



Doctoral Thesis

Advances in landscape genetic methods and theory lessons learnt from insects in agricultural landscapes

Author(s):

Strien, Maarten Jan van

Publication Date:

2013

Permanent Link:

<https://doi.org/10.3929/ethz-a-009761395> →

Rights / License:

[In Copyright - Non-Commercial Use Permitted](#) →

This page was generated automatically upon download from the [ETH Zurich Research Collection](#). For more information please consult the [Terms of use](#).

Diss. ETH No. 20832

**ADVANCES IN LANDSCAPE GENETIC
METHODS AND THEORY**
-
**LESSONS LEARNT FROM INSECTS
IN AGRICULTURAL LANDSCAPES**

A dissertation submitted to
ETH ZURICH

for the degree of
Doctor of Sciences

presented by
MAARTEN JAN VAN STRIEN
MSc in Biology, Wageningen University
born 12 December 1981
citizen of The Netherlands

accepted on the recommendation of

Prof. Dr. Rolf Holderegger, examiner
PD. Dr. Janine Bolliger, co-examiner
Prof. Dr. Jaboury Ghazoul, co-examiner
Prof. Dr. Pierre Taberlet, co-examiner

2013

Summary

To counteract the negative effects of habitat loss and fragmentation on the survival of populations and species, knowledge is needed on the effect of landscape elements on migration and gene flow. Such knowledge is gathered in the discipline of landscape genetics. Landscape effects on gene flow are usually assessed by correlating indirect measures of gene flow between population pairs with measures that quantify the intervening landscape between populations. However, there are some methodological and theoretical limitations in current landscape genetic practice that need to be solved to further develop the discipline. I have identified three such limitations. (i) Currently methods based on resistance surfaces or interpopulation transects are being used to quantify the landscape between two populations. While methods based on resistance surfaces usually require subjective decisions and often represent a single potential migration path, straight-line transects have the disadvantage that rectilinear migration routes are assumed. A combination of both methods could overcome these drawbacks. (ii) The statistical analysis of pairwise gene flow and landscape measures is complicated due to the non-independence of pairwise data. Despite some statistical progress on this topic, some issues remain to be solved. (iii) Although gene flow is the result of migration, both processes are bound to different parameters, which should be considered in landscape genetics. Compared to migration, gene flow takes place over larger distances, longer time spans and is more affected by population topology. In this thesis I address the above issues and propose some solutions. As a study system, I choose insect species in fragmented agricultural landscapes.

In chapter 2 of the thesis, I propose a new analytical approach to quantify the landscape between populations (least-cost transect analysis; LCTA), in which traditional methods based on resistance surfaces (i.e. least-cost paths) were used to direct the course of a transect. Compared to straight-line transects, these least-cost transects may better represent migration routes. I aimed to objectively parameterise resistance surfaces by selecting each landscape element in turn as migration habitat. From the transects, I calculated landscape predictor variables that were regressed against measures of indirect gene flow. For multiple regression I used mixed effect models of which the covariance structure was specifically designed to accommodate for the non-independence of the response and predictor variables. Applied to an empirical genetic dataset on the damselfly *Coenagrion mercuriale*, I showed that LCTA was an objective approach that identifies both the most probable migration habitat and the landscape elements that either inhibit or facilitate gene flow. Although I believe this statistical approach to be an improvement in landscape genetics statistical analyses, more stringent testing is needed.

In chapter 3, I study the effect that habitat configuration has on gene flow patterns. With a simulation program, I studied the evolution of isolation-by-distance patterns in different habitat configurations and for different maximum migration distances, by examining relationships between indirect gene flow measures (i.e. F_{ST}) and interpopulation Euclidean distance. I showed that the rate and likelihood of the appearance of certain isolation-by-distance patterns was strongly determined by habitat configuration and maximum migration distance. Furthermore, I determined that the maximum F_{ST} -distance correlation, from a subset of population pairs, was a good estimator of the maximum migration distance. I explain my findings by discussing the effect of population topology on gene flow rates. In appendix 1, we applied these findings to a landscape genetic analysis of the large marsh grasshopper (*Stethophyma grossum*). With LCTA, we were only able to determine a most likely migration habitat and facilitators and inhibitors to gene flow, by restricting our analysis to population pairs within migration range of each other and between which direct gene flow was likely.

