Project title:	Resistance and susceptibility in interactions between apple and woolly aphid
Project number:	
Project leader:	Cindayniah Godfrey (Harper Adams University/NIAB EMR)
Report:	Annual report, October 2021
Previous report:	Annual report, October 2020
Key staff:	Simon Segar (Harper Adams University), Michelle Fountain (NIAB EMR), Felicidad Fernández Fernández (NIAB EMR),
Location of project:	NIAB EMR
Industry Representative:	Loraine Boddington, National Association of Cider Makers, NACM Cider Makers Ltd., The Bounds, Much Marcle, Ledbury, Herefordshire, HR8 2NQ Charnee Butcher, Worldwide Fruit, Worldwide Fruit Ltd., Apple Way, Wardentree Lane, Spalding, Lincolnshire, PE11 3BB
Date project commenced:	23rd September 2019

DISCLAIMER

While the Agriculture and Horticulture Development Board seeks to ensure that the information contained within this document is accurate at the time of printing, no warranty is given in respect thereof and, to the maximum extent permitted by law the Agriculture and Horticulture Development Board accepts no liability for loss, damage or injury howsoever caused (including that caused by negligence) or suffered directly or indirectly in relation to information and opinions contained in or omitted from this document.

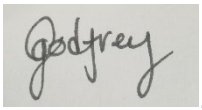
AUTHENTICATION

We declare that this work was done under our supervision according to the procedures described herein and that the report represents a true and accurate record of the results obtained.

Cindayniah Godfrey

PhD Student

NIAB EMR/Harper Adams University

Signature .....  ..... Date 19..11.2021.....

[Name]

[Position]

[Organisation]

Signature ..... Date .....

Report authorised by:

[Name]

[Position]

[Organisation]

Signature ..... Date .....

[Name]

[Position]

[Organisation]

Signature ..... Date .....

CONTENTS

To update the Contents Page, click to highlight the table, then right click on 'Update field', 'Update page numbers only'

GROWER SUMMARY

Headline

Woolly apple aphids in the UK are genetically different from in other countries and there is genetic variation within UK populations sampled. This is likely through sexual reproduction which was previously thought to not occur in the UK and may lead to spreading virulence genes such as for overcoming resistant rootstocks.

Background

Woolly apple aphid (WAA) is a sap-feeding pest which originates from North America but has been in the UK since the 18th century. In America WAA shows both sexual and asexual reproduction but has broadly lost the sexual stage in the rest of the world, leading to reduced genetic diversity.

There have been reports of sexual forms of WAA in a number of countries around the world, including New Zealand, India, and Australia, all in major apple growing regions. A number of commercially available rootstocks contain WAA resistance genes including MM106, M.116, and some individuals from the Geneva rootstock line. There are some isolated reports of WAA feeding on some of these resistant rootstocks, mostly in the southern hemisphere however it is not yet certain how prolific these resistance-breaking WAA are.

Sexual reproduction leads to genetic flow between individuals and between isolated populations which can allow for faster spread of traits undesirable for growers, such as potential genes for insecticide resistance, although this has not yet been documented. Some WAA populations have been found in other countries which can feed on rootstocks which were bred to be resistant to WAA. It is not known what gene(s), if any, are responsible for this resistance-breaking ability but a sexually reproducing WAA population will be more likely to be able to spread this type of virulence gene between individuals.

It is thought that WAA in the UK are exclusively asexual and therefore there would be little genetic flow between populations. Woolly apple aphid are not especially mobile so it is unlikely that separate populations will be able to mix without sexual reproduction. Understanding the genetic diversity of WAA in the UK and around the world will help to predict how WAA populations are mixing and to speculate about how virulence genes may be able to spread.

This project aims to both better understand the pest species to aid their control and to develop potential novel control methods. Breeding plant material that is resistant to aphid feeding is an effective control method for integrated pest management. A number of resistant rootstocks

are currently available but there have been some reports in other countries of aphids feeding on these rootstocks. Expanding the potential breeding programme for WAA resistance to include resistances from different sources will give a greater range of resistance genes to choose from. Such a resource reduces the likelihood that resistance-breaking aphids will be able to break every resistance. There is also the possibility to combine multiple resistance genes to give longer lasting resistance to WAA.

Summary

The main work package undertaken in the second year of this project was to analyse the genetic diversity of woolly apple aphid (WAA) samples collected from nineteen locations, mostly from around the south east of England. Samples were included from North America where the pest originates, and New Zealand where WAA is a serious pest, sexual WAA have been found, and populations has been observed feeding on rootstock material containing resistance genes.

