UNIL | Université de Lausanne

Genetics of sex determination in a parasitoid wasp

Master Project
Master in Molecular Life Sciences, Bioinformatics

Cyril Matthey-Doret
Supervised by: Casper Van der Kooi
Directed by: Tanja Schwander

Department of Ecology and Evolution
University of Lausanne - Switzerland

September 1, 2017

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Introduction

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Results

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Figures, tables and legends

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Discussion

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Materials and methods

Crossing experiments

A haploid male coming from an asexual family was crossed with an inbred sexual line. The offspring of asexual females at the Nth generation were used.

RAD-seq protocol

N samples of *Lysiphlebus fabarum* coming from X different asexual females kept in ethanol at -XX°C for XX months. They were sexed visually and prepared in 6 separate libraries, all following the same protocol from XXX. ddRAD-seq was performed using EcorI and MseI restriction enzymes and a fragments of 200-450 bp were size selected on agarose gel. Single-end sequencing was performed using Illumina Miseq (or Hiseq 2500???). Samples were multiplexed in each library following the TruSeq multiplexing design and libraries were pooled pairwise on the same Illumina lane using different adapters (iA06 or iA12).

STACKS pipeline

The raw reads were trimmed and demultiplexed using the "process radtags" module from the STACKS suite and two mismatches were allowed to detect adapters. The 96bp demultiplexed reads were mapped to the latest assembly of the *L. fabarum* genome using BWA aln with 4 mismatches allowed. Stacks were generated from SAM files using the pstacks module with a minimum stack depth (-m) of 3. The catalogue of loci was built with cstacks allowing for a distance (-n) of 3 mismatches between samples at each locus. Individuals with less than 10% radtags compared to the average across all samples. Populations was run pooling all samples together, requiring each locus to be present (-r) in at least 80% of samples. The different parameters were selected following guidelines in **Paris et al. 2017**.

Ploidy separation

Genome wide heteroyzgosity per individual was computed on all variant sites using the output VCF file from a separate populations run. Only high confidence loci were included by requiring a minimum sequencing depth (-m) of 20 reads in populations. A conservative threshold of 77% heterozygous sites was determined empirically and individuals above that threshold were considered haploid and excluded.

Categorizing families

The proportion of males among diploid offspring was used to group families by number of heterozygous CSD loci in the mother. The total number of diploid males was inferred from sequenced individuals by extrapolating the rate of haploidy from sequenced individuals to the total number of males. The proportion of males among diploid offspring was then computed using the inferred number of diploid males. The families were classified using 1-dimensional k-means clustering on the proportion of males among diploid offspring and 2 scenarios were considered to decide the number of categories; either 2 CSD loci, resulting in 3 categories (k=3), or 3 CSD loci resulting in 7 categories (k=7).

Association mapping

Case-control association mapping was carried separately in each category of family. The number of heterozygous males, heterozygous females, homozygous males and homozygous females was computed for every SNP and a two-sided Fisher exact-test was performed on the 2X2 contingency tables. P-values were corrected for multiple-testing using Benjamini-Hochberg correction.

Supplementary Material

Code and manuscript data hosted at: [INSERT URLS]