cdot s_{ik}) \cdot f_k^B \cdot m_k^C]}$  ; in which the probability of host plant  $j$  being chosen is calculated as a multiplication of terms divided by sum of the same multiplication for all  $n$  plants available, following the multinomial distribution and ensuring that the probabilities add to one. The term  $s_{ij}$  is the spatial distance between the host plant where individual  $i$  was last observed and the host plant  $j$ ;  $b_j$  is the total number of individuals of *L. undecimlineata* on the host plant  $j$ ;  $f_j$  is the number of females on the host plant  $j$ , and  $m_j$  is the number of males (so that  $b_j = f_j + m_j$ ). The values  $A_i$ ,  $B$ , and  $C$  are the model coefficients, and since this model also includes repeated measures of the same individuals, we included a random factor of individual identity in the distance coefficient  $A_i$ , so that for each individual we calculated a value of  $A_i$  (hence the index in this exponent). We estimated these individual factors using linear optimization procedures (Burnham & Anderson 1998).

We performed all statistical analysis using the software R (R Development Core Team 2014), and fitted all models using the packages *bbmle* (Bolker and R Development Core Team 2014), *MuMIn* (Barton 2014) and *optimx* (Nash and Varadhan 2011).

## Results

### WHEN TO MOVE?

Our analysis included a total of 919 female and 844 male recaptures, with a mean probability of plant switching of 0.307 for females and 0.377 for males. For both males and females the most plausible model of plant switches included plant centrality, mating frequency, and encounter rate with potential mates as explanatory variables, but no interaction between these variables. In both cases, the best model included mating frequency and encounter rate calculated for a single day prior to the focal recapture (Table 1). For both males and females, higher mating frequency and encounter rate with potential mates decreased the probability of plant switching (Fig. 1A-D), while plant centrality increased switching probability (Fig. 1E-F). For males, the strongest effect was that of encounter rate with females, whereas for females the mating frequency was the strongest explanatory variable of plant switches (Table 1). For both males and females, plant centrality had the smallest effect size (Table 1).



**Figure 1.** Relationship between plant switches by males and females of the leaf beetle *Leptisotarsa undecimlineata* and three predictor variables: mating frequency in the previous day (A = males, B = females), encounter rate with potential mates in the previous day (C = males, D = females), and plant spatial centrality (E = males, F = females). Encounter rate with potential mates was calculated as the number of opposite sex conspecifics observed in the same host plant as the focal individual. Plant centrality was calculated as the inverse of the mean distance from a host plant to all other host plants in the study area and is a measure of proximity to other plants. Grey lines represent the expected probabilities according to the best model of plant switches (see Table 1). In A, B, C and D the numbers above the points represent the number of observations with the same value. In E and F the points represent single observations, "flowers" represent multiple observations with the same value, and each "petal" represents one observation.

**Table 1.** Summary of the generalized linear mixed models of plant switches by males and females of the leaf beetle *Leptinotarsa undecimlineata*. Predictor variables have been centred and standardized prior to model fitting, so that model coefficients are comparable effect sizes. ER: encounter rate with potential mates; MF: mating frequency; SC: spatial centrality of the plant. The symbols + and \* indicate, respectively, additive and interactive effects between variables.  $\Delta$ AIC: difference in the Akaike Information Criterion between each model and the best model; df: degrees of freedom. We present only the models in which ER and MF were calculated based on the day previous to the focal observations. Models with  $\Delta$ AIC  $\leq$  2 are highlighted in bold.

Model description	Model coefficients						df	$\Delta$ AIC	Akaike weight
	(Int)	ER	MF	SC	SC*ER	SC*MF			
Male movement									
<b>Effect of ER+ MF + SC</b>	<b>-0.14</b>	<b>-0.66</b>	<b>-0.41</b>	<b>0.28</b>	-	-	<b>6</b>	<b>0</b>	<b>0.4</b>
<b>Effect of ER+ MF+ SC + SC*MF</b>	<b>-0.14</b>	<b>-0.66</b>	<b>-0.41</b>	<b>0.27</b>	-	<b>-0.08</b>	<b>7</b>	<b>1.46</b>	<b>0.19</b>
<b>Effect of ER + MF</b>	<b>-0.12</b>	<b>-0.66</b>	<b>-0.41</b>	-	-	-	<b>5</b>	<b>1.72</b>	<b>0.17</b>
Effect of ER+ MF+ SC + SC*ER	-0.14	-0.66	-0.41	0.28	-0.01	-	7	2	0.15
Effect of ER+ MF+ SC+ SC*ER + SC*MF	-0.14	-0.67	-0.4	0.29	0.09	-0.11	8	3.33	0.08
Effect of ER + SC	-0.25	-1.08	-	0.28	-	-	5	10.09	<0.01
Effect only of ER	-0.23	-1.09	-	-	-	-	4	12.07	<0.01
Effect of ER+ SC + SC*ER	-0.25	-1.08	-	0.29	0.02	-	6	12.08	<0.01
Effect of MF + SC	0.03	-	-0.62	0.29	-	-	5	13.38	<0.01
Effect of MF+ SC + SC*MF	0.03	-	-0.62	0.28	-	-0.08	6	14.82	<0.01
Effect only of MF	0.06	-	-0.63	-	-	-	4	15.17	<0.01
Effect only of SC	0.04	-	-	0.31	-	-	4	56.47	<0.01
Null model	0.07	-	-	-	-	-	3	58.96	<0.01
Female movement									
<b>Effect of ER+ MF + SC</b>	<b>-0.56</b>	<b>-0.42</b>	<b>-0.55</b>	<b>0.38</b>	-	-	<b>6</b>	<b>0</b>	<b>0.42</b>
<b>Effect of ER+ MF+ SC + SC*ER</b>	<b>-0.56</b>	<b>-0.41</b>	<b>-0.56</b>	<b>0.43</b>	<b>0.13</b>	-	<b>7</b>	<b>1.48</b>	<b>0.2</b>
<b>Effect of ER+ MF+ SC + SC*MF</b>	<b>-0.56</b>	<b>-0.42</b>	<b>-0.56</b>	<b>0.4</b>	-	<b>0.07</b>	<b>7</b>	<b>1.6</b>	<b>0.19</b>
Effect of ER+ MF+ SC+ SC*ER + SC*MF	-0.56	-0.41	-0.56	0.43	0.1	0.05	8	3.36	0.08
Effect of MF + SC	-0.43	-	-0.71	0.4	-	-	5	4	0.06
Effect of MF+ SC + SC*MF	-0.44	-	-0.71	0.42	-	0.08	6	5.53	0.03
Effect of ER + MF	-0.52	-0.43	-0.55	-	-	-	5	5.79	0.02
Effect only of MF	-0.39	-	-0.7	-	-	-	4	9.8	<0.01
Effect of ER + SC	-0.66	-0.95	-	0.36	-	-	5	18.26	<0.01
Effect of ER+ SC + SC*ER	-0.66	-0.96	-	0.41	0.11	-	6	20.01	<0.01
Effect only of ER	-0.62	-0.95	-	-	-	-	4	23.81	<0.01
Effect only of SC	-0.32	-	-	0.38	-	-	4	51.18	<0.01
Null model	-0.28	-	-	-	-	-	3	56.48	<0.01

For both males and females, some of the models with  $\Delta\text{AIC} < 2$  included the interaction between centrality and mating frequency or between centrality and encounter rate. For males, the effect size of the interaction between centrality and encounter rate was negligible (-0.01), but the interaction between centrality and mating frequency had slightly negative effect (-0.08). For females, both interactions were positive, and the effect size of the interaction between centrality and mating frequency was weaker (0.07) than the interaction between centrality and encounter rate with males (0.13).

## WHERE TO GO?

This analysis included 407 plant switches by 161 females and 495 plant switches by 215 males. The best model explaining plant choice between switches both for males and female was the model including both distance and total number of individuals per plant (Table 2). The median distance coefficient  $A_i$  in all models was negative, and its likelihood interval did not overlap zero, so that all models indicate that individuals are more likely to switch between nearby plants. The coefficient  $B$  in the models was positive, so that individuals were more likely to move to plants with more conspecifics.

**Table 2.** Summary of the multinomial network models of plant choice by males and females of the leaf beetle *Leptinotarsa undecimlineata*.  $\Delta\text{AIC}$ : difference in the Akaike Information Criterion between each model and the best model. The ‘Distance’ coefficient is the median of the individual coefficients.

Model description	Model coefficients				$\Delta\text{AIC}$	Akaike weight
	Distance	Conspecifics	Females	Males		
<b>Male Movement</b>						
Multiplicative effect of distance and conspecifics	-5.63	1.02	-	-	0	0.985
Multiplicative effect of distance, males and females	-5.64	-	0.53	0.86	8.3	0.015
Effect only of distance	-1.00	-	-	-	950.7	<0.01
<b>Female movement</b>						
Multiplicative effect of distance and conspecifics	-9.25	2.17	-	-	0	1
Multiplicative effect of distance, males and females	-9.23	-	4.21	-1.05	22.8	<0.01
Effect only of distance	-7.91	-	-	-	1179.3	<0.01

## Discussion

We studied the movement patterns of males and females of the leaf beetle *L. undecimlineata* to test the hypothesis that individual movements are driven by the search for mating opportunities. We found that individuals of both sexes are more likely to switch between host plants when their past mating frequency and encounter rate with potential mates are low and when they are located in plants surrounded by other plants. We also found that both males and females were more likely to move to nearby plants and to plants with high number of conspecifics, regardless of their sex. In what follows, we will discuss the implications of our findings to the relationship between individual movement and reproductive success in scramble competition polygynies.

In scramble competition polygynies, sexual selection should favour male locomotor ability (Andersson 1994). Papers that reported positive relationships between movement and mating frequency usually focused on species in which females have low density or are highly philopatric, sometimes even showing aggressive behaviours against other females (e.g. Stockley 1994, Sandell 1996, Lane et al. 2009). In these circumstances, males need to move to locate a receptive female, and once a male has copulated he must move away from the home range of that female to find another mate. The only two cases of a negative relationship between movement and mating frequency for males were reported for *L. undecimlineata* (Baena & Macías-Ordoñez 2015) and a population of the water snake *Nerodia sipedon* (Brown & Weatherhead 1999). In both cases, population density was very high, so that a male had the potential to find another mate after copulation without moving to distant areas. Whereas males would not be expected to move while they are obtaining copulations, males unable to mate would benefit from switching to areas where their chances of obtaining copulations are higher. Thus, we propose that the rule of thumb "stay where you are if you can mate, move if you cannot" could lead to different relationships between movement and mating frequency depending on the density and spatial distribution of potential mates.

