Inferring Male Population Structure of *Myotis lucifugus*

using Y-chromosome Markers

By
Benjamin Michael Perriman

A Thesis Submitted to
Saint Mary's University, Halifax, Nova Scotia
in Partial Fulfillment of the Requirements for
the Degree of Bachelor of Science (Honours Biology).

[April, 2014], Halifax, Nova Scotia

Approved: Dr. Hugh Broders
Internal Supervisor

Approved: Dr. Timothy Frasier
Internal Supervisor

Approved: Dr. Jeremy Lundholm
Reader

Date: April, 2014

Inferring Male Population Structure of *Myotis lucifugus*

using Y Chromosome Markers

by Benjamin Michael Perriman

Abstract

The little brown bat (*Myotis lucifugus)* is a vagile species occupying a large distributional throughout most of North America. *Myotis lucifugus* separate into female maternity colonies to rear their young, and males appear to be solitary during the summer. In the fall males and females congregate for mating at hibernacula, after which they hibernate for the winter. In many mammalian species males are the primary dispersers causing gene flow between populations. Y chromosome markers can be used to characterize male specific population dynamics. White-nose syndrome (WNS) is a fungal infection that disturbs hibernating bats, and is distributed through physical contact between bats. Y chromosome sequences that show high conservation allow for cross species amplification. The 7th intron of the DBY gene and the 1st intron of the SMCY gene show conserved primer binding sites allowing for amplification in species without a characterized Y chromosome. SMCY1 locus was split into the SMCY1a and SMCY1b loci due to its length for better sequencing product. After optimization of the sequencing reaction 111 individuals were sequenced at these three loci. Aligned sequences showed eight haplotypes and nine variable sites across both the DBY7 and SMCY1b loci. An analysis of molecular variance (AMOVA) for the DBY7 and SMCY1b loci showed no statistically significant level of genetic structuring of Y chromosome haplotypes. A low transition transversion ratio was found for both loci, which suggests mutation bias. Low variability in intron sequences can be combined with variable microsatellites to better characterize male population structure.

April, 2014

Introduction

Population structure is the subdivision of populations where two individuals are less likely to mate due to clustering of individuals. Mountains, large bodies of water, and human development may represent barriers for dispersal. These factors decrease the likelihood that individuals from different populations will reproduce (Wright, 1943). Subpopulations will also form from the formation of distinct colonies where distance separates them. The degree of isolation depends on the physical distance, genetic distance (Nei, 1972), and the geographic factors separating the subpopulations.

A single population may split into multiple subpopulations through dispersal to new suitable habitats. Once a population splits into multiple subpopulations genetic drift may cause their allele frequencies to change due to chance. With each population undergoing genetic drift independently the chance that two population's allele frequencies remain the same is low. Mutations may also alter ancestral population allele frequencies in each subpopulation independently. These mutations will be unique to each subpopulation and genetic drift will alter the new mutated alleles frequencies in the subpopulation.

With minor separation of subpopulations some individuals will move between groups causing gene flow (Kokko and Lopez-Sepulcre 2006). Gene flow is the transfer of alleles between groups through the migration of individuals between populations, and works contrary to speciation by reducing the genetic distance. Gene flow distributes unique alleles to other populations and changes allele frequencies towards Hardy-Weinberg equilibrium.

With subdivision of a population into subpopulations causes a reduced level of heterozygotes. This is known as the Walhund effect. Using F-statistics the reduction in heterozygosity can be calculated (Weir & Cockerham, 1984). The reduction in heterozygosity due to population structuring is known as F_{st}. Φ_{st} is a modification of F_{st} that accounts for genetic distance between haplotypes (Excoffier, Smouse, and Quattro, 1992). Φ_{st} can be calculated from analysis of molecular variance (AMOVA). Based on F_{st} and Φ_{st} the degree of structuring in a population can be determined. A limitation of F_{st} is it is requires subpopulations to be in equilibrium.

Population structuring can be determined with nuclear DNA and mtDNA, but the information provided from each type of DNA is different. Autosomal nDNA provides many loci given the greater overall length of the region and can provide individual level of resolution. In population dynamics autosomal nDNA can provide whole population movement dynamics, but sex specific movement can be missed. In many species females display site fidelity and males act as genetic dispersers (Agnell, R, et al, 2013; Greenwood, 1980). With autosomal nDNA this would appear as weak population structuring. Lineage markers like mtDNA and the Y chromosome can provide each sex's movement, but with lower levels of resolution. To determine female only population structure mtDNA is used because it is only passed from mother to daughter. Even though mtDNA is present in all individuals in a population males are genetic dead ends for mtDNA. Y chromosomal DNA can be used to determine male only population dynamics.

The inherent limitation of using the Y chromosome is that it is restricted to only male individuals, and if a male has no male offspring his Y chromosome lineage ends. In

a 50:50 sex split of a population all individuals contain two copies of autosomal DNA, one to two copies of the X-Chromosome, one copy of mtDNA, and half the population contains a single Y chromosome. With a strict father to son inheritance the Y chromosome can be used to identify close relatives (Petit, E., et al, 2012). The Y chromosome is a powerful tool in mapping paternal lineage for an individual, and can be used to trace back the evolution of the current Y chromosome haplotypes.