In chapter 4, I focus on predicting the effects of landscape change on gene flow between populations of the *S. grossum*. From transects drawn between all population pairs within migration range of each other, I calculated measures of landscape composition and configuration. These measures, together with measures of population topology, were used to predict gene flow. Taking into account the non-independence of response and predictor variables, I modified a leave-one-out-cross-validation approach to select the model with the highest predictive power. With this model, I predicted future gene flow under several hypothetical landscape change scenarios. Some scenarios were predicted to have a positive or negative effect on gene flow, while others had practically no effect. Furthermore, the predictive power of the models was strongly increased by including measures of population topology as predictor variable.

In summary, (i) I propose a new approach to quantify the landscape between population pairs and demonstrate this method on two insect species, (ii) I show how several, formerly unused, statistical methods can be implement or modified for landscape genetic analysis and (iii) I determine from simulated and empirical datasets that the effects of population topology on gene flow should not be overlooked in landscape genetics. Furthermore, this thesis demonstrates how landscape genetic models can be used to plan conservation measures mitigating the negative effects of landscape fragmentation and change.

Zusammenfassung

Um den negativen Effekten von Habitatsverlust und Fragmentierung entgegenzuwirken, sind Kenntnisse über den Einfluss von Landschaftselementen auf die Ausbreitung und den Genfluss von Populationen und Arten notwendig. Mit Hilfe landschaftsgenetischer Methoden können entsprechende Informationen erhalten werden. Meist werden Landschaftseffekte auf den Genfluss durch die Korrelation von indirekt gemessenem Genfluss und Daten über die Landschaftszusammensetzung zwischen Populationspaaren erfasst. Allerdings gibt es in der Landschaftsgenetik einige methodische und konzeptionelle Probleme zu lösen, damit die Disziplin weiterentwickelt werden kann. Drei dieser Probleme habe ich in meiner Arbeit behandelt. (i) Zurzeit werden vor allem Methoden zur Quantifizierung der Landschaft zwischen zwei Populationen angewandt welche entweder auf Resistenz-Oberflächen oder auf Transekten basieren. Resistenz-Oberflächen Methoden basieren meist auf subjektiven Entscheidungen und beziehen sich nur auf einen einzigen Ausbreitungspfad. Lineare Transekte hingegen haben den Nachteil, dass sie nur geradlinige Ausbreitungspfade berücksichtigen. Durch die Kombination beider Ansätze könnten diese Nachteile überwunden werden. (ii) Da paarweise Daten nicht voneinander unabhängig sind, ist die statistische Analyse von Genfluss und Landschaftsdaten zwischen Populationspaaren komplex. Trotz Verbesserungen in der Statistik zur Analyse solcher Daten sind noch immer einige Aspekte ungelöst. (iii) Obwohl Genfluss das Resultat von Ausbreitung ist, sind diese beiden Prozesse von unterschiedlichen Parametern abhängig. Dies sollte in der Landschaftsgenetik berücksichtigt werden. Im Gegensatz zur Ausbreitung findet Genfluss über längere Distanzen und längere Zeiträume statt und ist somit stärker durch die Topologie der Populationen beeinflusst. In dieser Arbeit befasse ich mich mit den drei oben genannten Themen und schlage verschiedene Lösungsansätze vor. Als Untersuchungssystem wählte ich zwei verschiedene Insektenarten, welche in einem fragmentierten Landwirtschaftsgebiet vorkamen.

In Kapitel 2 dieser Arbeit schlage ich einen neuen analytischen Ansatz zur Quantifikation der Landschaft zwischen Populationen vor (least-cost transect analysis; LCTA). Dieser Ansatz verwendet traditionelle Methoden welche auf Resistenz-Oberflächen basieren (d.h. least-cost paths). Mit den Resistenz-Oberflächen wird der Verlauf von Ausbreitungstransekten bestimmt ("least-cost" Transekte), welche im Gegensatz zu geradlinigen Transekten eher die tatsächlichen Ausbreitungsrouten repräsentieren. Da ich die Resistenz-Oberflächen objektiv parametrisieren wollte, verwendete ich jedes Landschaftselement abwechselnd als Ausbreitungshabitat. Die für die Transekte berechneten Landschaftselemente verwendete ich als unabhängige Variablen in Regressionsmodellen. Indirekte Messungen von Genfluss wurden als

