



Universiteit
Leiden
The Netherlands

Calculating hazard rates of introgression with branching processes

Ghosh, A.

Citation

Ghosh, A. (2012, May 22). *Calculating hazard rates of introgression with branching processes*. Retrieved from <https://hdl.handle.net/1887/18976>

Version: Not Applicable (or Unknown)

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/18976>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/18976> holds various files of this Leiden University dissertation.

Author: Ghosh, Atiyo

Title: Calculating hazard rates of introgression with branching processes

Date: 2012-05-22

ENGLISH SUMMARY

The majority of the world's most important crops hybridise readily with their wild relatives. With the advent of genetically modified crops, the possible consequences of such hybridisation has come under increasing scientific scrutiny. One important possible consequence is introgression. Introgression is the permanent incorporation of genes from the genome of one population into another. Will genetically modified genes introgress from cultivated plants into their wild relatives? If so, when will this happen? What is a suitable measure for the risk for introgression? To answer these questions requires a combination of experimental and theoretical studies. Experimental approaches can help determine the fitness effects that a crop gene will have when placed in wild individuals, and theoretical approaches can then use this information to forecast the population growth of foreign genes in a wild population. The work in this thesis provides a suitable theoretical framework for quantifying the risk of occurrence of introgression, and companion projects present complementary experimental methods.

Chapter 1 shows that if an invading gene has fitness benefit, then repeated hybridisation will result in the gene invading after some time. This is not to say that the risk of introgression of the gene is high—the gene will eventually invade, but it might take a long time to do so. The risk of introgression occurrence is governed by how long it takes for a permanent lineage to be formed. Since the number of invaders can be initially small, then randomness in the number of offspring (so-called demographic stochasticity) is a crucial factor in determining when permanent lineages form. The theory of branching processes is used to develop a methodology for calculating the time at which permanent invading lineages are formed. From this time distribution, a measure of introgression risk is introduced: the hazard rate. The hazard rate of introgression is the probability that introgression occurs at a certain time given that it has not occurred before. A sample calculation of the hazard rate is shown for plants which can either flower after one year or delay flowering. It is shown that introgression risks can sometimes be higher if plants delay flowering instead of flowering immediately.

Chapter 2 uses the results of chapter 1 and combines them with the preliminary results from a companion study to estimate introgression risk from cultivated carrots into their wild relatives. A sensitivity analysis was performed to determine the most important factors driving introgression. The combination of the experiments and theory in this chapter hints at ongoing work, which uses the carrot (*Daucus carota*) as a model species in developing a comprehensive methodology to estimate introgression risks.

Chapter 3 elaborates on the the results from chapter 1 by allowing for hybridisation rates to change in time, allowing for the incorporation of management strategies such as crop-rotation to be included into risk calculations. The case where it takes several generations before a fitness advantage is seen is also investigated.. Hazard rates of introgression can change in time during crop rotations,

so the task of choosing a suitable level of introgression risk is complicated. Procedures for averaging the hazard rate over time are presented. The average hazard rate could be used as a measure for introgression risk, but it might be misleading since it can significantly underestimate introgression risks during some time periods. Randomly varying hybridisation rates, due to chance changes in pollinator activity or shifts in weather, are also investigated.

Chapter 4 presents a framework to incorporate genetics into the previous methodologies. Furthermore, procedures for calculating hazard rates using computer simulations are shown, in addition to the mathematical methods presented in previous chapters. In order to retain mathematical tractability when using branching processes, the work in previous chapters assumed that introgression was occurring into (infinitely) large wild populations. The use of simulation-based techniques allowed tests to be done on how large a wild population has to be before the assumptions from previous chapters hold. Both branching process and computational approaches give similar predictions for population sizes on the order of 100. For small wild populations, introgression is primarily driven by chance and is less dependent on fitness effects of invading genes. For large wild populations, introgression is driven by selection and branching processes are an efficient tool to calculate hazard rates. Genetic linkage is found to be an important factor affecting introgression risk.