The Y chromosome partially recombines with the X chromosome, the other sex chromosome, in the pseudo-autosomal region with the rest of the Y chromosome being inherited whole, and unmodified (Butler, J, 2012). This means that the Y chromosome of the son is almost identical to that of the father. Without recombination the main source of new alleles or haplotypes is from mutation (Igea, J., et al, 2010). Since the Y chromosome is nuclear DNA, the Y chromosome undergoes proofreading after transcription, and therefore the mutational rate is less than mitochondrial DNA, which lacks proofreading machinery (Butler, J, 2005). A decreased mutation rate limits the power of discrimination of the Y chromosome since there are fewer variable sites total than autosomal DNA. For autosomal DNA individual level resolution comes from independent segregation of loci on different chromosomes, and for loci far apart on the same chromosome. This allows for the use of the product rule when combining loci.

Each set of loci must be treated as a haplotype since it is inherited as a single unit instead of each locus being inherited nearly independently, or completely independently depending on chromosome position. The Y chromosome has been used to map the out of Africa theory of human evolution (Mendez, F, et al, 2013), and has shown a disparity

between mitochondrial DNA variability and Y chromosome variability in wolves (Iacolina, L, 2010). The extensive use of microsatellites has led to a large DNA database for forensic use, and the development of Y chromosome microsatellite markers to be used for identification when autosomal microsatellites are too degraded or the sample is contaminated. When using the Y chromosome to test the 'out of Africa' theory they calculated the genetic distance of many individuals from many genetic backgrounds at the common human Y chromosome microsatellites. This suggested a large period of time where the Y chromosome changed, but the location didn't change. This showed a that many mutational events occurred inside Africa and then branching out into Europe, and Asia confirming the out of Africa theory (Mendez, F, et al, 2013). The Thomas Jefferson case questioned if there were any illegitimate children of Thomas Jefferson with an African-American woman. Since they lacked any DNA sample from Thomas Jefferson they found a male descendent and used his Y chromosome and compared it to the supposed illegitimate child's Y chromosome. This test showed that they were related, but it they could not identify if he, or a sibling was the father (Butler, J, 2005). This high profile case demonstrates both the power and weakness of Y chromosome markers in their ability to show paternity across many generations, and lack of discrimination between close individuals. Their use in ecology can help give broad senses of movement from the distribution of haplotypes.

Extensive work with the wolf Y chromosome has been performed to see if results concur with conclusions based on mitochondrial DNA's lack of variability (Iacolina, L, 2010). After testing many individuals they found a four haplotypes of the Y chromosome

compared to the single mitochondrial haplotype allowing them to infer that a greater bottleneck occurred for females than males with the use of these lineage markers.

There are intersexual differences in the behaviour of bats. Females require large easily accessible food supplies for feeding young, and while pregnant (Papadatou, E, et al, 2009). In some species male bats avoid competition with the females to increase offspring longevity, and are able to thrive with difficult feeding environments (Papadatou, E, et al, 2009). In some populations males disperse during swarming to distant swarming sites, and females remain relatively stationary with some movement (Senior, P, et al, 2005). This form of migration ensures that gene flow occurs and keeps the females in a safe place to rear the young. These differences in population structure and movement can be determined with Y chromosome lineage markers to determine paternal lines.

White-nose syndrome (WNS) is a fungal infection in hibernating bats that invades skin tissue causing infected bats to arouse from hibernating more often (Marshall 1998; Gianni et al. 2003). The fungus implicated with WNS is *Pseudogymnoascus destructans*, formerly *Geomyces destructans* (Lindner, D, 2013). *P. destructans* is transmitted by direct physical contact between individuals (Brownlee-Bouboulis & Reeder., 2013), and grows at temperatures of 3-5°C. These conditions are similar to bat hibernaculum, allowing the fungus to thrive in the close quarters of a hibernacula. Frequently waking up during hibernation from irritation from the fungus invading the skin tissue (Cryan et al., 2010), and lack of resources for active bats means that the infected bats will deplete energy reserves at a faster rate. During hibernation there are no resources to replenish these fats stores leaving the bat to die from exhaustion, or starvation. *P. destructans* causes a high

mortality rate of over 75% (Blehert, Hicks, Behr, et al, 2009). Transmission through physical contact means that population level movement dynamics is important to understand to predict the movement of this pathogen.

Myotis lucifugus, commonly known as the little brown bat, is a small, highly vagile hibernating bat distributed throughout North America (Broders, et al. 2003, 2006, Norquay, Martinez-Nunez & Dubois, 2013). During the winter season *M. lucifugus* hibernates in caves and abandoned mines. After emerging from hibernation the female bats move to maternity colonies. In these maternity colonies female bats raise their young. During this time male bats form loose summering sites. In late summer to early fall the bats form swarming colonies where mating occurs. During mating females will mate with many males. During swarming it is suspected that gene flow occurs here with males migrating between colonies, and females showing philopatry.