Financial Benefits

The UK apple industry is worth approximately £190M and makes up a significant proportion of the 605000 tonnes of top fruit sold in the UK annually (https://agri-epicentre.com/impact/extending-the-availability-and-flavour-of-uk-apples/). A 2011 report found that without pesticides the UK apple industry would suffer serious yield losses and become commercially unviable. The costs of plant protection services have increased in the fresh fruit sector since 2015 and with likely further pesticide restrictions in the future it is important to consider alternate pest control options. Resistant rootstocks offer a one-time investment at the beginning of an orchard's life to control WAA underground for that rootstock's use.

Action Points

No action points for growers at present.

SCIENCE SECTION

Introduction

Woolly apple aphid (WAA; *Eriosoma lanigerum* Hausmann.) originates from North America and has now spread across the world. In North America WAA show a sexual lifecycle, alternating between apple and American Elm (*Ulmus americana* L.) but in the rest of the world WAA largely shows an asexual lifecycle feeding exclusively on apple. When reproduction is exclusively asexual it can lead to reduced genetic diversity within the population because individuals are not exchanging genes during reproduction.

Based on this we assume that within North America WAA is sexually reproducing and therefore genetically diverse but that in the rest of the world reproduction is solely asexual and therefore expected to have little genetic diversity. If WAA are sexually reproducing in the UK, then it may enable the spread of undesirable genes. Several other aphid species have developed resistance to pesticides, including to the active ingredient of Batavia (spirotetramat), which is recommended for WAA control. This presents another potential area in which spreading of a particular gene could have broader consequences. Rootstocks which were bred to carry genes for resisting aphid feeding are commercially available, for example M.116 and MM106. These prevent aphids feeding below-ground where it can be hard to detect them and there is some evidence that this resistance may be able to travel to a grafted interstock or scion. There have been reports of WAA feeding on some of these rootstocks. We assume that this ability is caused by resistance-breaking genes in the aphid, but the exact gene or genes are unknown.

This work set out to answer three key research questions: Does a population of WAA exist in the UK which is genetically distinct from populations in other countries; To what extent is there genetic variation within this UK population? Or is it possible that there are several distinct populations? In the instance of genetic variation, how likely is it that this variation is caused by sexual reproduction?

Materials and methods

Sample collection

Aphids were collected from sites detailed in Table 1 by brushing aphids from plant material with a paintbrush into tubes filled with silica gel for drying. Each sample was collected from a single colony, assuming that one colony is genetically identical and come from on single asexual mother. Eight samples of the elm balloon-gall aphid (*Eriosoma lanuginosum* Hartig) were included to act as an outgroup.

Table 1- Population identifiers and their location of collection

Population	Collection site	No. samples
1	Elm balloon-gall aphid (*E. lanuginosum*)	10
2	New Zealand Plant & Food, NZ (-39.65363452085866, 176.8760303275441)	3
3	Geneva, New York State, USA (42.902758, -77.029434)	6
4	Sodus, New York State, USA (43.210338, -77.015819)	7
5	Glasshouse -NIAB EMR, Kent, UK (51.284623, 0.449558)	10
6	Gene bank - NIAB EMR, Kent, UK (51.287592, 0.441731)	6
7	WAA culture - NIAB EMR, Kent, UK (51.285916, 0.453081)	19
10	National Fruit Collection, Kent, UK +0C tunnel (51.296155, 0.882794)	9
11	National Fruit Collection, Kent, UK +2C tunnel (51.296231, 0.881980) - 2020	7
12	National Fruit Collection, Kent, UK +2C tunnel (51.296231, 0.881980) - 2021	29
13	National Fruit Collection, Kent, UK +4C tunnel (51.296224, 0.882149) - 2020	5
14	National Fruit Collection, Kent, UK +4C tunnel (51.296224, 0.882149) - 2021	31
15	Walton-on-Thames, Surrey, UK (51.386154, -0.431309)	1
16	West Malling, Kent, UK	1
17	Honoton Farm, Paddock Wood, Kent, UK (51.146229, 0.412999)	2
18	Clockhouse Farm, Kent, UK (51.227388, 0.498149)	1
19	Whitstable, Kent, UK (51.357181, 1.018644)	2

DNA extraction

DNA extracted from multiple individuals using the Quiagen DNeasy Blood and Tissue kit. DNA extraction products were diluted using distilled water and then amplified by Polymerase Chain Reaction (PCR) using microsatellite markers from Lavandero *et al.*, 2009b. PCR products were analysed using ABI PRISM® DNA Sequencing Analysis Software and the resulting peaks were classified using GeneScan® and Genotyper® Analysis Softwares (Applied Biosystems Inc).