Females of *L. undecimlineata* behaved similarly to the males and were more likely to switch plants when not mating or not meeting potential mates. Thus, female movement can at least partly be explained by mate search (Table 1). This relationship was expected in this species because the optimal number of copulations in chrysomelid beetles may be as high as five (Fan et al. 2015), and because large females have an overall higher mating rate, probably due to male mate choice (see discussion in Baena & Macías-Ordoñez 2015). These

characteristics are not unique of our study system; polyandry has been shown to have benefits in many species (Arnqvist & Nilsson 2000; Parker & Birkhead 2013), and male mate choice seems to be widespread in insects (Bonduriansky 2001) and other taxa (Edward & Chapman 2011). Although female mate search is not a well-studied topic, we argue that the hypothesis that females move in response to mating frequency and in search of further mates deserves more attention. According to the rationale exposed above, we expect that females will engage in mate search when polyandry is beneficial and when there is male mate choice, so that less attractive females need to move more to find a mate.

The most important predictor of plant switch for females was mating frequency, while for males it was encounter rate (Table 1). We suggest that this difference between sexes is related to the Bateman gradient of males and females. Even in polyandrous species the Bateman gradient is steeper for males, so that they gain more reproductive success per additional copulation than females (e.g. Fritzsche & Arnqvist 2013). Therefore, males should be more sensitive than females to the social environment and respond more strongly to the availability of potential mates. Males that are not encountering females should move to other mating sites to increase their chances of acquiring copulations. Females, on the other hand, should respond to the actual number of copulations and move to other mating sites to avoid the risk of mate failure.

Our results also show another sexual difference in the movement patterns. The second and third best models of female plant switching included positive interactions between plant centrality and encounter rate, and between plant centrality and mating frequency, respectively (Table 1). Females that encountered many males or mated recently were more likely to move when they were in central plants. Perhaps, once a female has guaranteed fertilization she seeks additional males or good oviposition sites on nearby plants. For males, the second best model of plant switching included a negative interaction between centrality and mating frequency (Table 1). Males with high mating frequency in central plants were less likely to switch plants than successful males in peripheral plants. If central plants are more visited than peripheral plants, successful males in central plants should remain stationary waiting for the arrival of females to avoid the costs of moving. Taken together, our results suggest that although both sexes may search for mates, there are subtle differences in movement patterns of males and females.

Both males and females moved preferentially to nearby plants and to plants with higher number of conspecifics, regardless of their sex. This pattern suggests that *L.*

*undecimlineata* individuals are able to detect conspecifics at distance, but cannot differentiate their sex, which is in accordance with experiments with other chrysomelid beetles in which males were attracted to plants baited with conspecifics regardless of their sex (Nahrung & Allen 2004). Given this sensorial constraint, the best mate search strategy is apparently to seek crowded plants nearby where the number of potential mates is perceived to be high. The coefficient of distance for females was more negative than that of males (Table 2), suggesting that females are more restricted than males to switch between nearby plants, perhaps because it is more costly for females to move given their larger body size when bearing eggs. Moving to close sites decreases the costs of movement, while searching for crowded plants increases the potential mating pay-off. Thus, both males and females of *L. undecimlineata* seem to be moving strategically in order to optimize their reproductive benefits.

The concept of optimization has permeated the food search literature for decades because the net benefit of feeding incorporates both costs associated with finding food and benefits derived from the quality and quantity of this food (Schoener 1971). The same rationale could also be applied to generate new predictions regarding mating strategies. If we assume that mate search is costly or risky (e.g. Polis et al. 1998; Byers et al. 2005; Lane et al. 2010), and that reproductive success increases with the number of copulations for males and also for females (e.g. Bergeron et al. 2012; Fritzche & Arnqvist 2013), sexual selection should favour strategic movements that minimize costs and maximize mating opportunities for individuals of both sexes. As we showed here, the optimum movement strategy may depend on individual features (mating frequency), the social environment (encounter rates), and also the spatial distribution of conspecifics.

Andersson (1994) describes *scramble* as a mechanism of mate competition in which males compete to be the first to locate a female, and argues that scramble competition polygynies are mating systems in which scramble is the main mechanism of mate competition. Following this definition, however, the scramble mechanism may also be present in other types of mating systems. For example, in a defence polygyny mating system, males compete for females via territorial contests, but frequently there are sneaker males that obtain copulations moving between territories invading them to copulate furtively (Oliveira et al. 2008). Sneaker males compete via scramble, since their copulations depend on locating territories with females and subsequently invading them. In a similar way, males seeking extra-pair copulations in monogamous species can compete via scramble, since they must locate fertile paired females to copulate (Westneat & Stewart 2003). Thus, we would expect to

find strategic movement in the search for mates not only in scramble competition polygynies, but also in any mating system in which individuals compete for mates, at least in part, via scramble. This is a new way of looking at mate search that provides testable hypotheses and can generate useful insights into the study of home ranges, spatial distribution of individuals, individual movement, and mate search.

### **Acknowledgements**

DGM and GM were sponsored by São Paulo Research Foundation (FAPESP 2011/12675-2 and 2012/50229-1). GM has a research grant from National Council for Scientific and Technological Development (CNPq).

### **References**

- Alcock, J. (1980) Natural selection and the mating systems of solitary bees. *American Scientist*, **68**, 146–153.
- Amundsen, T. & Forsgren, E. (2001) Male mate choice selects for female coloration in a fish. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 13155–13160.
- Andersson, M. (1994) *Sexual Selection*. Princeton University Press, Princeton, NJ.
- Arnqvist, G. & Nilsson, T. (2000) The evolution of polyandry: multiple mating and female fitness in insects. *Animal Behaviour*, **60**, 145–164.
- Baena, M.L. & Macías-Ordóñez, R. (2012) Phenology of scramble polygyny in a wild population of chrysolemid beetles: the opportunity for and the strength of sexual selection. *PLoS ONE*, **7**, e38315.
- Baena, M.L. & Macías-Ordóñez, R. (2015) Mobility and mating frequency in the scramble competition polygyny of a chrysomelid beetle. *Behavioral Ecology*, **26**, 416–424.
- Barton, K. (2014) MuMIn: Multi-model inference. R package version 1.10.5. Available from: <http://CRAN.R-project.org/package=MuMIn>

- Bergeron, P., Montiglio, P., Réale, D., Humphries, M.M. & Garant, D. (2012) Bateman gradients in a promiscuous mating system. *Behavioral Ecology and Sociobiology*, **66**, 1125–1130.
- Bolker, B. & R Development Core Team (2014). *bbmle: Tools for general maximum likelihood estimation*. R package version 1.0.17. Available from: <http://CRAN.R-project.org/package=bbmle>.
- Bonduriansky, R. (2001) The evolution of male mate choice in insects: a synthesis of ideas and evidence. *Biological Reviews of the Cambridge Philosophical Society*, **76**, 305–339.
- Börger, L., Dalziel, B.D. & Fryxell, J.M. (2008) Are there general mechanisms of animal home range behaviour? A review and prospects for future research. *Ecology Letters*, **11**, 637–50.
- Brown, G.P. & Weatherhead, P.J. (1999) Female distribution affects mate searching and sexual selection in male northern water snakes (*Nerodia sipedon*). *Behavioral Ecology and Sociobiology*, **47**, 9–16.
- Byers, J., Wiseman, P., Jones, L. & Roffe, T. (2005) A large cost of female mate sampling in pronghorn. *The American Naturalist*, **166**, 661–668.
- Edward, D.A. & Chapman, T. (2011) The evolution and significance of male mate choice. *Trends in ecology & evolution*, **26**, 647–54.
- Fan, H., Wang, Y., Li, J. & Zhang, G. (2014) Exposure to males reduces the benefit gained from multiple mating in female *Galerucella birmanica* Jacoby (Coleoptera: Chrysomelidae). *Behavioral Ecology and Sociobiology*, **69**, 109–116.
- Fritzsche, K. & Arnqvist, G. (2013) Homage to Bateman: sex roles predict sex differences in sexual selection. *Evolution*, **67**, 1926–36.
- Gelman, A. & Hill, J. (2006) *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, Cambridge, UK.
- Hill, G. (1993) Male mate choice and the evolution of female plumage coloration in the house finch. *Evolution*, **47**, 1515–1525.
- Lane, J.E., Boutin, S., Gunn, M.R. & Coltman, D.W. (2009) Sexually selected behaviour: red squirrel males search for reproductive success. *Journal of Animal Ecology*, **78**, 296–304.

- Lane, J.E., Boutin, S., Speakman, J.R. & Humphries, M.M. (2010) Energetic costs of male reproduction in a scramble competition mating system. *Journal of Animal Ecology*, **79**, 27–34.
- Nash, J.C. & Varadhan, R. (2011). Unifying optimization algorithms to aid software system users: optimx for R. *Journal of Statistical Software*, **43**, 1-14.
- Kasumovic, M.M., Bruce, M.J., Herberstein, M.E. & Andrade, M.C.B. (2006) Risky mate search and mate preference in the golden orb-web spider (*Nephila plumipes*). *Behavioral Ecology*, **18**, 189–195.
- Kelly, C. & Jennions, M.D. (2011) Sexual selection and sperm quantity: meta - analyses of strategic ejaculation. *Biological Reviews*, **86**, 863–84.
- Kokko, H., Klug, H. & Jennions, M.D. (2013) Mating systems. *The Evolution of Insect Mating Systems* (eds D.M. Shuker & L.W. Simmons), pp. 42–58. Oxford University Press, Oxford, UK.
- Kokko, H. & Wong, B. (2007) What determines sex roles in mate searching? *Evolution*, **61**, 1162–75.
- Madsen, T., Shine, R., Loman, J. & Håkansson, T. (1993) Determinants of mating success in male adders, *Vipera berus*. *Animal Behaviour*, **45**, 491–499.
- Nahrung, H.F. & Allen, G.R. (2004) Sexual selection under scramble competition: mate location and mate choice in the eucalypt leaf beetle *Chrysophtharta agricola* (Chapuis) in the field. *Journal of Insect Behavior*, **17**, 353–366.
- Oliveira, R.F., Taborsky, M. & Brockman, H.J. (2008) *Alternative Reproductive Tactics: An Integrative Approach*. Cambridge University Press, Cambridge.
- Parker, G.A. & Birkhead, T.R. (2013) Polyandry: the history of a revolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **368**, 20120335.
- Polis, G., Barnes, J. & Seely, M. (1998) Predation as a major cost of reproduction in Namib Desert tenebrionid beetles. *Ecology*, **79**, 2560–2566.
- R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available from: <http://www.R-project.org/>.
- Rank, N.E., Yturralde, K. & Dahlhoff, E.P. (2006) Role of contests in the scramble competition mating system of a leaf beetle. *Journal of Insect Behavior*, **19**, 699–716.

Rhainds, M. (2010) Female mating failures in insects. *Ecological Entomology*, **18**, 211–226.

Schoener, T. (1971) Theory of feeding strategies. *Annual Review of Ecology and Systematics*, **2**, 369–404.

Thornhill, R. & Alcock, J. (1983) *The Evolution of Insect Mating Systems*. Harvard University Press, Cambridge, MA.