Current genetic work with *M. lucifugus* focuses on mtDNA and autosomal markers. Autosomal microsatellites and mtDNA sequences are being combined for a better understanding of the bat's population structure. The full genome of a female *M. lucifugus* has been sequenced at 7x coverage giving access to more autosomal DNA markers if needed (Lindblad-Toh, K, et al, 2010). Y chromosome work on close relatives of *M. lucifugus* has focused on the identification of cryptic species, not on male population structuring (Claire, E, 2011; Lim, B., Engstrom, M., et al, 2008).

When the Y-Chromsome of an organism is not characterized, using conserved regions for primer binding sites is beneficial in capturing a region of the DNA. Exon regions of genes are conserved across species allowing the amplification of introns in

genes(Nakahori, Y., Tamura, T., et al 1991). DBY gene is a Y chromosome specific gene that is part of the minor histocompatibility complex (Vogt, van den Muijsenberg & Goulmy, et al, 2002) . As a part of the minor histocompatibility complex it functions in immunity. Due to this important function it is highly conserved across species. The SMCY gene is a Y chromosome homolog of the SMCX gene (Agulnik, Longepied, Ty, Bishop & Mitchell, 1999). The two genes diverged millions of years ago, and still perform the same function as the SMCX gene avoids X-inactivation meaning both males and females have two copies of the gene being expressed. The SMCY gene codes for a minor transplant antigen epitope.

Y chromosome markers can be used to identify male population structure of swarming colonies. If males function as genetic dispersers then haplotypes will be dispersed between swarming colonies evenly. If males show site fidelity then there will be isolation of haplotypes to specific swarming sites.

Methods

Bats were captured at 15 swarming sites in Nova Scotia, New Brunswick, and Quebec from 2009 to 2011. Bat tissue samples of approximately 9mm^2 were obtained from uropatagium (Faure, P. A., et al, 2009; Wilmer, B. & Barratt, E., 1996). Tissue samples were stored in 20% DMSO in a freezer prior to extraction. Samples were extracted using phenol:chloroform extraction (Sambrook J, Russell D, 2001). DNA quantity was assessed using spectrophotometry on a Nanodrop 2000. DNA quality was assessed using a 2% agarose gel stained with SYBR Green I. DNA was then diluted to

5ng/µL for PCR.

Eleven primers described in Hellborg, L., and Ellegren, H, (2003) that amplified Y chromosome introns in *Myotis daubentoni* were tested for amplification at 45ºC, 50ºC, 55ºC, and 60ºC annealing temperature in a 20µL PCR using two known male samples and a known female sample to show male specific amplification. The PCR contained 10ng of template DNA, 1X PCR Buffer ((20 mM Tris–HCl ph 8.4, 50 mM KCl;Promega), 0.2mM dNTPS (Invitrogen), 1.5mM $MgCl_2$ (Promega), 0.3µM of forward and reverse primer, 0.16mg/mL bovine serum albumin (BSA), 0.05U/µL Taq DNA polymerase (Promega). The cycling conditions were as follows: an initial denaturing step of 94°C for 5 minutes; 30 cycles of 94°C for 30 seconds, annealing temperature for 1 minute, and 72°C for 1 minute; followed by a final extension step of 60°C for 45 minutes. All PCR was conducted on Veriti® 96-well thermal cyclers (Applied Biosystems). PCR products were then quantified on a 2% agarose gel stained with ethidium bromide.

Primer pairs that showed successful amplification based on the agarose gels (DBY7, DBY8, UTY11, and SMCY1) were subsequently used to amplify 25 males. The optimal annealing temperatures were identified as follows: at 55°C for DBY7 and DBY8, 50°C for UTY11, and touchdown PCR from 55°C to 45°C decreasing 0.5°C every cycle over 20 cycles then 30 cycles at an annealing temperature of 45°C for SMCY1. Template DNA and Taq concentrations were increased to 20ng DNA for DBY7, DBY8, and UTY11, and 40ng DNA for the SMCY1 locus, and 0.1 U/µL of Taq for all loci. PCR product was quantified using 2% agarose gels stained with ethidium bromide, and diluted to 2.5ng/100bp, 5ng/100bp, and 10ng/100bp in 5µL. ExoSAP was performed using

0.65µL Antarctic phosphatase buffer (New England Biolabs), 0.1µL Antarctic phosphatase (New England Biolabs), and 0.03µL Exonuclease I (New England Biolabs) per reaction to clean the PCR product of excess dNTPs and primers for sequencing. ExoSAP was then conduted in a Veriti® 96-well thermal cycler for 15 minutes at 37°C, then for 15 minute at 80°C. A 15µL sequencing reaction was made using 1X Big Dye Terminator Buffer (Applied Biosystems), 0.25X Big Dye Terminator Reaction Mix (Applied Biosystems), 0.3µM of one primer for the locus, and 5.78µL ExoSAP product. Each sample was sequenced separately with the forward and reverse primer to test for better sequencing product. The cycling conditions were as follows: an initial denaturing step of 96°C for 2 minutes; 30 cycles of 20 seconds at 96°C, 20 seconds at 50°C, and 4 minutes at 60°C. Sequencing product was then de-salted with ethanol precipitation and size-separated and visualized on an ABI 3500xl capillary-based genetic analyzer (Applied Biosystems). Raw sequences were run through MEGA 6.0 (Tamura K, 2013) to call bases. All generated sequences were confirmed visually then run through ClustalX (Larkin MA, 2007) to align sequences and identify variable sites in each locus.