Data analysis

Population structure was inferred using the software STRUCTURE version 2.3 (Pritchard *et al*., 2000; Hubisz *et al*., 2009) assuming a range of populations numbers from 1-17 based on the number of locations samples were collected from. The data output from this software was further analysed to determine which potential number of populations is the most likely, ΔK. The ΔK values were analysed by an analysis of variance (ANOVA) using R Studio version 1.2.5019 (RStudio Team, 2019). Principal Component Analysis (PCA) was calculated using R Studio version 1.2.5019 (RStudio Team, 2019). A number of other statistical analyses were calculated using the Excel plugin GenAlEx: Hardy Weinberg equilibrium values, identification of private alleles and population assignment likelihoods were identified using GenAlEx version 6.5 (Peakall & Smouse, 2006, 2012). To enable these analyses the dataset was

further reduce, removing populations with a sample size less than five (sampling locations 2 and 15-19), and marker loci with there was ≥50% missing data for at least nine of the seventeen sampling locations.

Results

STRUCTURE analysis

STRUCTURE HARVESTER found the most likely number of populations (K) from those tested, to be thirteen (Figure 1). No significant difference was found in values of ΔK between the tested values of K (p=0.454). Figure 2 shows the genetic differences of thirteen different population, as suggested by the software STRUCTURE HARVESTER (Figure 1). The elm balloon-gall aphid and the New Zealand WAA samples form distinct genetic clusters which do not appear at any other sampling locations. There are other sampling locations which show only one genetic cluster (e.g. 13, 11, 16,17) but in these instances the genotypes of those clusters are found within other populations. Both sampling locations in the USA (populations three and four) show the same two genetic clusters. The other samples collected from the UK suggest there may be as many as nine genetic clusters within the UK, some of which are present in many samples, and some of which are only present in a small number of samples. This suggests both that there are genetically distinct populations of WAA outside

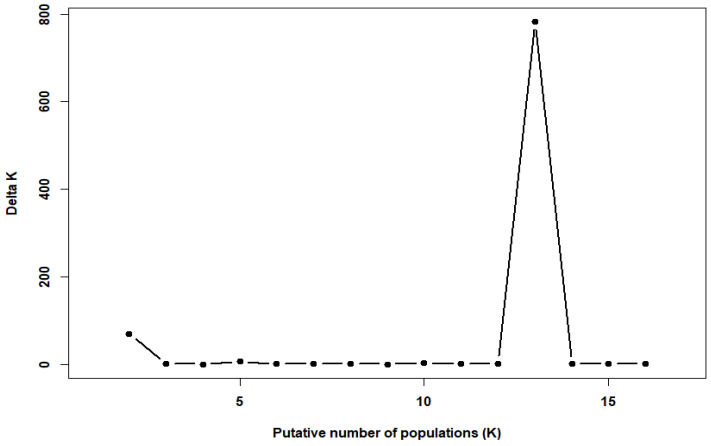


Figure 2- ΔK values for tested numbers of populations from 2-16 as generated by STRUCTURE HARVESTER

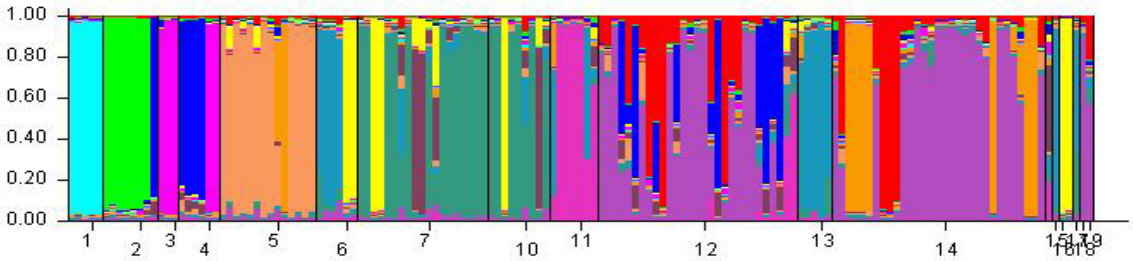


Figure 1- bar plot output of STRUCTURE software. Each vertical line represents an individual sample. Sampling locations are indicated along the x axis (see Table 1) and each colour represents the value of K tested, in this case 13.