Wedell, N., Gage, M.J.G. & Parker, G.A. (2002) Sperm competition, male prudence and sperm-limited females. *Trends in Ecology & Evolution*, **17**, 313–320.

Westneat, D.F. & Stewart, I.R.K. (2003) Extra-pair paternity in birds: causes , and conflict correlates. *Annual Review of Ecology, Evolution, and Systematics*, **34**, 365–396.

## Capítulo 4

### Mate sampling influences the intensity of sexual selection and the evolution of costly sexual ornaments<sup>5</sup>

Danilo G. Muniz<sup>1,2</sup>, Glauco Machado<sup>2</sup>



1. Programa de Pós-Graduação em Ecologia, Departamento de Ecologia, Instituto de Biociências, Universidade de São Paulo
2. LAGE do Departamento de Ecologia, Instituto de Biociências, Universidade de São Paulo

---

<sup>5</sup> Capítulo formatado de acordo com as normas da revista *Ecology Letters*.

## **Abstract**

Mate sampling, the fact that females must choose their mates from a limited sample of males in the population, influences many aspects of mate choice and mating systems. We used spatially explicit individual based simulations to investigate the effect of mate sampling on the intensity of sexual selection and its long term evolutionary effects. We found that an increase in the number of males sampled by each female promotes an increase in the intensity of sexual selection, as measured by both opportunity for sexual selection ( $I_s$ ) and selection gradient. We also ran evolutionary simulations in which females had unbounded preference for large male trait values, whereas males were under stabilizing natural selection for that same trait. As a result, an increase in the number of males sampled by each female promoted more exaggerated male traits after 50 generations. In addition, we analyzed empirical data on  $I_s$  data for a wide variety of species and found that  $I_s$  is positively correlated with an index of mate sampling based on several ecological and behavioral traits. Mate sampling may explain why lekking species are among those with the most exaggerated male ornaments, since male aggregations allow females to sample a large number of males. We argue that factors that influence the number of males sampled by each female (including, for instance, spatial distribution of individuals, movement patterns, female perception range) are important modulators of the intensity and evolutionary result of sexual selection in natural populations. We conclude that mate sampling is an important part of the mate choice process that cannot be ignored neither in empirical nor in theoretical investigations.

**Key-words:**  $I_{\text{mates}}$ , individual-based model, information filtering, mate choice, natural selection, selection gradient.

## Introduction

"80% of success is just showing up."

Woody Allen

In a population where females have preference for large male ornaments, and where females have access to all available males, we expect that males with large ornaments would copulate with most of the females (Andersson 1994). In this scenario, there would be extreme variation in male mate success, and consequently strong sexual selection (Shuster & Wade 2003). Moreover, the relationship between male relative fitness and variation in ornament size (i.e., the selection gradient *sensu* Lande & Arnold 1983), would be expected to be steep. However, in a population where females have the same preference for large male ornaments, but only have access to a small sample of all males, each female would have to choose the best male among limited options. In this second scenario, mating success would be more evenly distributed among males because even males with relatively small ornaments would have a chance of being selected by females with limited options. Therefore, the intensity of sexual selection would be weaker than in the first scenario. The selection gradient, in turn, is expected to be less steep than in the first scenario because the best option in a small sample of males is not necessarily the male with the largest ornament in the population. The limitation described in this second scenario may be called "mate sampling", i.e., the fact that during mate choice, females have access only to a sample of the males present in the mating pool at that moment.

Classically, the main factors thought to influence the evolution of male ornaments are female choosiness, which favors the exaggeration of the ornaments, and natural selection, which may act against the exaggeration of these ornaments (Fisher 1930, O'Donald 1962). However, mate sampling generates an information filtering process (Mossa et al. 2002), according to which females must take mating decisions based on limited information. As outlined above, we expect this information filtering to weaken sexual selection on male ornaments promoted by female choosiness. Consequently, we expect intense female choosiness to result in strong sexual selection and exaggerate male ornaments only when females have access to a large sample of males in the population. Under natural conditions, the size of the sample females take from the male population is likely to vary, but it is

frequently below 10 males (reviewed in Roff & Fairbairn 2014). Moreover, empirical studies have shown that mate search may be energetically costly for females, because they invest much time either looking for the best mating partners (e.g. Byers et al. 2005) or trying to escape from male harassment (e.g. Wikelski et al. 2001). These costs may explain why some females take mating decisions based on limited amount of information.

Although mate sampling was included in early models investigating the efficiency of mate choice tactics (Janetos 1980), it has been greatly overlooked in more recent theoretical work on sexual selection (but see Kokko et al. 2015). Here we use individual based simulations to investigate what are the effects of mate sampling on the intensity of sexual selection and the evolution of costly male ornaments. To investigate the short-term impact of mate sampling on the intensity of sexual selection, we simulate populations where females have preference for males with large ornaments, but have limited access to the males available for mating. To investigate the evolutionary long-term effects of mate sampling, we simulate populations where male ornament imposes a survival cost to males, so that the ornament evolves by both natural and sexual selection. In these simulations we vary the limitation of mate sampling and the intensity of natural selection promoted by female choosiness to investigate the strength of evolutionary change in male ornament throughout the time. Finally, we contrast the predictions of our model with empirical results available for a large dataset of animal species for which there is information on the intensity of sexual selection under natural conditions.

## **Methods**

### *Ecological simulations*

To investigate the effect of mate sampling on the intensity of sexual selection we built an individual based simulation (Grimm & Railsback 2005) that mimics a simplified mating system with conventional sex roles. The simulation is spatially explicit, so that males and females in the population occupy a squared landscape of  $3 \times 3$  arbitrary units where individuals of both sexes are randomly distributed. Each simulation represents one independent mating season of an entire population, which includes 1000 individuals with an even sex ratio. Each female copulates with a single male, which she must choose among the males available within a radius  $r$  from her own spatial coordinate. A female will mate with the closest available male if there is no male within the radius  $r$ . Males have a sexual

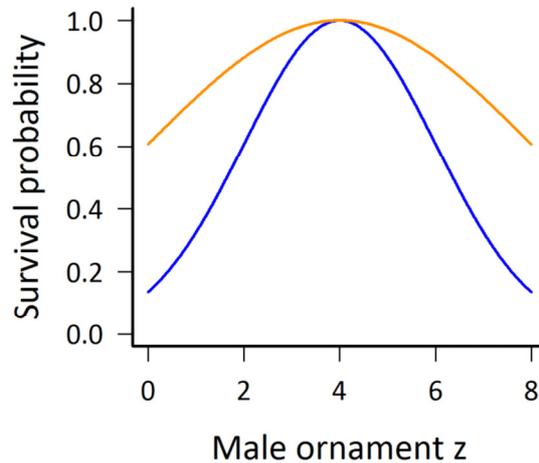
ornament  $z$  that follows a truncated normal distribution with mean  $z_{mean} = 4$ , standard deviation  $z_{sd} = 2$ , and minimum value  $z_{min} = 1$ . Females perform comparative mate choice and have directional mate preference for males with large ornaments, so that the probability of a male  $j$  being chosen by any female is proportional to  $z_j^B$ , where  $z_j$  is the male trait and  $B$  measures the intensity of female choosiness.

We ran simulations with five different values of  $r$  (0.1, 0.2, 0.4, 0.8 and 1.5). With  $r = 0.1$  females can sample, on average, only two males, representing intense information filtering, whereas with  $r = 1.5$  females can sample almost all males in the population. We also ran simulations with four different values of  $B$  representing different intensities of female choosiness, ranging from random mating ( $B = 0$ ) to weak, moderate, and strong mate choice ( $B = 1, 2, \text{ and } 4$ , respectively). We ran 50 simulation for each combination of the parameters  $r$  and  $B$  and measured the intensity of sexual selection in each using the potential for sexual selection  $I_s$  and the selection gradient on male ornament  $z$ .  $I_s$  is calculated as the standardized variance in male mating success (Arnold & Wade 1994, Shuster & Wade 2003), whereas the selection gradient is the slope of a linear regression between male ornament and standardized reproductive success (Lande & Arnold 1983). In our case, reproductive success was measured simply as mating success because we are assuming no post-copulatory process.

### *Evolutionary simulations*

To evaluate the long-term effects of mate sampling on the evolution of male ornaments we also ran a set of evolutionary simulations in which male ornaments have a survival cost and evolve under both natural and sexual selection. The model has discrete non-overlapping generations and the life-cycle can be summarized as: (1) natural selection on males; (2) mate choice; (3) mating, and (4) reproduction. After birth, all females that were born survive to reproduce, while males are subject to stabilizing natural selection, which operates as mortality conditional to the size of the sexual ornament. The probability that a male survives to reproduce is determined by a Gaussian function and is equal to  $\frac{f(z; \mu, \sigma)}{f(\mu; \mu, \sigma)}$ . In this formula,  $f(z; \mu, \sigma)$  is the probability density of  $z$  in a normal distribution with mean  $\mu$  and standard deviation  $\sigma$ , and  $f(\mu; \mu, \sigma)$  is the probability density of the mean  $\mu$  in the same distribution. This way,  $\mu$  is the value of  $z$  favored by stabilizing natural selection, where survival probability = 1. The parameter  $\sigma$  determines the intensity of natural selection, *i.e.*,

how penalized in terms of mortality are  $z$  values away from  $\mu$ . In all simulations we maintained  $\mu = 4$  and ran simulations with two intensities of natural selection: weak ( $\sigma = 4$ ) and strong ( $\sigma = 2$ ) (Fig. 1).



**Figure 1.** Survival probability of males according to ornament size ( $z$ ) in the individual based simulations of ornament evolution under natural and sexual selection. Orange: weak stabilizing natural selection; blue: strong stabilizing natural selection.

After the operation of natural selection, females perform mate choice as described in the *Ecological simulations* and copulate with their chosen males. At the beginning of each generation, 500 offspring of each sex were generated, maintaining population size and sex ratio constant throughout the simulation. For each offspring of the new generation, we sampled a female from the previous generation, with replacement, to be its mother (following Raimundo et al. 2014). This is a Poisson process and ensures that each female generates on average one male and one female offspring. Each male offspring inherits the  $z$  value of his father plus a small Gaussian error (mean = 0, SD = 0.1), females do not participate in the inheritance of  $z$ . The initial distribution of  $z$  in the populations followed a truncated normal distribution with mean  $z_{mean} = 4$ , standard deviation  $z_{sd} = 0.5$ , and minimum value  $z_{min} = 1$ . We ran the evolutionary simulations with moderate female choice ( $B = 2$ ) and all five different  $r$  values considered in the ecological simulations (Table 1). We allowed populations to evolve during 50 generations and registered the mean rate of change in  $z$  during the first 10 generation and the mean value of  $z$  after 50 generations. We measured the rate of change in  $z$  as the difference in the mean  $z$  of the population in subsequent generations. We programmed all simulations using the language R (R Core Team 2014).