Based on the sequences for SMCY1 three internal primers to pair with SMCY1-F and three internal primers to pair with SMCY1-R were developed to allow sequencing of the SMCY1 locus with less template DNA. The developed primers were matched to the original primer based on melting point calculation. The six primer pairs were tested at annealing temperatures of 45°C, 50°C, and 55°C in a 20µL PCR using two known male samples and a female sample to show male specific binding. The PCR contained 10ng of template DNA, 1X PCR Buffer (20 mM Tris pH 8.4, 50 mM KCl), 0.2mM dNTPS,

1.5mM $MgCl_2$, 0.3µM of forward and reverse primer, 0.16mg/mL bovine serum albumin (BSA), 0.05U/µL Taq DNA polymerase. The cycling conditions were as follows: an initial denaturing step of 94°C for 5 minutes; 30 cycles of 94°C for 30 seconds, annealing temperature for 1 minute, and 72°C for 1 minute; followed by a final extension step of 60°C for 45 minutes. The primer pairs SMCY1-int1-F and SMCY1-R, and SMCY1-int2-R and SMCY1-F were then diluted to 2.5ng/100bp, 5ng/100bp, and 10ng/100bp in 5µL. ExoSAP was performed using 0.65µL Antarctic phosphatase buffer, 0.1µL Antarctic phosphatase, and 0.03µL Exonuclease I per reaction to clean the PCR product of excess dNTPs and primers for sequencing. Samples were then run in a Veriti® 96-well thermal cycler for 15 minutes at 37°C, then for 15 minute at 80°C. A 15µL sequencing reaction was performed using 1X Big Dye Terminator Buffer, 0.25X Big Dye Terminator Reaction Mix, 0.3µM of one primer for the locus, and 5.78µL ExoSAP product. Each sample was sequenced separately with the forward and reverse primer to test for better sequencing product. The cycling conditions were as follows: an initial denaturing step of 96°C for 2 minutes; 30 cycles of 20 seconds at 96°C, 20 seconds at 50°C, and 4 minutes at 60°C. Sequencing product was then de-salted with ethanol precipitation and size-separated and visualized on an ABI 3500xl capillary-based genetic analyzer (Applied Biosystems). Raw sequences were run through MEGA 6.0 to edit sequences. All generated sequences were confirmed visually then run through ClustalX to align sequences and compare to original SMCY1 sequences to confirm the correct region was sequenced.

Ninety-two samples from 15 swarming sites were amplified at loci SMCY1a, SMCY1b, and DBY7. SMCY1a and SMCY1b had an annealing temperature of 50°C, and

DBY7 had an annealing temperature of 55°C. The PCR contained 10ng of template DNA, 1X PCR Buffer (20 mM Tris pH 8.4, 50 mM KCl), 0.2mM dNTPS, 1.5mM $MgCl_2$, 0.3µM of forward and reverse primer, 0.16mg/mL bovine serum albumin (BSA), 0.05U/µL Taq DNA polymerase. The cycling conditions were as follows: an initial denaturing step of 94°C for 5 minutes; 30 cycles of 94°C for 30 seconds, annealing temperature for 1 minute, and 72°C for 1 minute; followed by a final extension step of 60°C for 45 minutes. All PCR was conducted on Veriti® 96-well thermal cyclers (Applied Biosystems). PCR products were then quantified on a 2% agarose gel stained with ethidium bromide. Each locus was then diluted to 10ng/100bp in 5µL. ExoSAP was performed using 0.65µL Antarctic phosphatase buffer, 0.1µL Antarctic phosphatase, and 0.03µL Exonuclease I per reaction to clean the PCR product of excess dNTPs and primers for sequencing. Samples were then run in a Veriti® 96-well thermal cycler for 15 minutes at 37°C, then for 15 minute at 80°C. A 15µL sequencing reaction was performed using 1X Big Dye Terminator Buffer, 0.25X Big Dye Terminator Reaction Mix, 0.3µM of one primer for the locus, and 5.78µL ExoSAP product. Each sample was sequenced with the primers DBY7-F, SMCY-int1-F, and SMCY1-int2-R. The cycling conditions were as follows: an initial denaturing step of 96°C for 2 minutes; 30 cycles of 20 seconds at 96°C, 20 seconds at 50°C, and 4 minutes at 60°C. Sequencing product was then de-salted with ethanol precipitation and size-separated and visualized on an ABI 3500xl capillary-based genetic analyzer (Applied Biosystems). Raw sequences were run through MEGA 6.0 to call bases. All generated sequences were confirmed visually then run through ClustalX to align sequences and compare positive controls to original sequences.