of the USA and that there is genetic variability within UK populations. This suggested population clustering is not reflected by the PCA (Figure 3) which appears to cluster samples into approximately three groups with no clear clustering of samples from the same collection

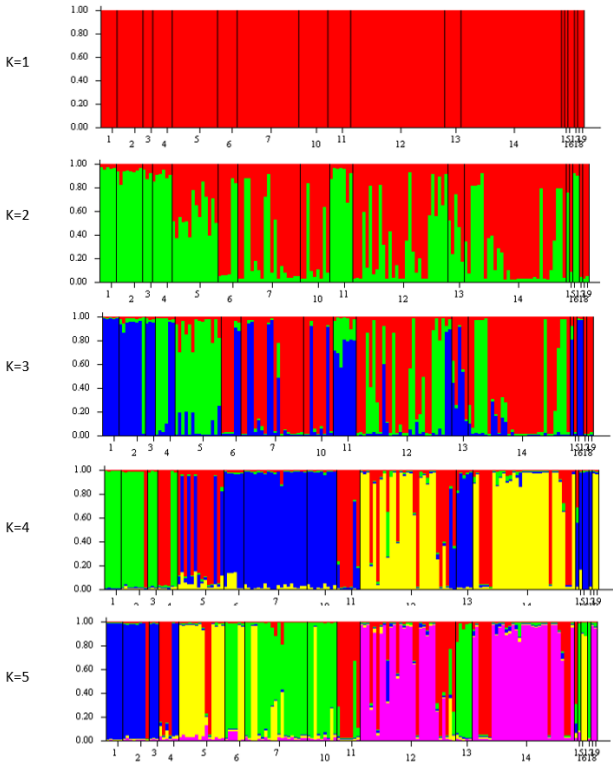


Figure 3- boxplot outputs from STRUCTURE for putative K values of 1-5. Each individual sample is represented by a single vertical line and original sampling populations are indicated by vertical black lines.

population together. This may however reflect the uneven sample sizes across sampling sites leading to differential occurrence of genetic clusters within different sampling locations.

The results of the ΔK analysis suggest that within the samples there are thirteen distinct genetic populations present within the seventeen locations sampled, suggesting that there is genetic variation occurring in the UK, against what has been believed. In this instance only eight genetic markers were used and, although the size of the populations from which samples were taken cannot be known, the number of samples analysed varied (Table 1). This may mean that the estimated likely thirteen genetic clusters is not accurate in this instance. Increasing the number of samples or the marker coverage in the future may help to refine the ΔK analysis.

Given this variation among UK WAA samples it is reasonable to ask why this is the case. There are several potential explanations: spontaneous mutation or random genetic drift in

obligate asexual populations; multiple points of invasion of WAA from the USA where it does sexually reproduce; and lastly that there is WAA sexual reproduction in the UK.

Population assignment in GenAlEx

The allele frequencies for each population were calculated at each marker locus for all sampling populations using GenAlEx and then assigned to the most likely of these populations, including the original sampling population, assuming random mating (Paetkau *et al.,* 2004). Just over half of the samples remain assigned to their original population with the remainder assigned to other populations (Table 3. Before the removal of populations with small sample size and loci with a high proportion of missing data 82% of samples were assigned to their original population and 18% to other populations (data not presented).

Table 2- population assignment outcomes from GenAlEx showing allocation of individuals to either their original sampling population or to another sampling location.

Sampling location	Assigned to original population	Assigned to another population
1	10	
3	6	
4	1	6
5	4	6
6	1	5
7	9	10
10	3	6
11	5	2
12	9	20
13	4	1
14	25	6
Total	77	62
Percent	55%	45%

Private allele summaries

Private allele summaries were calculated using GenAlEx for the reduced data set used. Twenty-two private alleles were identified across four of the five loci tested and in seven of the remaining eleven populations (Table 5). The populations with private alleles were *E. lanuginosum*, both USA populations, one population from NIAB EMR, and two populations from the National Fruit Collection.

Table 3- Summary of private alleles identified for each population of a reduced dataset calculated with GenAlEx. The locus and allele of the private allele, and the frequency at which it occurs, are given.