### *Analysis of empirical data*

To test the effect of information filtering generated by mate sampling in the intensity of sexual selection in natural population we analyzed a dataset from the literature including estimates for several animal species belonging to a wide variety of taxa. We adopted as an estimate of sexual selection the potential for sexual selection  $I_s$ , which is the standardized variance in reproductive success (Shuster & Wade 2003). We analyzed a subset of the database compiled by Moura & Peixoto (2013), in which we included only data from non-manipulative studies conducted under natural conditions (see Supplementary Material). We included  $I_s$  estimates based on both mating success and reproductive success accessed by parentage analysis. We excluded studies in which the data collection protocol prevented the detection of males with no reproductive success, since their exclusion greatly underestimated the potential for sexual selection (Shuster & Wade 2003). In most of the dataset  $I_s$  values refer to sexual selection acting on males, but we also included in the analysis a few species from family Syngnathidae (seahorses and pipefish) and Jacanidae (jacana birds) with reversed sex roles, in which  $I_s$  refer to sexual selection acting on females.

The predictor variable was an index of mate sampling that we computed based on seven binary variables (Table 1). These variables include, for instance, the presence/absence of reproductive aggregations, characteristics of the mating signal transmission between the sexes, vagility of the individuals that perform mate search, and the type of habitat used by the species (for more details, see Table 1). We calculated the index value as the mean of all variables scored for each species in the dataset, so that the index value varied between zero and one. Zero indicates that choosing individuals may sample a small number of potential mates, whereas one indicates that choosing individuals may have access to large samples of potential mates in the population. We adopted this simple index because it is easily interpretable and allows for missing data. We included some species for which we had only incomplete data ( $n = 3$  species), but excluded those in which more than two variables were missing ( $n = 1$  species).

**Table 1.** Binary proxies of mate sampling used to calculate the mate sampling index.

<b>1. Signal transmission:</b>
(0) Short range (tactile, CHC, seismic) (1) Long range (acoustic, visual, air-borne pheromone)
Rationale: Long range signals emitted by the senders (usually males) probably allow individual receptors (usually females) to sample a larger number of mating partners in the population because the mating signals may be transmitted for longer distances.
<b>2. Modality of sexual signals:</b>
(0) Unimodal (1) Multimodal
Rationale: The use of multimodal communication, in which the same information is emitted using more than one signal with differing transmission properties, may increase the chance that sexual signals will be detected against a background of environmental noise (efficacy-base hypothesis <i>sensu</i> Hebets & Papaj 2005). Thus multimodality is expected to decrease the information filtering imposed by the environment and to increase mate sampling.
<b>3. Movement:</b>
(0) Ambulatory or slow swimmer (1) Volant or fast swimmers
Rationale: We assume that flying and fast swimmer species are able to perform longer movements when compared to ambulatory and slow swimmer species. Thus, choosing individuals (usually females) of flying and fast swimmer species may sample a large number of mating partners.
<b>4. Mate search:</b>
(0) Performed by one sex (1) Performed by both sexes
Rationale: We assume that if both sexes perform mate choice, the encounter rate between males and females will be higher, allowing choosing individuals to sample a larger number of individuals from the opposite sex.
<b>5. Vagility of the sex that performs mate search:</b>
(0) Low (territorial, cavity dwellers, nest-site defenders, trap builders) (1) High (other than the above)
Rationale: Higher vagility implies that individuals are able to move longer distances, and thus the choosing individuals may sample a large number of mating partners.
<b>6. Reproductive aggregations:</b>
(0) Absent (1) Present (leks, chorus, swarms)
Rationale: When individuals of one sex (usually males) form reproductive aggregations, the individuals of the choosing sex (usually females) can sample several mating partners at once.
<b>7. Habitat:</b>
(0) Closed (forests, kelp forest, turbid water) (1) Open (savannah, grassland, deserts, shore)
Rationale: In open habitats, signals may travel longer distances, suffering less degradation and attenuation as they travel through the environment (Hebets & Papaj 2005). Thus, individual receptors (usually females) may sample a large number of mating partners because the signals may be transmitted for longer distances.

**Footnote:** We searched for information on population density, which could also affect mate sampling. However, for most species in the dataset (see Supplemental Material), this information was not available to be included in our mate sampling index.

We analyzed the empirical data using a generalized linear mixed models (Gelman & Hill 2006), in which  $I_s$  was the response variable and the index of mate sampling was the predictor variable. Since many species in the dataset contributed to more than one value of  $I_s$ , we added species as a random factor in the analysis. To control for possible phylogenetic effects, we also included taxonomic order as a random factor, whitening which the species were nested. We compared three alternative models: (1) a null model in which  $I_s$  is independent of the mate sampling index, (2) a model in which  $I_s$  increases linearly with the index of mate sampling, and (3) a model in which  $I_s$  increases quadratically with the index of mate sampling. We ranked the models according to Akaike Information Criterion (AIC) and considered models with  $\Delta AIC < 2$  as equally plausible (Bolker 2008). We performed these analysis in the software R (R Core Team 2015) using the package *lme4* (Bates et al. 2014).

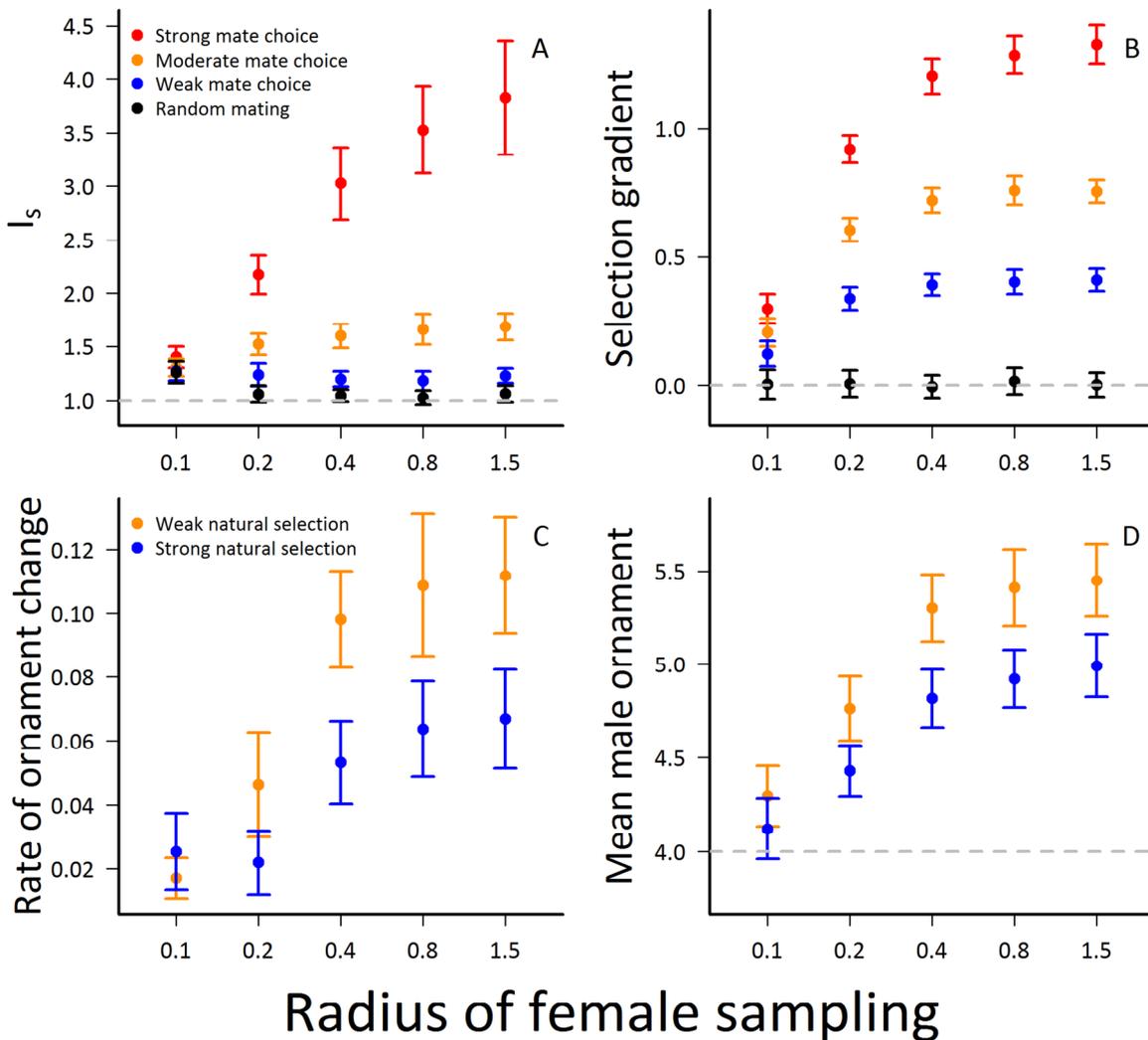
## Results

### *Ecological simulations*

In the simulations with random mating,  $I_s$  was close to 1 irrespective of the mate sampling radius (Fig. 2A). When females were choosy,  $I_s$  increased with female choosiness and also with the radius of female mate sampling (Fig. 2A). We observed similar results with the selection gradient, which was zero with random mating and increased both with female choosiness and the radius of female mate sampling (Fig. 2B).

### *Evolutionary simulations*

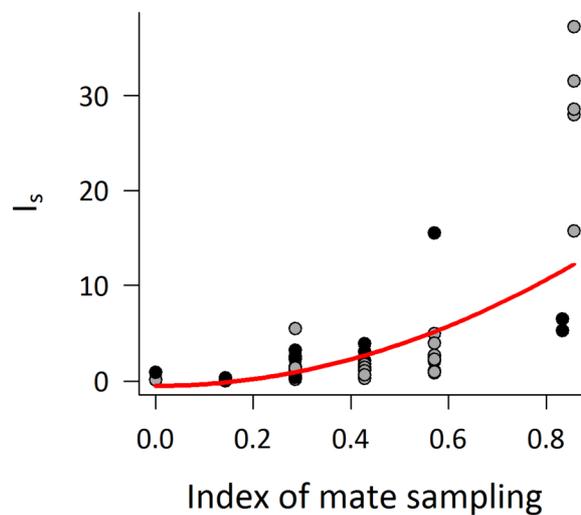
In the simulations with the most restricted radius of female sampling ( $r = 0.1$ ), the rate of evolution during the first 10 generations was close to zero, and ornament size after 50 generations did not deviate much from the value favored by natural selection (Fig. 2D). The increase in the radius of mate sampling promoted an increase in the rate of evolution (Fig. 2C), which resulted in more exaggerated male ornaments after 50 generations (Fig. 2D). Natural selection constrained both the rate of evolution and ornament size after 50 generations, so that when natural selection was relaxed male ornaments became even more exaggerated (Fig. 2D).