FaBox (Villesen, P, 2007) was used to assign haplotypes to individuals from aligned sequences. SMCY1b sequences were run through DnaSP (Rozas, J. 2009) to convert to a roehl file (.rdf) then opened as a Network project file. In Network (Bandelt H-J, 1999) two median-joining networks were constructed for the SMCY1b locus from the roehl file with default weights for each mutation with epsilon=0, and recommended weights of 10 for transitions, 20 for indels, and 30 for transversions with epsilon=10. The two generated networks were identical.

Aligned sequences were run through ModelGenerator (Keane, T, et al, 2006) to estimate the model of evolution most appropriate for my data set, and the transition transversion ratio. An analysis of molecular variance (AMOVA) was conducted for each locus in Arlequin (Excoffier, L, 2010) using the model and transition transversion ratio obtained from ModelGenerator, and 1000 permutations.

Haplotypes for the DBY7 and SMCY1b locus were combined for each individual, and combined into a multilocus haplotype. Haplotypes of individuals were overlayed onto a map of individual capture locations.

Results

Of the 11 loci that amplified in *Myotis daubentoni* based on (Hellborg, L., Ellegren, H, 2003), only four loci, DBY7, DBY8, UTY11, and SMCY1, amplified selectively in male little brown bats. Primers DBY7-F, DBY8-F, UTY11-R, SMCY1-F, and SMCY1-R were determined best for sequencing from peak height and clarity of electropherogram in MEGA 6.1. After sequencing 25 individuals with each primer the UTY11, and DBY8 loci were found to be monomorphic. Sequencing with SMCY1-F

captured less than half of the SMCY1 locus, along with SMCY1-R. Six primers shown in Table 1 were developed based on SMCY1 sequence data to pair with SMCY1-F and SMCY1-R. Melting points of the developed primers were made to match with the melting points of the original primers they were to pair with. Of the six primer pairs SMCY1-int1-F / SMCY1-R, and SMCY1-int2-R / SMCY1-F amplified selectively in males. The sequenced regions were designated SMCY1a for primer pair SMCY1-int1-F and SMCY1-R and SMCY1b for primer pair SMCY1-int2-R and SMCY1-F with each being approximately 750 base pairs in length.

Table 1. Developed primer sequences from sequence results in the SMCY1 locus and the melting point SMCY1-int(1-3)-F pairs with SMCY1-R and SMCY1-int(1-3)-R pairs with SMCY1-F.

Primers	Primer Sequence	Melting point Tm (°C)
SMCY1-int1-F	5'-GATCTTCTGGTGCCAGGG-3'	58
SMCY1-int2-F	5'-GGCATGTGTCCTGACCC-3'	56
SMCY1-int3-F	5'-TGACCCAGAATCAAATCCG-3'	56
SMCY1-int1-R	5'-GGTTTGATCCATGGTGAGGG-3'	62
SMCY1-int2-R	5'-CACATGCCCAGTTTTTGGG-3'	58
SMCY1-int3-R	5'-TGGGTCAGGACACATGCC-3'	58

After amplification of 92 individuals from 15 swarming sites the SMCY1a locus was determined to be monomorphic. The DBY 7 locus was found to have two haplotypes with one variable site, and locus SMCY1b had seven haplotypes with eight variable sites shown in Table 2. Of the two haplotypes in DBY7 locus there is a single C/T mutation at base position 66. Of the seven haplotypes in the SMCY1b locus one haplotype was abundant with the other haplotypes appearing in low frequencies, as shown in Table 3. Eight different haplotypes were identified by combining the DBY7 and SMCY1b loci

together as shown in Table 4.

Table 2. Shows the sizes, number of variable sites, and number of haplotypes for the regions of the Y chromosome successfully amplified in *Myotis lucifugus*.

Region	Size (bp)	Variable sites	Number of Haplotypes
DBY7	~400	1	2
DBY8	~200	0	1
UTY11	~750	0	1
SMCY1	~2000	8	7
SMCY1a	~750	0	1
SMCY1b	~750	8	7

Table 3. The different nucleotides and base positions for the SMCY1b haplotypes.

	Position (bp)								
SMCY1b Haplotypes	97	98	159	162	163	164	285	503	Number of Individuals
H1	A	-	G	C	A	A	A	T	62
H2	C	-	G	C	A	A	A	T	2
H3	A	-	C	-	-	-	A	T	2
H4	A	A	G	C	A	A	A	T	7
H5	C	-	G	C	A	A	-	T	1
H6	A	-	C	C	A	A	A	T	1
H7	A	-	G	C	A	A	A	C	1

Table 4. How haplotypes of the two loci were combined for the GIS map.

Y Chromosome Haplotype	SMCY1b Haplotype	DBY7 Haplotype	Number of Individuals
1	1	1	50
2	3	1	1
3	4	1	6
4	1	2	1
5	2	1	1
6	7	1	1
7	6	1	1
8	5	1	1

After combining the haplotypes of the two loci, DBY7 and SMCY1b, for a Y chromosome haplotype each haplotype was paired with the individual's capture location, and placed over their geographic location on a map showing the haplotype frequencies per swarming site. The number of individuals with each haplotype separated into swarming sites and the number of samples are shown in Table 5. The map of the two locations were split into a Nova Scotia/New Brunswick map and a Quebec map for greater magnification and resolution of the swarming sites shown in Figure 1 and 2 respectively.