Chapter 5 generalises the approaches from chapters 1 and 3 to allow for random environments. In such scenarios, the hazard rate can change randomly in time. This means that there may be some periods where introgression risk may be higher than in others. This leads to a practical challenge in choosing an acceptable level of introgression risk: should one choose an average hazard rate, a maximum hazard rate or some other level? If an average hazard rate is used, introgression risk might be severely underestimated for some periods of time. Increasing flowering probabilities can either increase or decrease the average hazard rate, depending on the environment.

SAMENVATTING

Het merendeel van de belangrijkste cultuurgewassen in de wereld hybridiseren gemakkelijk met hun wilde verwanten. Met de komst van genetisch gemodificeerde gewassen, is het belangrijk om te onderzoeken wat de mogelijke gevolgen van dergelijke hybridisatie zijn. Een belangrijk mogelijk gevolg is introgressie. Introgressie is de permanente incorporatie van genen van het genoom van een bepaalde populatie in een andere populatie. Zullen genetisch gemodificeerde genen van cultuurgewassen op den duur permanent in hun wilde verwanten voorkomen? Zo ja, wanneer zal dit gebeuren? Wat is een geschikte maat voor het risico op introgressie? Om deze vragen te beantwoorden is een combinatie van experimenteel en theoretisch onderzoek nodig. Experimentele benaderingen zijn nodig om de fitness-effecten te bepalen, die een bepaald gen zal hebben indien het in het genoom van een wilde verwant terecht komt. Deze informatie kan dan vervolgens in theoretisch onderzoek worden gebruikt om het verloop van de frequentie van zo'n gen in een natuurlijke populatie te voorspellen. Het onderzoek dat in dit proefschrift wordt beschreven biedt een geschikt theoretisch kader voor het kwantificeren van het risico op het optreden van introgressie. Dit onderzoek maakt deel uit van een groter programma, met complementaire empirische deelprojecten.

In *Hoofdstuk 1* wordt uitgelegd dat wanneer een gen een fitness voordeel heeft, hoe klein ook, herhaaldelijke hybridisatie uiteindelijk zal resulteren in introgressie. Dit betekent echter niet noodzakelijk dat het risico op introgressie van het gen ook hoog is. Het kan namelijk ook heel lang duren voordat dit gebeurt. Het risico dat introgressie optreedt wordt bepaald door de duur van de periode voordat er een lijn van nakomelingen wordt geproduceerd waarin het gen blijft voorkomen. Dit noemen we een 'permanente lijn'. Aangezien het aantal hybrides aanvankelijk erg klein is, is inter-individuele variatie in overlevingskans en nakomelingen (zogenoemde demografische stochasticiteit) een essentiële factor bij dit proces. De theorie van vertakking processen wordt gebruikt om een methode af te leiden waarmee de verdeling van de tijd tot initiatie van een permanente lijn kan worden bepaald. De hazard rate van introgressie is de kans per tijdseenheid dat er een permanente lijn ontstaat, gegeven dat dat nog niet eerder is gebeurd. In dit hoofdstuk wordt een voorbeeld van een berekening van de hazard rate gegeven, voor planten die eenmalig bloeien, na een of meerdere jaren. Er wordt aangetoond dat het uitstellen van de bloei het risico van introgressie kan verhogen.

Hoofdstuk 2 maakt gebruik van de resultaten van hoofdstuk 1 en combineert ze met de voorlopige resultaten van een empirisch onderzoek naar introgressie van genen van gecultiveerde peen in hun wilde verwanten. Een gevoeligheidsanalyse werd uitgevoerd om de belangrijkste factoren die introgressie beïnvloeden vast te stellen. De combinatie van de experimenten en de theorie in dit hoofdstuk verwijst naar lopend onderzoek, waarin de wortel wordt gebruikt (*Daucus carota*) als modelsoort in de ontwikkeling van een uitgebreide methodologie om introgressie risico's te bepalen.