**Figure 2.** Results of the ecological simulations of mate sampling (A and B) and of the evolutionary simulations of male ornament evolution under natural and sexual selection (C and D). Simulations were spatially explicit and during mate choice each female could sample all males within a radius  $r$  from its own spatial position. With  $r = 0.1$  females can sample, on average, only two males, whereas with  $r = 1.5$  females can sample almost all males in the population. In all plots, points and bars represent mean and standard deviation of 50 simulated populations. (A) Potential for sexual selection over males ( $I_s$ ) under varying levels of female choosiness. (B) Selection gradient on male ornament. (C) Rate of ornament change during the first 10 generations. This rate was quantified as the mean change in the population mean of the male ornament per generation. (D) Mean male ornament after 50 generations. In (A) and (B) the grey line highlights the theoretical expectation under random mating. In (C) the grey line highlights the value favored by natural selection.

### Analysis of empirical data

This analysis included 105 estimates of  $I_s$ , from 37 different studies. The dataset included populations from 32 species belonging to 17 orders, including insects, crustaceans, fish, amphibians, lizards, mammals, and birds (Supplementary Material). The observed index of mate sampling followed a nearly normal distribution, with values varying between 0 and 0.857 (mean = 0.412). The model that best explained the relationship between  $I_s$  and the index of mate sampling was the quadratic model, so that  $I_s$  increased quadratically with the index of mate sampling (Table 2, Fig. 3).



**Fig. 3.** Relationship between the potential for sexual selection  $I_s$  and an index of mate sampling (see Table 1). Zero indicates that choosing individuals may sample a small number of potential mates, whereas one indicates that choosing individuals may have access to large samples of potential mates in the population. The red line represents the prediction from the most likely model of relationship between these variables (quadratic model, see Table 2). Black circles: values calculated from reproductive success inferred from parentage analysis. Grey-filled circles: values calculated from mating success (see Supplementary Material).

**Table 2.** Summary of the models relating the potential for sexual selection  $I_s$  and an index of mate sampling (see Table 1).  $\Delta$ AIC is the difference in the Akaike information criterion between each model and the most likely model; df: degrees of freedom.

Model	$\Delta$ AIC	df	Akaike weight
Quadratic	0	5	0.995
Linear	10.8	5	0.004
Null model	26.1	3	<0.001

## Discussion

Mate sampling was already present in theoretical models devoted to explain the best strategies of female choice and mate search (Janetos 1980). More recently, theoretical models showed the importance of mate sampling in the evolution of mate choice (Kokko et al. 2015) and the correlation between female choice and male ornament (Roff & Fairbairn 2014). In our simulations, we fixed female choice and investigated the role of mate sampling for the intensity of sexual selection acting on male ornaments and also for the evolution of these costly male traits. Our findings show that the intensity of sexual selection on male ornaments was influenced both by female choosiness for these ornaments and by the size of the sample of males taken by females in the population; sexual selection is more intense with increasing female choosiness and with increasing sample size. Furthermore, in the evolutionary simulations the rate of evolutionary change and the size of the male ornament after 50 generations increased with increasing mate sampling. The evolutionary trajectory was also influenced by natural selection, so that when natural selection was more intense, the evolutionary change was slower and resulted in less exaggerated male ornaments. The theoretical results are strongly supported by empirical data on a wide variety of animal taxa. In species where females can potentially sample a large number of males in the population the intensity of sexual selection is stronger than in species where females are somehow constrained to sample a small number of males.

Both in the ecological and in the evolutionary simulations, we varied the size of the sample taken by females by changing the spatial radius of mate sampling, which is analogous to changing the radius of female perception or the mate searching effort. However, in our mate sampling index, we included not only the presence/absence of reproductive aggregations, but also information about individual mobility, signal transmission, mate search effort, and type habitat (Table 1), in an attempt to capture the highest amount of biological traits related mate sampling. We did so because many factors other than the spatial distribution of individuals probably influence the number of potential partners an individual can sample during mate choice in natural populations (e.g. Rintamäki et al. 1995, Backwell & Passmore 1996). One of the main factors is probably the cost of mate search (e.g. Polis et al. 1998, Byers et al. 2005, Lane et al. 2010). Independent of the sex performing mate search, the more costly in terms of energy or survival risks is mate search, the more advantageous it is for individuals to diminish their mate searching efforts and mate with the closest available option. In fact, this type of mate sampling tactic has already been

described for several species, including fish (Forsgren 1997), crickets (Beckers & Wagner 2011), and sea iguanas (Wikelski et al 2001).

The evolutionary effect of mate sampling observed in our simulations may be a possible explanation on why lekking species are among those with the most exaggerated sexual ornaments in nature (Höglund & Alatalo 1995). Male aggregations in leks allow females to sample a large number of males, which according to our model, should intensify sexual selection on male ornaments. This rationale is not new, especially in the literature of sexual selection in birds (e.g., Wagner 1998). However, as we showed here, even when males are not spatially aggregated the size of the sample taken by females can influence sexual selection. Leks, therefore, represent only an extreme example in which male aggregation allows females to sample many males, resulting in a very strong bias in mating success among males (e.g., McDonald 1989). Most interestingly, our simulations and the analysis of empirical data show that mate sampling can influence sexual selection even when there are no reproductive aggregations. For instance, some species emit long range acoustic signals that allow females to access information about males from long distances (e.g., Pröhl & Hödl 1999), whereas in other species females emit airborne pheromones, attracting a large number of males (e.g., Mendoza-Cuenca & Macías-Ordóñez 2010). Such mechanisms of long-range communication, as well as differences in mate searching effort, mobility, and density are likely to modulate mate sampling in natural populations.

In our analysis of empirical data, we found a quadratic increase of the intensity of sexual selection with increasing potential for mate sampling (Fig. 3). This response goes in the same direction as predicted by our simulation, but with a different shape, since  $I_s$  in our simulations increased asymptotically with female sampling radius (Fig. 1A). One possible explanation for this discrepancy is that mate sampling and female choosiness in natural populations are positively correlated, whereas in our simulations they are independent traits. The model by Roff & Fairbairn (2014) predicts that when females are allowed to sample more males, increased female choosiness will evolve. Thus, if the populations with increased mate sampling index are also those with the more selective females, these two factors would contribute synergistically to an increase in the intensity of sexual selection. In this scenario, the expected pattern would be of a quadratic increase in sexual selection with increasing sampling by females.

Recent empirical work has explored mate sampling tactics employed by females, focusing mostly on the cognitive and perceptual processes associated to mate choice, and also on the effect of sampling experience on female choosiness (e.g. Buchholz 2004, Beckers & Wagner 2011). Our results, along with the findings of previous theoretical studies (see also Roff & Fairbairn 2014, Kokko et al. 2015), point out that mate sampling is an important part of the mate choice process that cannot be ignored in the study of mate choice either by theoreticians or empiricists. The simplifying assumption that females have access to all males in the mating pool (or at least to a representative sample) during mate choice is most likely not true in real populations. As we showed here, limitations in mate sampling can have drastic effects in both ecological and evolutionary time scales.

### **Acknowledgements**

The authors were funded by the São Paulo Research Foundation (FAPESP grants FAPESP 2011/12675-2, 2012/50229-1, 2012/20468-4 and 2013/13632-5). GM has a research grant from National Council for Scientific and Technological Development (CNPq). We are thankful to Paulo E. C. Peixoto for the help with the empirical dataset.

### **References**

- Andersson, M. (1994). *Sexual selection*. Princeton University Press, Princeton.
- Backwell, P.R.Y. & Passmore, N.I. (1996). Time constraints and multiple choice criteria in the sampling behaviour and mate choice of the fiddler crab, *Uca annulipes*. *Behav. Ecol. Sociobiol.*, 38, 407–416.
- Bates, D., Maechler, M., Bolker, B. & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7. Available from: <http://CRAN.R-project.org/package=lme4>.
- Beckers, O.M. & Wagner, W.E. (2011). Mate sampling strategy in a field cricket: Evidence for a fixed threshold strategy with last chance option. *Anim. Behav.*, 81, 519–527.
- Bolker, B.M. (2008). *Ecological models and data in R*. Princeton University Press, Princeton.

- Buchholz, R. (2004). Effects of parasitic infection on mate sampling by female wild turkeys (*Meleagris gallopavo*): Should infected females be more or less choosy? *Behav. Ecol.*, 15, 687–694.
- Byers, J., Wiseman, P., Jones, L. & Roffe, T. (2005). A large cost of female mate sampling in pronghorn. *Am. Nat.*, 166, 661–668.
- Coltman, D.W., Festa-Bianchet, M., Jorgenson, J.T. & Strobeck, C. (2002). Age-dependent sexual selection in bighorn rams. *Proc. R. Soc. London B.*, 269, 165–172.
- Fisher, R.A. (1930). *The genetical theory of natural selection*. Oxford University Press, Oxford.
- Forsgren, E. (1997). Mate sampling in a population of sand gobies. *Anim. Behav.*, 53, 267–276.
- Gelman, A. & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press, Cambridge.
- Grimm, V. & Railsback, S.F. (2005). *Individual-based modeling and ecology*. Princeton University Press, Princeton.
- Hebets, E. a., Papaj, D.R., 2005. Complex signal function: Developing a framework of testable hypotheses. *Behav. Ecol. Sociobiol.*, 57, 197–214.
- Höglund, J. & Alatalo, R. V. (1995). *Leks*. Princeton University Press, Princeton.
- Janetos, A.C. (1980). Strategies of female mate choice: a theoretical analysis. *Behav. Ecol. Sociobiol.*, 12, 107–112.
- Kokko, H., Booksmythe, I. & Jennions, M.D. (2015). Mate-sampling costs and sexy sons. *J. Evol. Biol.*, 28, 259–266.
- Lande, R., Arnold, S.J., 1983. The measurement of selection on correlated characters. *Evolution*, 37, 1210–1226.
- Lane, J.E., Boutin, S., Speakman, J.R. & Humphries, M.M. (2010). Energetic costs of male reproduction in a scramble competition mating system. *J. Anim. Ecol.*, 79, 27–34.
- Mendoza-Cuenca, L. & Macías-Ordóñez, R. (2010). Female asynchrony may drive disruptive sexual selection on male mating phenotypes in a *Heliconius* butterfly. *Behav. Ecol.*, 21, 144–152.
- Mossa, S., Barthélémy, M., Eugene Stanley, H. & Nunes Amaral, L. (2002). Truncation of power law behavior in “scale-free” network models due to information filtering. *Phys. Rev. Lett.*, 88, 138701.