Table 5. Number of individuals with each Y chromosome haplotype divided into swarming site capture locations.

Site	Province	Sample Size	haplotype1	haplotype2	haplotype3	haplotype4	haplotype5	haplotype6	haplotype7	haplotype8
Hayes Cave	NS	4	3	1	0	0	0	0	0	0
Cheverie Cave	NS	5	5	0	0	0	0	0	0	0
Cave of the Bats	NS	3	3	0	0	0	0	0	0	0
Lake Charlotte	NS	2	1	0	1	0	0	0	0	0
Lear	NS	6	6	0	0	0	0	0	0	0
Rawdon Gold Mine	NS	11	6	0	2	1	1	1	0	0
Mine aux Pipistelles	QC	2	2	0	0	0	0	0	0	0
Copperstream-Frontenac	QC	1	1	0	0	0	0	0	0	0
Howes Cave	NB	2	1	0	0	0	0	0	1	0
Whites Cave South	NB	4	2	0	1	0	0	0	0	1
Glenelg Mine	NS	6	5	0	1	0	0	0	0	0
Vault Cave	NS	8	8	0	0	0	0	0	0	0
Minasville Cave	NS	6	5	0	1	0	0	0	0	0
Mont St. Hilare	QC	2	2	0	0	0	0	0	0	0

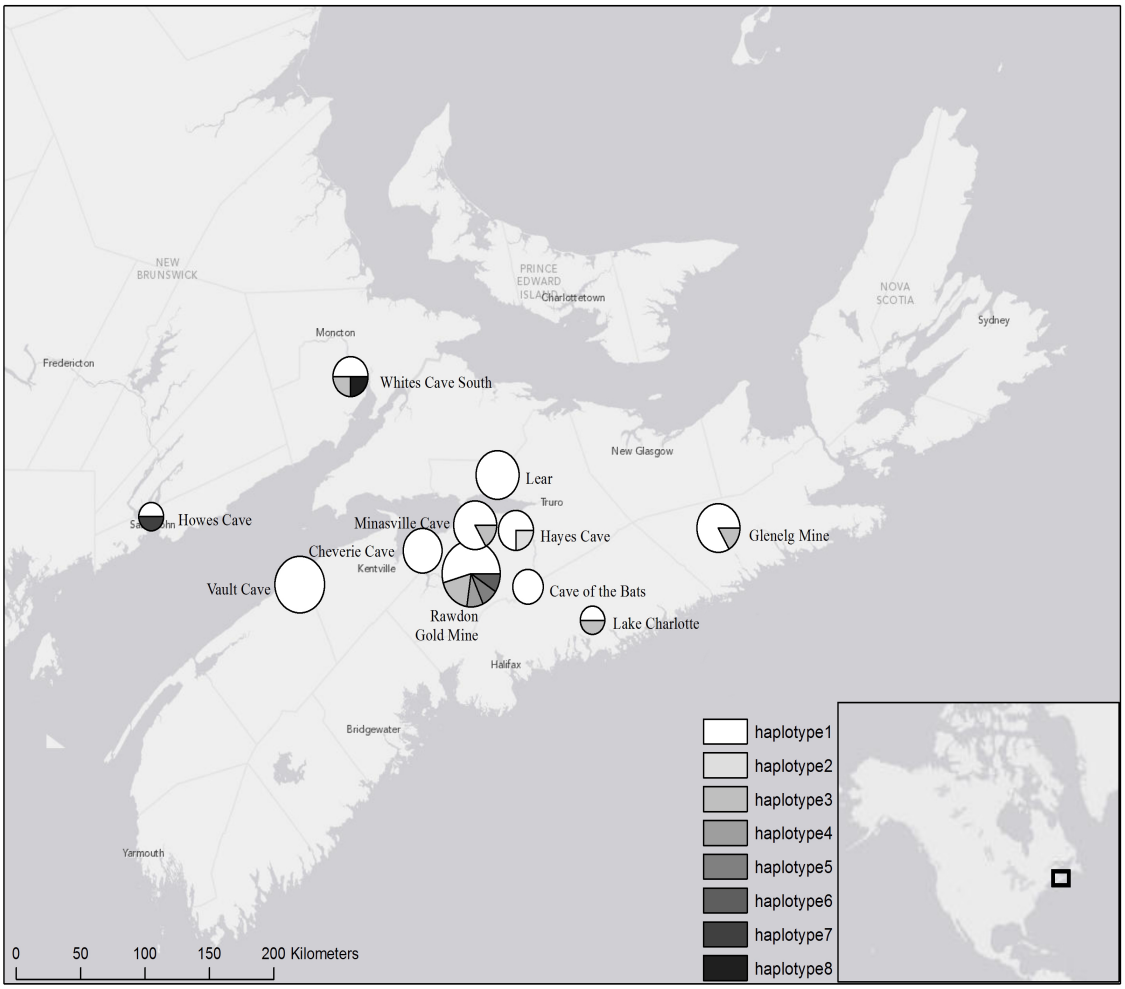

Figure 1. A graphical breakdown of haplotypes paired with the swarming site capture locations in Nova Scotia and New Brunswick. Size of the circle indicates sample number for the swarming site. Numbers shown in Table 5. Berryton, New Brunswick not shown due to lack of samples with both loci sequenced. Cartography credits by Ben Perriman, Esri, DeLorme, NAVTEQ.