abhängige Variable gewählt. Für die multiple Regression verwendete ich “mixed effect models“ für welche die Struktur der Kovarianz unter Berücksichtigung der Abhängigkeit der Variablen berechnet wurde. Mit einem empirischen, genetischen Datensatz der Kleinlibelle *Coenagrion mercuriale* konnte ich zeigen, dass LCTA als objektiver Ansatz zur Ermittlung des wahrscheinlichsten Ausbreitungshabitats sowie der Landschaftselemente, die für die Libelle Genfluss hindern oder fördern, verwendet werden kann. Obwohl der verwendete statistische Ansatz eine Verbesserung bisheriger statistischer Analysen in der Landschaftsgenetik erbringt, sollte die hier vorgeschlagene Methode in Zukunft noch genauer getestet werden.

In Kapitel 3 analysiere ich die Auswirkungen der Habitatzusammensetzung auf Genflussmuster. Mit einem Simulationsprogramm untersuchte ich die Entwicklung von “isolation-by-distance“ Mustern für verschiedene Habitatzusammensetzungen und für verschiedene maximale Ausbreitungsdistanzen einer Art. Dafür berechnete ich jeweils den Zusammenhang zwischen indirektem Genfluss (F_{ST}) und Euklidischer Distanz. Ich konnte zeigen, dass die Rate und die Wahrscheinlichkeit bestimmter “isolation-by-distance“ Muster stark von der Zusammensetzung des Habitats und der maximalen Ausbreitungsdistanz bestimmt waren. Zudem konnte ich zeigen, dass das Maximum der Korrelation zwischen F_{ST} und Euklidischer Distanz (für einen Teil der Populationspaare) eine gute Schätzung der maximalen Ausbreitungsdistanz ergab. Diese Ergebnisse konnte ich durch den Einfluss, welcher die Populationstopologie auf den Genfluss hat, erklären. In Appendix 1 haben wir diese Erkenntnisse in einer landschaftsgenetischen Studie über die Sumpfschrecke (*Stethophyma grossum*) verwendet. Mit LCTA konnten wir das wahrscheinlichste Ausbreitungshabitat sowie jene Landschaftselemente, die den Genfluss von *S. grossum* entweder hindern oder fördern nur dann ermitteln, wenn wir die Analyse auf Populationspaare beschränkten, welche innerhalb der maximalen Ausbreitungsdistanz der Art lagen und zwischen welchen direkter Genfluss wahrscheinlich war.

In Kapitel 4 konzentriere ich mich auf die Voraussage von Effekten von Landschaftsveränderungen auf den Genfluss zwischen Populationen von *S. grossum*. Für die Transekte aller Populationspaare, welche innerhalb der Ausbreitungsdistanz der Art lagen, berechnete ich die Anordnung und Zusammensetzung der Landschaft. Diese Werte wurden zusammen mit Schätzwerten der Populationstopologie für die Voraussage des Genflusses verwendet. Damit ich die fehlende Unabhängigkeit der Variablen berücksichtigen konnte, modifizierte ich einen sogenannten “leave-one-out-cross-validation“ Ansatz um jeweils das Modell mit der besten Voraussagekraft auszuwählen. Mit diesem Modell konnte ich den zukünftigen Genfluss unter verschiedenen hypothetischen Szenarien der Landschaftsveränderung voraussagen. Einige Szenarien zeigten einen positiven oder negativen Effekt der Landschaftsveränderungen auf den Genfluss an, andere hingegen zeigten keinen Effekt.

Ausserdem konnte die Aussagekraft der Modelle stark gesteigert werden, wenn die Populationstopologie als unabhängige Variable in die Modelle einbezogen wurde.

Zusammenfassend (i) schlage ich einen neuen Ansatz zur Quantifikation der Landschaft zwischen Populationspaaren vor und wende diese Methode exemplarisch an zwei Insektenarten an, (ii) demonstriere, ich wie verschiedene, bis anhin nicht benutzte statistische Methoden für landschaftsgenetische Studien modifiziert werden können und (iii) zeige ich mit simulierten und empirischen Datensätzen, dass die Populationstopologie Auswirkungen auf den Genfluss hat und deshalb in der Landschaftsgenetik berücksichtigt werden sollte. Zudem konnte meine Arbeit zeigen, wie landschaftsgenetische Modelle in der Planung von Naturschutzmassnahmen, welche den Effekten der Landschaftsfragmentierung und -veränderung entgegenwirken, verwendet werden können.