- Moura, R.R. & Peixoto, P. (2013). The effect of operational sex ratio on the opportunity for sexual selection: a meta-analysis. *Anim. Behav.*, 86, 675–683.
- O'Donald, P. (1962). The theory of sexual selection. *Heredity*, 17, 541–552.
- Polis, G.A., Barnes, J.D., Seely, M.K., Henschel, J.R. & Enders, M.M. (1998). Predation as a major cost of reproduction in Namib Desert tenebrionid beetles. *Ecology*, 79, 2560–2566.
- Pröhl, H. & Hödl, W. (1999). Parental investment, potential reproductive rates, and mating system in the strawberry dart-poison frog, *Dendrobates pumilio*. *Behav. Ecol. Sociobiol.*, 46, 215–220.
- R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from: <http://www.R-project.org/>
- Raimundo, R.L.G., Gibert, J.P., Hembry, D.H. & Guimarães Jr, P.R. (2014). Conflicting selection in the course of adaptive diversification: the interplay between mutualism and intraspecific competition. *Am. Nat.*, 183, 363–75.
- Rintamaki, P.T., Alatalo, R. V., Hoglund, J. & Lundberg, a. (1995). Mate sampling behaviour of black grouse females (*Tetrao tetrix*). *Behav. Ecol. Sociobiol.*, 37, 209–215.
- Roff, D.A. & Fairbairn, D.J. (2014). The evolution of phenotypes and genetic parameters under preferential mating. *Ecol. Evol.*, 4, 2759–2776.
- Shuster, S.M. & Wade, M.J. (2003). *Mating systems and strategies*. Princeton University Press, Princeton.
- Wagner, R.H. (1998). Hidden leks: sexual selection and the clustering of avian territories. *Ornithol. Monogr.*, 49, 123–145.
- Wikelski, M., Carbone, C., Bednekoff, P. a., Choudhury, S. & Tebbich, S. (2001). Why is female choice not unanimous? Insights from costly mate sampling in marine iguanas. *Ethology*, 107, 623–638.

## Supplemental Material

1  
2  
3  
4  
5  
6  
7

**Table S1.** Full data used in the analysis of potential for sexual selection from the literature and the studies from which  $I_s$  data was retrieved. MSI is the mate sampling index, which is the mean of the seven mate sampling proxies (see Table 1 in the manuscript for details). Values of  $I_s$  highlighted with an asterisk were calculated using reproductive success data accessed by parentage analysis, the other  $I_s$  values were calculated from mating success.

Species (Order)	Proxies of mate sampling							MSI	$I_s$	Reference
	1	2	3	4	5	6	7			
Amphibians										
<i>Allobates femoralis</i> (Anura)	1	1	0	0	1	0	0	0.429	2.35*	Ursprung et. al. 2011
<i>Eleutherodactylus coqui</i> (Anura)	1	1	0	0	1	0	0	0.429	0.69	Townsend & Stewart 1994
<i>Hyla versicolor</i> (Anura)	1	1	0	0	0	1	1	0.571	4.97	Sullivan & Hinshaw 1992
<i>Hyla versicolor</i> (Anura)	1	1	0	0	0	1	1	0.571	2.56	Sullivan & Hinshaw 1992
<i>Hyla versicolor</i> (Anura)	1	1	0	0	0	1	1	0.571	0.99	Sullivan & Hinshaw 1992
<i>Incilius valliceps</i> (Anura)	1	0	0	0	1	1	1	0.571	4	Wagner & Sullivan 1992
<i>Incilius valliceps</i> (Anura)	1	0	0	0	1	1	1	0.571	2.7	Wagner & Sullivan 1992
<i>Incilius valliceps</i> (Anura)	1	0	0	0	1	1	1	0.571	2.73	Wagner & Sullivan 1992
<i>Incilius valliceps</i> (Anura)	1	0	0	0	1	1	1	0.571	2.28	Wagner & Sullivan 1992
<i>Lithobates catesbeianus</i> (Anura)	0	0	0	0	1	1	1	0.429	1.38	Arnold & Wade 1984
<i>Lithobates catesbeianus</i> (Anura)	0	0	0	0	1	1	1	0.429	0.21*	Arnold & Wade 1984
<i>Lithobates catesbeianus</i> (Anura)	0	0	0	0	1	1	1	0.429	0.16*	Arnold & Wade 1984
<i>Oophaga pumilio</i> (Anura)	1	1	0	0	1	0	0	0.429	1.76	Pröhl & Hödl 1999

<i>Oophaga pumilio</i> (Anura)	1	1	0	0	1	0	0	0.429	1.48	Pröhl 2002
<i>Oophaga pumilio</i> (Anura)	1	1	0	0	1	0	0	0.429	1.08	Pröhl 2002
<i>Oophaga pumilio</i> (Anura)	1	1	0	0	1	0	0	0.429	0.61	Pröhl 2002
<i>Oophaga pumilio</i> (Anura)	1	1	0	0	1	0	0	0.429	0.33	Pröhl 2002
<i>Oophaga pumilio</i> (Anura)	1	1	0	0	1	0	0	0.429	1.76	Pröhl 2002
<i>Oophaga pumilio</i> (Anura)	1	1	0	0	1	0	0	0.429	1.48	Pröhl 2002
<i>Oophaga pumilio</i> (Anura)	1	1	0	0	1	0	0	0.429	1.08	Pröhl 2002
<i>Taricha granulosa</i> (Caudata)	0	0	0	0	0	1	1	0.286	4.48*	Jones et. al. 2002

Birds										
<i>Chiroxiphia linearis</i> (Passeriformes)	1	1	1	0	1	1	1	0.857	28	McDonald 1989
<i>Chiroxiphia linearis</i> (Passeriformes)	1	1	1	0	1	1	1	0.857	31.5	McDonald 1989
<i>Chiroxiphia linearis</i> (Passeriformes)	1	1	1	0	1	1	1	0.857	28.6	McDonald 1989
<i>Chiroxiphia linearis</i> (Passeriformes)	1	1	1	0	1	1	1	0.857	15.8	McDonald 1989
<i>Chiroxiphia linearis</i> (Passeriformes)	1	1	1	0	1	1	1	0.857	37.2	McDonald 1989
<i>Geospiza fortis</i> (Passeriformes)	0	1	1	0	1	0	1	0.571	15.6	Price 1984
<i>Geospiza fortis</i> (Passeriformes)	0	1	1	0	1	0	1	0.571	2.11	Price 1984
<i>Geospiza fortis</i> (Passeriformes)	0	1	1	0	1	0	1	0.571	1.12	Price 1984
<i>Icterus bullockii</i> (Passeriformes)	0	0	1	-	1	0	1	0.5	0.17*	Richardson & Burke 2001
<i>Jacana jacana</i> (Charadriiformes)	0	0	0	0	0	0	1	0.143	0.34	Emlen & Wrege 2004
<i>Molothrus ater</i> (Passeriformes)	1	1	1	0	1	0	0	0.571	1.52*	Woolfenden et al. 2002
<i>Sula granti</i> (Suliformes)	0	0	1	0	1	0	1	0.429	0.85*	Maness & Anderson 2007
<i>Sula granti</i> (Suliformes)	0	0	1	0	1	0	1	0.429	0.77*	Maness & Anderson 2007
<i>Sula granti</i> (Suliformes)	0	0	1	0	1	0	1	0.429	0.77*	Maness & Anderson 2007
<i>Sula granti</i> (Suliformes)	0	0	1	0	1	0	1	0.429	0.81*	Maness & Anderson 2007

<i>Sula granti</i> (Suliformes)	0	0	1	0	1	0	1	0.429	0.85*	Maness & Anderson 2007
<i>Sula granti</i> (Suliformes)	0	0	1	0	1	0	1	0.429	0.83*	Maness & Anderson 2007
<i>Sula granti</i> (Suliformes)	0	0	1	0	1	0	1	0.429	1.12*	Maness & Anderson 2007
<i>Sula granti</i> (Suliformes)	0	0	1	0	1	0	1	0.429	1.38*	Maness & Anderson 2007
<b>Crustaceans</b>										
<i>Paracerceis sculpta</i> (Isopoda)	0	0	0	0	1	0	1	0.286	1.06	Shuster & Wade 1991a
<i>Paracerceis sculpta</i> (Isopoda)	0	0	0	0	1	0	1	0.286	1.21	Shuster & Wade 1991a
<i>Paracerceis sculpta</i> (Isopoda)	0	0	0	0	1	0	1	0.286	0.58	Shuster & Wade 1991a
<i>Paracerceis sculpta</i> (Isopoda)	0	0	0	0	1	0	1	0.286	1.05	Shuster & Wade 1991a
<i>Paracerceis sculpta</i> (Isopoda)	0	0	0	0	1	0	1	0.286	1.5	Shuster & Wade 1991a
<i>Paracerceis sculpta</i> (Isopoda)	0	0	0	0	1	0	1	0.286	2.28	Shuster & Wade 1991a
<i>Paracerceis sculpta</i> (Isopoda)	0	0	0	0	1	0	1	0.286	1.43	Shuster & Wade 1991a
<i>Paracerceis sculpta</i> (Isopoda)	0	0	0	0	1	0	1	0.286	2.55	Shuster & Wade 1991a
<i>Paracerceis sculpta</i> (Isopoda)	0	0	0	0	1	0	1	0.286	3.25	Shuster & Wade 1991a
<i>Paracerceis sculpta</i> (Isopoda)	0	0	0	0	1	0	1	0.286	1.47	Shuster & Wade 1991a
<i>Paracerceis sculpta</i> (Isopoda)	0	0	0	0	1	0	1	0.286	1.07	Shuster & Wade 1991a
<i>Paracerceis sculpta</i> (Isopoda)	0	0	0	0	1	0	1	0.286	0.44	Shuster & Wade 1991a
<i>Paracerceis sculpta</i> (Isopoda)	0	0	0	0	1	0	1	0.286	1.39	Shuster & Wade 1991b
<b>Fishes</b>										
<i>Hippocampus subelongatus</i> (Gasterosteiformes)	0	0	0	0	1	0	1	0.286	0.27	Kvarnemo et al. 2007
<i>Salmo trutta</i> (Salmoniformes)	0	0	1	0	1	1	1	0.571	6.64*	Serbezov 2010
<i>Salmo trutta</i> (Salmoniformes)	0	0	1	0	1	1	1	0.571	4.05*	Serbezov 2010
<i>Salmo trutta</i> (Salmoniformes)	0	0	1	0	1	1	1	0.571	5.46*	Serbezov 2010
<i>Syngnathus floridae</i> (Gasterosteiformes)	0	0	0	0	0	0	0	0	0.15*	Mobley & Jones 2007