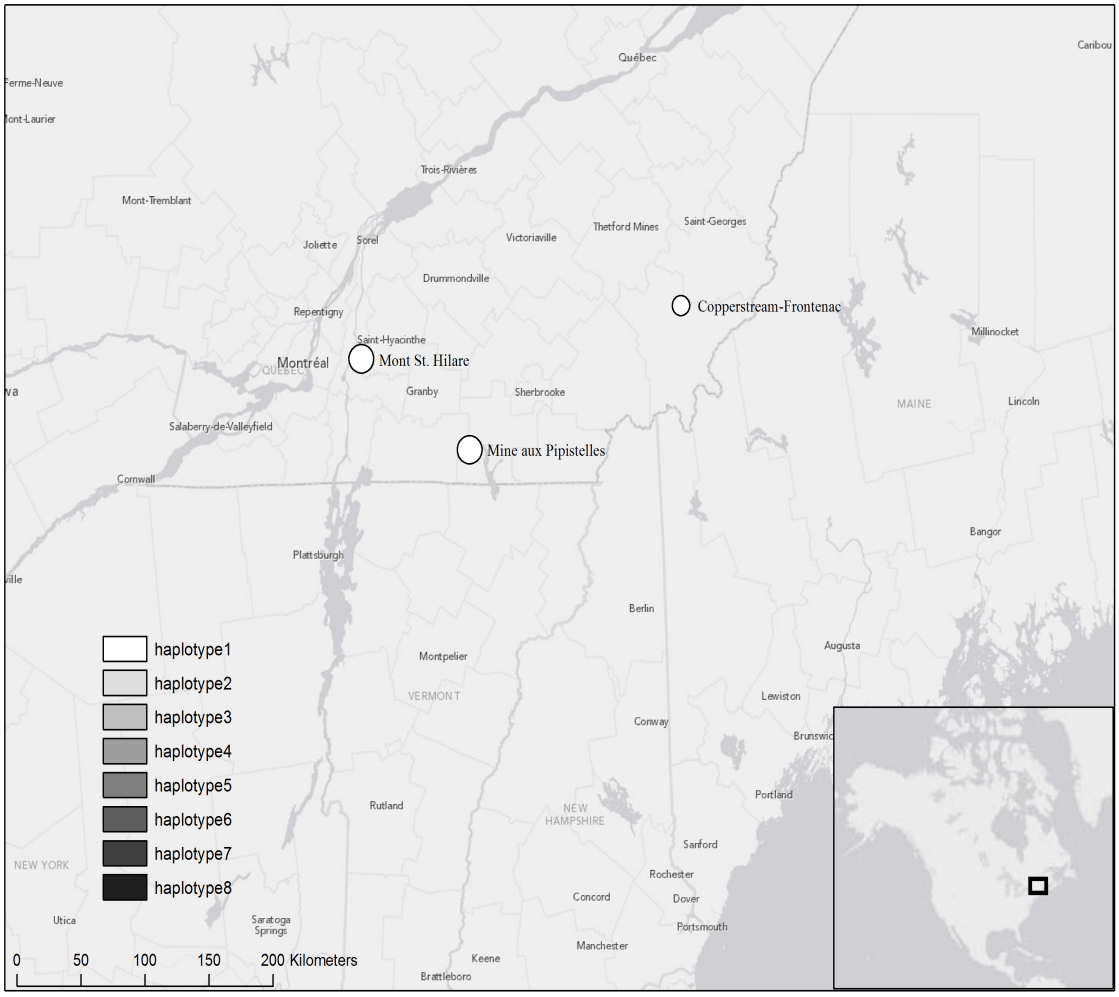

Figure 2. A graphical breakdown of haplotypes paired with the swarming site capture locations in Quebec. Size of the circle indicates sample number for the swarming site. Numbers shown in Table 5. Cartography credits by Ben Perriman, Esri, DeLorme, NAVTEQ.