Population	Locus	Allele	Freq
1	*Erio20*	146	0.143
1	*Erio20*	158	0.143
1	*Erio20*	186	0.071
1	*Erio20*	187	0.286
1	*Erio33*	142	0.500
1	*Erio33*	222	0.167
1	*Erio33*	223	0.333
1	*Erio72*	221	0.125
3	*Erio72*	172	0.300
3	*Erio75*	152	0.250
4	*Erio20*	166	0.143
4	*Erio75*	145	0.125
7	*Erio33*	169	0.033
7	*Erio75*	163	0.059
12	*Erio20*	161	0.020
12	*Erio20*	170	0.020
12	*Erio20*	172	0.020
13	*Erio75*	155	0.167
14	*Erio20*	162	0.017
14	*Erio20*	171	0.017
14	*Erio20*	178	0.133
14	*Erio75*	137	0.021

Private alleles are those found only in one (sub-)population (Neel, 1973) and can be indicative of heritable alleles which are restricted to, for example a single family unit. Eight of the twenty-two private alleles identified, almost one third, are present in the samples of *E. lanuginosum* which is to be expected as a separate species from WAA. Each of the samples collected from the USA (populations three and four) show two private alleles apiece, as did population seven collected at NIAB EMR. This suggests that these populations have some degree of isolation

from other populations. Whilst this result was expected for the American samples it is surprising that only one population of WAA collected at NIAB EMR showed the presence of private alleles. This collection site was observed to have some WAA feeding on supposedly resistant rootstocks which may be related to the presence of genetic isolation as indicated by the presence of private alleles. Another eight of the private alleles were found in samples collected from three locations at the National Fruit Collection (populations twelve, thirteen, and fourteen). The largest sample sizes of the data were collected at locations twelve and fourteen and it has been found that increasing the number of individuals per population increases the number of private alleles identified which in turn suggests higher levels of gene flow (Slatkin, 1985). This suggests that collecting more, larger data sets of WAA genomic material may reveal further genetic variation between populations than previously thought.

Principal Component Analysis

Principal Component Analysis (PCA) clustered most samples closely together with no apparent clustering of individuals from different sampling populations (Figure 3). There does however appear to be clustering of individuals from multiple sampling populations into circa three or four groups, with some individuals not assigned to clusters. Both alleles of Erio33 and an allele each of Erio72 and Erio75 show trends diverging from the central.

The allocation of individuals to their own or another population by GenAlEx assignment analysis found that approximately half of individual samples were assigned to a population other than their own.