<i>Syngnathus floridae</i> (Gasterosteiformes)	0	0	0	0	0	0	0	0	0.74*	Mobley & Jones 2007
<i>Syngnathus floridae</i> (Gasterosteiformes)	0	0	0	0	0	0	0	0	0.13	Mobley & Jones 2008
<i>Syngnathus floridae</i> (Gasterosteiformes)	0	0	0	0	0	0	0	0	0.15*	Mobley & Jones 2009
<i>Syngnathus floridae</i> (Gasterosteiformes)	0	0	0	0	0	0	0	0	0.53*	Mobley & Jones 2009
<i>Syngnathus floridae</i> (Gasterosteiformes)	0	0	0	0	0	0	0	0	0.37*	Mobley & Jones 2009
<i>Syngnathus floridae</i> (Gasterosteiformes)	0	0	0	0	0	0	0	0	0.74*	Mobley & Jones 2009
<i>Syngnathus floridae</i> (Gasterosteiformes)	0	0	0	0	0	0	0	0	0.59*	Mobley & Jones 2009
Insects										
<i>Colpula lativentris</i> (Hemiptera)	0	0	0	0	0	0	1	0.143	0.12	Nishida 1987
<i>Colpula lativentris</i> (Hemiptera)	0	0	0	0	0	0	1	0.143	0.32	Nishida 1987
<i>Colpula lativentris</i> (Hemiptera)	0	0	0	0	0	0	1	0.143	0.04	Nishida 1987
<i>Colpula lativentris</i> (Hemiptera)	0	0	0	0	0	0	1	0.143	0.05	Nishida 1987
<i>Enallagma hageni</i> (Odonata)	0	0	0	0	1	1	1	0.429	1.15	Fincke 1986
<i>Enallagma hageni</i> (Odonata)	0	0	0	0	1	1	1	0.429	3.08	Fincke 1986
<i>Enallagma hageni</i> (Odonata)	0	0	0	0	1	1	1	0.429	2.14	Fincke 1986
<i>Heliconius charithonia</i> (Lepidoptera)	1	0	1	0	1	0	0	0.429	3.95	Mendoza-Cuenca & Macías-Ordóñez 2010
<i>Ischnura gemina</i> (Odonata)	0	0	1	0	1	1	1	0.571	0.87	Hafernik & Garrison 1986
<i>Labidomera clivicollis</i> (Coleoptera)	0	0	0	1	1	0	0	0.286	0.53	Dickinson 1988
<i>Neacoryphus bicrucis</i> (Hemiptera)	0	0	1	0	1	0	1	0.429	0.67*	McLain 1986
<i>Neacoryphus bicrucis</i> (Hemiptera)	0	0	1	0	1	0	1	0.429	0.38*	McLain 1986
<i>Neacoryphus bicrucis</i> (Hemiptera)	0	0	1	0	1	0	1	0.429	0.29*	McLain 1986
<i>Tetraopes tetrophthalmus</i> (Coleoptera)	0	0	0	1	0	0	1	0.286	0.19	McCauley 1983
<i>Tetraopes tetrophthalmus</i> (Coleoptera)	0	0	0	1	0	0	1	0.286	0.55	McLain & Boromisa 1987
Lizards										

<i>Anolis carolinensis</i> (Squamata)	0	0	0	0	0	0	-	0	0.94	Ruby 1984
<i>Norops garmani</i> (Squamata)	0	1	0	0	1	0	0	0.286	5.5	Arnold & Wade 1984
<b>Mammals</b>										
<i>Capreolus capreolus</i> (Artiodactyla)	0	0	0	0	1	0	1	0.286	0.75*	Vanpé et al. 2007
<i>Mirounga leonina</i> (Carnivora)	1	1	0	-	1	1	1	0.833	5.32	Galimberti et. al. 2002
<i>Mirounga leonina</i> (Carnivora)	1	1	0	-	1	1	1	0.833	6.5	Galimberti et. al. 2002
<i>Ovis canadensis</i> (Artiodactyla)	0	0	0	0	1	1	1	0.429	2.88*	Coltman 2002
<i>Ovis canadensis</i> (Artiodactyla)	0	0	0	0	1	1	1	0.429	2.78*	Coltman 2002
<i>Ovis canadensis</i> (Artiodactyla)	0	0	0	0	1	1	1	0.429	3.6*	Coltman 2002
<i>Ovis canadensis</i> (Artiodactyla)	0	0	0	0	1	1	1	0.429	8.32*	Coltman 2002
<i>Ovis canadensis</i> (Artiodactyla)	0	0	0	0	1	1	1	0.429	7.09*	Coltman 2002
<i>Ovis canadensis</i> (Artiodactyla)	0	0	0	0	1	1	1	0.429	2.45*	Coltman 2002
<i>Rhinolophus ferrumequinum</i> (Chiroptera)	0	0	1	0	1	1	1	0.571	1.72*	Rossitier et al. 2006
<i>Rhinolophus ferrumequinum</i> (Chiroptera)	0	0	1	0	1	1	1	0.571	3.57*	Rossitier et al. 2006
<i>Rhinolophus ferrumequinum</i> (Chiroptera)	0	0	1	0	1	1	1	0.571	2.69*	Rossitier et al. 2006
<i>Rhinolophus ferrumequinum</i> (Chiroptera)	0	0	1	0	1	1	1	0.571	2.09*	Rossitier et al. 2006
<i>Rhinolophus ferrumequinum</i> (Chiroptera)	0	0	1	0	1	1	1	0.571	3.49*	Rossitier et al. 2006
<i>Rhinolophus ferrumequinum</i> (Chiroptera)	0	0	1	0	1	1	1	0.571	4.35*	Rossitier et al. 2006
<i>Rhinolophus ferrumequinum</i> (Chiroptera)	0	0	1	0	1	1	1	0.571	3.06*	Rossitier et al. 2006
<i>Rhinolophus ferrumequinum</i> (Chiroptera)	0	0	1	0	1	1	1	0.571	2.95*	Rossitier et al. 2006
<i>Rhinolophus ferrumequinum</i> (Chiroptera)	0	0	1	0	1	1	1	0.571	2.08*	Rossitier et al. 2006
<i>Rhinolophus ferrumequinum</i> (Chiroptera)	0	0	1	0	1	1	1	0.571	1.56*	Rossitier et al. 2006
<i>Tamias amoenus</i> (Rodentia)	1	0	0	0	1	0	1	0.429	0.86*	Schulte-Hostedde et. al. 2004
<i>Tamias amoenus</i> (Rodentia)	1	0	0	0	1	0	1	0.429	1.23*	Schulte-Hostedde et. al. 2004

<i>Tamias amoenus</i> (Rodentia)	1	0	0	0	1	0	1	0.429	1.55*	Schulte-Hostedde et. al. 2004
<i>Xerospermophilus tereticaudus</i> (Rodentia)	0	0	0	0	1	0	1	0.286	1.6*	Munroe & Koprowski 2011

---

## References

- Arnold, S. & Wade, M. (1984). On the measurement of natural and sexual selection: applications. *Evolution*, 38, 720–734.
- Dickinson, J.L. (1988). Determinants of paternity in the milkweed leaf beetle. *Behav. Ecol. Sociobiol.*, 23, 9–19.
- Emlen, S.T. & Wrege, P.H. (2004). Size dimorphism, intrasexual competition, and sexual selection in wattled Jacana (*Jacana jacana*), a sex-role-reversed shorebird in Panama. *Auk*, 121, 391.
- Fincke, O.M. (1986). Lifetime reproductive success and the opportunity for selection in a nonterritorial damselfly (Odonata: Coenagrionidae). *Evolution*, 40, 791–803.
- Galimberti, F., Fabiani, A. & Sanvito, S. (2002). Opportunity for selection in southern elephant seals (*Mirounga leonina*): the effect of spatial scale of analysis. *J. Zool.*, 256, 93–97.
- Hafernik Jr., J.E. & Garrison, R.W. (1986). Mating success and survival rate in a population of damselflies: results at variance with theory? *Am. Nat.*, 128, 353–365.
- Jones, A.G., Arguello, J.R. & Arnold, S.J. (2002). Validation of Bateman's principles: a genetic study of sexual selection and mating patterns in the rough-skinned newt. *Proc. R. Soc. London B.*, 269, 2533–2539.
- Kvarnemo, C., Moore, G.I. & Jones, A.G. (2007). Sexually selected females in the monogamous Western Australian seahorse. *Proc. R. Soc. London B.*, 274, 521–525.
- Maness, T.J. & Anderson, D.J. (2007). Serial monogamy and sex ratio bias in Nazca boobies. *Proc. R. Soc. London B.*, 274, 2047–2054.
- McCauley, D.E. (1983). An estimate of the relative opportunities for natural and sexual selection in a population of milkweed beetles. *Evolution*, 37, 701–707.
- McDonald, D.B. (1989). Cooperation under sexual selection: age-graded changes in a lekking bird. *Am. Nat.*, 134, 709–730.
- McLain, D.K. & Boromisa, R.D. (1987). Male choice, fighting ability, assortative mating and the intensity of sexual selection in the milkweed longhorn beetle, *Tetraopes tetraophthalmus* (Coleoptera, Cerambycidae). *Behav. Ecol. Sociobiol.*, 20, 239–246.

- Mclain, D.K. (1986). Resource patchiness and variation in the intensity of sexual selection in a resource defending polygynous insect species. *Oikos*, 47, 19–25.
- Mendoza-Cuenca, L. & Macías-Ordóñez, R. (2010). Female asynchrony may drive disruptive sexual selection on male mating phenotypes in a *Heliconius* butterfly. *Behav. Ecol.*, 21, 144–152.
- Mobley, K.B. & Jones, a. G. (2007). Geographical variation in the mating system of the dusky pipefish (*Syngnathus floridae*). *Mol. Ecol.*, 16, 2596–2606.
- Mobley, K.B. & Jones, A.G. (2009). Environmental, demographic, and genetic mating system variation among five geographically distinct dusky pipefish (*Syngnathus floridae*) populations. *Mol. Ecol.*, 18, 1476–1490.
- Munroe, K.E. & Koprowski, J.L. (2011). Sociality, Bateman's gradients, and the polygynandrous genetic mating system of round-tailed ground squirrels (*Xerospermophilus tereticaudus*). *Behav. Ecol. Sociobiol.*, 65, 1811–1824.
- Nishida, T. (1987). Measurement of the opportunity for natural and sexual selection in a breeding population of the coreid bug (*Colpula lativentris*). *Res. Pop. Ecol.*, 29, 271–289.
- Price, T. (1984). Sexual selection on body size, territory and plumage variables in a population of Darwin's finches. *Evolution*, 38, 327–341.
- Pröhl, H. & Hödl, W. (1999). Parental investment, potential reproductive rates, and mating system in the strawberry dart-poison frog, *Dendrobates pumilio*. *Behav. Ecol. Sociobiol.*, 46, 215–220.
- Pröhl, H. (2002). Population differences in female resource abundance, adult sex ratio, and male mating success in *Dendrobates pumilio*. *Behav. Ecol.*, 13, 175–181.
- Pröhl, H. (2005). Clutch loss affects the operational sex ratio in the strawberry poison frog *Dendrobates pumilio*. *Behav. Ecol. Sociobiol.*, 58, 310–315.
- Richardson, D.S. & Burke, T. (2001). Extrapair paternity and variance in reproductive success related to breeding density in Bullock's orioles. *Anim. Behav.*, 62, 519–525.
- Rossiter, S.J., Ransome, R.D., Faulkes, C.G., Dawson, D.A. & Jones, G. (2006). Long-term paternity skew and the opportunity for selection in a mammal with reversed sexual size dimorphism. *Mol. Ecol.*, 15, 3035–3043.