Hoofdstuk 3 worden de resultaten van hoofdstuk 1 gegeneraliseerd, door toe te staan dat de verdeling van het aantal gevormde hybriden kan veranderen in de tijd. Hierdoor kunnen effecten van gewas management, zoals gewas-rotatie op introgressie risico worden bepaald. Hazard rates van introgressie kunnen hierdoor variëren in de tijd. Dit compliceert het bepalen van een geschikt niveau van introgressie risico. In dit hoofdstuk wordt een procedure om het gemiddelde van de hazard rate te berekenen gepresenteerd. Dit zou kunnen worden gebruikt als een maat voor introgressie risico, maar dat zou misleidend kunnen zijn, aangezien er periodes zijn waarin het risico aanzienlijk hoger is. Het effect van toevalsfluctuaties in hybridisatie-snelheden, bijvoorbeeld ten gevolge van variatie in de activiteit van bestuivers, en/ of weersomstandigheden, worden ook onderzocht. In dit hoofdstuk worden ook situaties onderzocht waarin het enkele generaties duurt voordat er een fitness voordeel is.

Hoofdstuk 4 biedt een framework om meer gecompliceerde genetische mechanismen te beschouwen. Bovendien worden procedures voor de berekening van de hazard rate op grond van computersimulaties gepresenteerd, naast de wiskundige methoden van de voorgaande hoofdstukken. In de vorige hoofdstukken werd ervan uit gegaan dat de wilde populatie groot genoeg was om de kans op interactie tussen hybride individuen te verwaarlozen. Met behulp van simulatie-gebaseerde technieken kan worden gedaan hoe groot een wilde populatie moet zijn voordat deze aanname een redelijke benadering geeft. Het blijkt dat de methode gebaseerd op vertakkingsprocessen en de computersimulaties vergelijkbare uitkomsten geven zodra de populatieomvang van de orde van grootte van 100 individuen of meer is. Voor kleine wilde populaties wordt introgressie in de eerste plaats gedreven door toeval en is het minder afhankelijk van fitness-effecten van de invasie van genen. Voor grote wilde populaties, wordt introgressie gedreven door selectie, en zijn vertakking processen een doeltreffend instrument om risico's te berekenen. Genetic linkage tussen loci van cultuurgewassen blijkt een belangrijke factor invloed op introgressie risico te hebben. Dat betekent dat koppeling tussen een gemodificeerd gen en een gen dat in de natuur nadelig is een bruikbare strategie kan zijn om introgressie risico te verlagen.

Hoofdstuk 5 generaliseert de benaderingen van de hoofdstukken 1 en 3 naar situaties met omgevingen met toevalsfluctuaties. In dergelijke scenario's kan de hazard rate willekeurig veranderen in de tijd. Het blijkt dat er perioden kunnen zijn waarin introgressierisico's behoorlijk hoger zijn dan gemiddeld. Dit leidt tot een probleem bij het kiezen van een aanvaardbaar niveau van introgressie risico: moet men kiezen voor een gemiddelde hazard rate, een maximale hazard rate of een ander niveau? Indien de gemiddelde hazard rate wordt gebruikt, kan introgressie risico ernstig worden onderschat enige tijd. Ook blijkt dat het effect van life history parameters, zoals de kans op bloei, afhangt van de manier waarop de omgeving varieert in de tijd.

ACKNOWLEDGMENTS

Several people were instrumental in bringing this work to its current state. Patsy Haccou's contribution to every chapter should not be understated. A collaboration with Patrick Meirmans brought chapter 4 into fruition, and chapters 3 and 5 were very much a team effort involving São Serra.

Tom de Jong, Wil Tamis, Klaas Vrieling, Jun Rong and Cilia Grebenstein provided helpful guidance throughout my project.

The work presented would not have been possible without funding from the Ecology Regarding Genetically Modified Organisms (ERGO) program, which is managed by the Earth and Life Sciences Council (ALW) of the Netherlands Organisation for Scientific Research (NWO).

CURRICULUM VITAE

Atiyo Ghosh was born on 24 September 1984 in Toronto, Canada. In 2003, after completing his A-Levels at Harrow School, London, he pursued Bachelor's and Master's degrees in Physics at the University of Cambridge. During these degrees, his interests in Biology and stochasticity were piqued by courses in Evolution and Behaviour and Quantum Physics. In 2008, after completing his studies at Cambridge, he moved to the Institutes of Biology and Environmental Sciences at Leiden University to start the work contained within this thesis. He is planning to continue with post-doctoral research on stochastic modelling with biological applications in Japan.