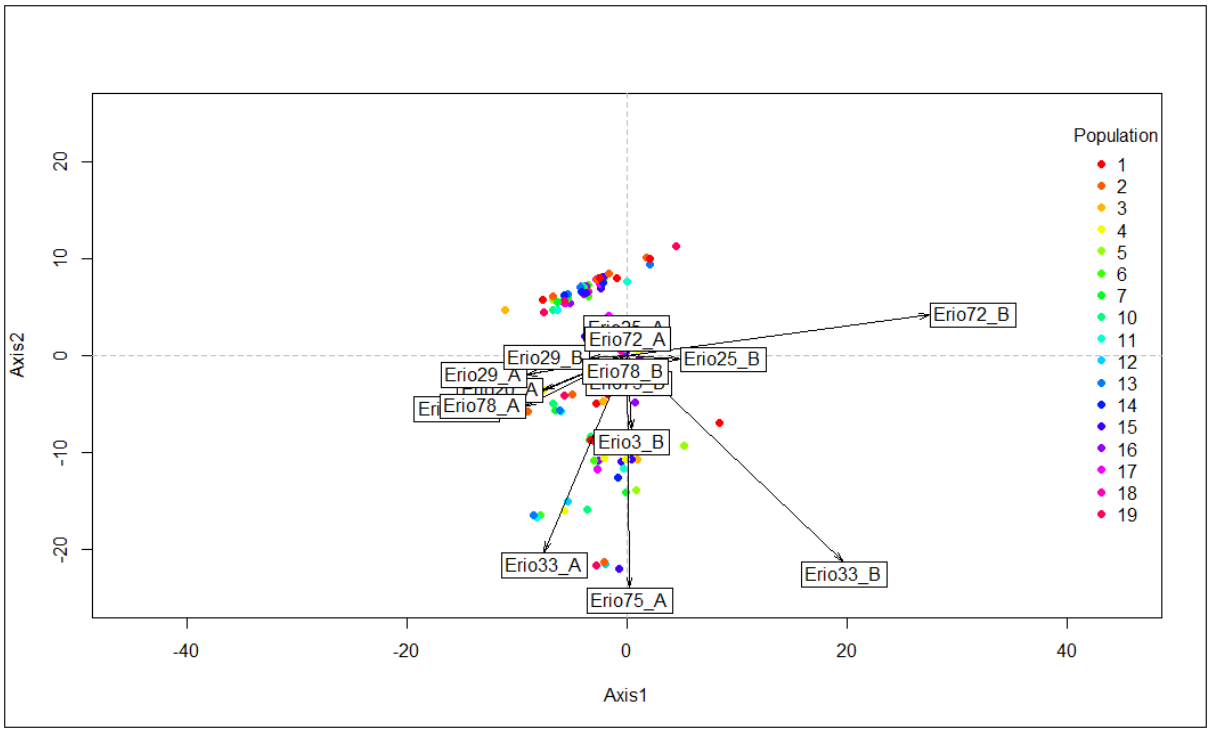


Figure 4- results of a Principal Component Analysis (PCA) output showing clustering of WAA samples based on collection locations (see Table 1). Projected patterns of data points for each microsatellite locus are indicated with arrows. For each microsatellite marker two alleles are given, indicated A and B.

Discussion

GenAlEx requires populations to be pre-defined before analysis and can only compare samples to these defined populations. Assignment of individuals to their original sampling population implies they were born to this population whereas assignment to another population implies immigration to the destination population (Paetkau *et al*., 2004). Only two populations showed true classification to their original sampling population: *E. lanuginosum* and one of the two American samples, suggesting that those genotypes are not shared within other populations. The relatively even split between original and other organisms suggests both clonal reproduction and migration of WAA samples between populations. It is not clear whether this migration is between UK populations or whether they may be separate invasions of WAA from the USA without more extensive sampling from the USA to determine American genotypes to compare against UK samples.

Organisms which reproduce both sexually and asexually can exist in populations which are wholly sexual or asexual (cyclical parthenogenesis and obligate parthenogenesis respectively), or partially sexual populations (Delmotte *et al.,* 2002). It is possible that high heterozygosity and low allelic polymorphisms in sexual populations of *R. padi* is a result of

either long-term asexuality leading to high genetic diversity (as described by Bengtsson, 2003), or as a result of multiple origins of asexual lineage which have since hybridised (Delmotte *et al.,* 2002). It is possible that distinct genetic populations are the result of multiple invasions of WAA from the USA at different time points. Assuming sexual reproduction had been occurring in the USA continuously in the meantime, each invasion to the UK which then became functionally asexual would have a different genotype than the previous invasion. This may explain genetic variation between different locations within the UK but is more tenuous when considering that some sampling locations of this study are enclosed in a small area and still showed allelic polymorphisms, where it would be unlikely that multiple invasions had occurred.

Conclusions

Woolly apple aphid collected in England is genetically different from populations collected in New Zealand and the USA. Variation was also found between WAA samples from within the UK, suggesting a much greater range of genetic difference than was expected based on previous assumptions that WAA only reproduces asexually in the UK. This leads the conclusion that WAA is likely sexually reproducing in the UK.

Knowledge and Technology Transfer

Glossary

References

Bengtsson, B.O. (2003). Genetic variation in organisms with sexual and asexual reproduction. Journal of Evolutionary Biology 16:189–199.

Delmotte, F., Leterme, N., Gauthier, J.P., Rispe, C., Simon, J.C. (2002). Genetic architecture of sexual and asexual populations of the aphid *Rhopalosiphum padi* based on allozyme and microsatellite markers. Molecular Ecology 11:711–723.

Hubisz, M.J., Falush, D., Stephens, M., Pritchard, J.K. (2009). Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources 9:1322–1332.

Neel, J.V. (1973). “Private” genetic variants and the frequency of mutation among South American Indians. PNAS 70:3311–3315.

Paetkau, D., Slade, R., Burden, M., Estoup, A. (2004). Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. Molecular Ecology 13:55–65.

Peakall, R., Smouse, P.E. (2006). genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6:288–295.

Peakall, R., Smouse, P.E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics 28:2537–2539.

Pritchard, J.K., Stephens, M., Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Slatkin, M., (1985). Rare alleles as indicators of gene flow. Evolution 39:53–65.

Appendices