- Ruby, D.E. (1984). Male breeding success and differential access to females in *Anolis carolinensis*. *Herpetologica*, 40, 272–280.
- Schulte-Hostedde, A.I., Millar, J.S. & Gibbs, H.L. (2004). Sexual selection and mating patterns in a mammal with female-biased sexual size dimorphism. *Behav. Ecol.*, 15, 351–356.
- Serbezov, D., Bernatchez, L., Olsen, E.M. & VØllestad, L. a. (2010). Mating patterns and determinants of individual reproductive success in brown trout (*Salmo trutta*) revealed by parentage analysis of an entire stream living population. *Mol. Ecol.*, 19, 3193–3205.
- Shuster, S.M. & Wade, M.J. (1991a). Equal mating success among male reproductive strategies in a marine isopod. *Nature*, 350, 608–610.
- Shuster, S.M. & Wade, M.J. (1991b). Female copying and sexual selection in a marine isopod crustacean, *Paracerceis sculpta*. *Anim. Behav.*, 41, 1071–1078.
- Sullivan, B.K. & Hinshaw, S.H. (1992). Female choice and selection on male calling behaviour in the grey treefrog *Hyla versicolor*. *Anim. Behav.*, 44, 733–744.
- Townsend, D.S. & Stewart, M.M. (1994). Reproductive ecology of the Puerto Rican frog *Eleutherodactylus coqui*. *J. Herpetol.*, 28, 34–40.
- Ursprung, E., Ringler, M., Jehle, R. & Hödl, W. (2011). Strong male/male competition allows for nonchoosy females: High levels of polygynandry in a territorial frog with paternal care. *Mol. Ecol.*, 20, 1759–1771.
- Vanpé, C. (2007). Mating systems and sexual selection in ungulates - New insights from a territorial species with low sexual size dimorphism: the European roe deer (*Capreolus capreolus*). Doctorate Thesis. CEFÉ-INRA / Grimsö Wildl. Res. Station.
- Wagner Jr, W.E., Sullivan, B.K. & Wagner, W.E. (1992). Chorus organization in the gulf coast toad (*Bufo valliceps*): male and female behavior and the opportunity for sexual selection. *Copeia*, 1992, 647–658.
- Woolfenden, B.E., Gibbs, H.L. & Sealy, S.G. (2002). High opportunity for sexual selection in both sexes of an obligate brood parasitic bird, the brown-headed cowbird (*Molothrus ater*). *Behav. Ecol. Sociobiol.*, 52, 417–425.

## Conclusões gerais<sup>6</sup>

*If you can steal an idea, why can't you plant one there instead?*

*Saito (Inception, 2010)<sup>7</sup>*

Ao longo dos quatro capítulos que compõem esta tese, usamos simulações computacionais e dados empíricos de espécies com diferentes tipos de sistemas de acasalamento para tentar entender melhor a influência do "espaço no sexo". Se você leu toda a tese, espero que esteja disposto a comprar a ideia de que a distribuição espacial dos indivíduos pode ser um elemento importante, que influencia tanto a organização dos sistemas de acasalamento quanto a intensidade da seleção sexual. Caso você não tenha lido a tese, vou sintetizar a seguir os principais resultados que obtivemos em cada capítulo.

No primeiro capítulo, ao estudar a competição espermática no opilião *Serracutisoma proximum*, observamos que machos dos morfos pequeno e grande diferem quanto à intensidade de competição espermática enfrentada. Enquanto machos do morfo pequeno enfrentam, em média, maior intensidade de competição espermática, a variação na intensidade enfrentada por machos do morfo grande é muito maior. Além disso, machos do morfo grande que possuem mais fêmeas em seus haréns enfrentam, em média, menos competição espermática do que machos com menos fêmeas. Não observamos um efeito da localização dos territórios sobre a competição espermática enfrentada pelos machos do morfo grande. Entretanto, machos do morfo pequeno tendem a invadir e copular com fêmeas pertencentes a haréns espacialmente próximos.

No segundo capítulo, propusemos um modelo probabilístico de escolha de parceiros que permite a inclusão de variáveis par-a-par como preditoras de

---

<sup>6</sup> Assim como na introdução geral, as citações seguem o padrão do periódico *Advances in Space Research*.

<sup>7</sup> *Inception* é um filme de 2010, escrito e dirigido por Christopher Nolan e distribuído no Brasil como *A Origem*.

acasalamentos. Com isso, nosso modelo permite a investigação de hipóteses sobre acasalamento assortativo e permite a inclusão de distância espacial entre os indivíduos como uma das variáveis preditoras. Testamos o desempenho do nosso modelo e o comparamos com modelos alternativos analisando dados gerados por simulações baseadas em indivíduos. Descobrimos que os métodos estatísticos tradicionalmente usados para investigar a escolha de parceiros subestimam a seletividade das fêmeas. Isso ocorre, pois tais métodos ignoram a estrutura espacial das populações, pressupondo erroneamente que cada fêmea tem acesso a todos os machos da população. Ao incluir explicitamente as distâncias entre os indivíduos, nosso modelo considera que os encontros entre vizinhos devem ser mais prováveis do que entre indivíduos distantes. Com isso, somos capazes de produzir estimativas mais acuradas de seletividade feminina.

No terceiro capítulo, investigamos a movimentação de machos e fêmeas do besouro *Leptinotarsa umdecimlineata* e encontramos evidências de que os movimentos de machos e fêmeas são, ao menos em parte, motivados pela busca por parceiros. Observamos que os indivíduos estão mais propensos a se mover quando não encontram parceiros em potencial, quando não conseguem obter cópulas e quando estão em plantas-hospedeiras cercadas por muitas outras plantas. Dependendo da densidade dos membros do sexo oposto, essa mesma regra pode gerar uma relação positiva ou negativa entre mobilidade e taxa de acasalamentos. Adicionalmente, observamos que machos e fêmeas parecem se movimentar estrategicamente, preferindo sempre mover-se para plantas próximas (diminuindo o custo energético) e com grande número de indivíduos co-específicos (aumentando a chance de obter um parceiro). Com isso, propusemos o conceito de *movimentação estratégica* na busca por parceiros, que acreditamos ser aplicável a diferentes sistemas de acasalamento.

Finalmente, no último capítulo, investigamos como a seleção sexual e a evolução de caracteres sexualmente selecionados são influenciados pela filtragem de informação gerada por restrições espaciais, ou seja, o fato de que indivíduos estão limitados a escolher seus parceiros sexuais entre aqueles disponíveis em sua vizinhança. Para isso, realizamos simulações computacionais espacialmente explícitas com papéis sexuais convencionais (fêmeas seletivas e machos não seletivos). Observamos que quanto maior o número de machos uma fêmea é capaz de amostrar durante a escolha de parceiros, mais intensa é a seleção sexual. Adicionalmente,

quando permitimos que um ornamento sexualmente selecionado dos machos evoluísse ao longo das gerações, observamos que quanto maior a amostra tomada pelas fêmeas, mais exagerado o ornamento dos machos se torna após 50 gerações. Finalmente, analisamos um grande conjunto de dados empíricos sobre a intensidade de seleção sexual em populações naturais e observamos que a seleção sexual se torna mais intensa quando os indivíduos podem amostrar mais parceiros em potencial, assim como previsto por nossas simulações.

A premissa básica que permeou todos os capítulos foi que um indivíduo só pode interagir com aqueles com os quais ele se encontra — e um indivíduo não se encontra com todos os indivíduos da população! Essa pode parecer uma premissa trivial e até óbvia. Entretanto, ela difere radicalmente da abordagem de campo médio frequentemente usada em ecologia, que pressupõe que cada indivíduo tem a mesma probabilidade de interagir com qualquer outro indivíduo em sua população (Durrett & Levin 1994). Ao quebrar essa premissa, descobrimos que efetivamente o espaço influencia a organização de sistemas de acasalamento, a intensidade da seleção sexual e a evolução de caracteres sexualmente selecionados, tais como ornamentos. Portanto, a principal conclusão que podemos tirar desta tese é que o espaço importa em estudos sobre seleção sexual. Populações não são misturas homogêneas de indivíduos e isso tem consequências tanto a curto quanto a longo prazo. Para alguns ecólogos comportamentais, incluir o espaço em suas análises pode parecer apenas preciosismo estatístico. Como mostramos aqui, porém, é possível gerar hipóteses baseadas na distribuição espacial dos indivíduos de forma que o espaço não é apenas um ruído de fundo incômodo, mas um fator essencial na compreensão de alguns fenômenos relacionados à seleção sexual, tais como a intensidade da competição espermática e a busca por parceiros.

Por fim, gostaríamos de apontar neste último parágrafo alguns dos caminhos que poderíamos seguir para estender e generalizar os resultados que aqui apresentamos. Em primeiro lugar, ainda existem muitas perguntas para responder sobre como a distribuição espacial dos indivíduos e a amostragem de parceiros (*mate sampling*) influenciam sistemas de acasalamento e quais fatores influenciam a amostragem de parceiros. Duas perguntas que me parecem particularmente interessantes são: (1) Quais são os efeitos da filtragem de informação quando o critério de escolha das fêmeas não é consistente *entre* fêmeas como, por exemplo, em espécies

com escolha assortativa de parceiros? (2) Como a filtragem de informação influencia a seleção sexual em uma população em que fêmeas podem usar múltiplos caracteres para acessar a qualidade dos machos? Em segundo lugar, ressaltamos que a probabilidade de encontro entre indivíduos é fortemente influenciada pelo espaço. Ao longo da tese, abordamos apenas interações relacionadas à reprodução, incluindo a escolha e busca de parceiros e a competição entre machos. Entretanto, assim como o modelo probabilístico proposto no Capítulo 2 pode ser aplicado à escolha em contextos não sexuais, a distribuição espacial dos indivíduos deve influenciar outros tipos de interações ecológicas, como polinização e predação. É claro que já existem investigações neste sentido (e.g., Fortuna et al. 2008), mas acreditamos que ainda há muito que explorar neste campo. Finalmente, ao longo de toda a tese, abordamos as consequências da distribuição espacial dos indivíduos sem nos preocuparmos com os mecanismos que geram esta distribuição. Ao propor o conceito de movimentação estratégica na busca por parceiros no Capítulo 3, abrimos a possibilidade de investigar *feedbacks* entre a distribuição espacial e o sistema de acasalamento. Esperamos que as ideias propostas aqui encontrem terreno fértil na mente do leitor e que gerem novos e estimulantes desdobramentos no campo teórico e empírico.

## Referências

- Durrett, R., Levin, S., 1994. The importance of being discrete (and spatial). *Theor. Popul. Biol.* 46, 363–394.
- Fortuna, M. A., García, C., Guimarães Jr., P.R., Bascompte, J., 2008. Spatial mating networks in insect-pollinated plants. *Ecol. Lett.* 11, 490–498.