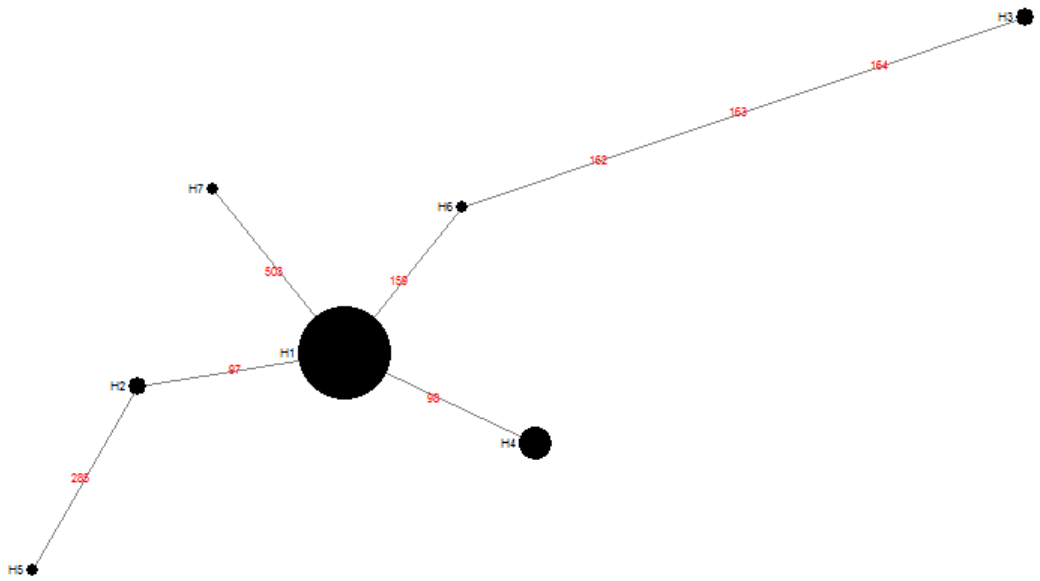

Figure 3. A median-joining network of *M. lucifugus* haplotypes of the SMCY1b locus with mutation positions shown in the links. Size of the circles indicates number of individuals assigned to the haplotype from across the 15 swarming sites.

The generated median-joining network shown in Figure 3 has few nodes and shows a rapid decline in number of individuals with each mutation from the H1 haplotype. A median-joining network could not be constructed for the DBY7 locus because it requires three haplotypes to generate a network.

The model of evolution most appropriate for this data set as calculated by ModelGenerator, was the F81 model, with a transition transversion ratio of 0.4 for DBY7, and 0.43 for SMCY1b.

F_{st} based on haplotype frequencies yielded a value of F_{st} = -0.03475 with p = 0.55718±0.01631 for DBY7 locus, and F_{st} = 0.06651 with p = 0.16227±0.01180 for the

SMCY1b locus. The Φ_{st} for the DBY7 locus was Φ_{st} = -0.03475 with p=0.14076±0.01238. For the SMCY1b locus Φ_{st} = 0.10269 with p = 0.14076±0.01238.

Discussion

Of the eleven tested primer pairs for amplification of Y chromosome loci that were originally used in *Myotis daubentoni,* only four amplified in *Myotis lucifugus*. This could be due to mutations in the binding site, or that given enough time and effort modifying the PCR conditions the loci may work. The benefit of getting the extra loci to amplify is the potential to allow for better resolution by adding more variable sites, but whether these loci are polymorphic is unknown. A better chance of identifying population structure with genetic data, and polymorphic loci is through analyses of microsatellites on the Y chromosome. Microsatellites are hypervariable and accumulate mutations at a higher rate than single base pair substitutions.

Of the variable sites seen there is a lower than expected transition transverison. Across the human genome the transition transversion ratio is approximately 2:1, and the expected ratio to be 0.5. This discrepancy between observed and expected ratios is due to the likelihood of a mutation from a purine to purine, or pyramidine to pyrimidine mutation, a transition, being much greater than a transversion. Transversions are mutations where a nucleotide changes from a purine to pyramidine, or a pyramidine to purine. The expected transition transversion ratio with each mutational event occurring with equal probability is 0.5 since there are twice as many transversions possible than transitions.

The observed transition transversion ratio of 0.4 for DBY7, and 0.43 for SMCY1 suggests neutral mutation probability where the likelihood of a transition is the same as a transversion, or a slight bias towards transversions. Explanations of this result are that these intron sequences are under selection pressures. Another study in snow voles using the SMCY gene, but with the 7^{th} and 8^{th} introns, had 11 transversions and 12 transitions. No transition transversion ratio was published, but the high level of transversions suggests that SMCY introns do not show transition mutation bias (Wandeler & Camenisch, 2011). More species would need to be have SMCY introns sequenced to get a comprehensive understanding of SMCY intron mutations and the possibility of selection in intron sequences. Selection pressure could be the result of alternate splicing with introns incorporated into the expressed gene (Croft et al. 2000), or through intron regulatory sequences (Hare & Palumbi, 2003).

The number of variable sites is low for the two loci resulting in low resolution for detecting any patterns of differential male movements. Only eight haplotypes were found over two loci with each locus having a single allele in high frequency.

AMOVA analyses performed allow for statistical tests of population structuring with genetic data. F_{st} calculations returned a value of less than zero for DBY7, which means that based on haplotype frequencies among subpopulations there is no indication of population structuring. This suggests that there is no population structure. With the SMCY1b locus the F_{st} value is 0.07 and a p-value of 0.16. The Φ_{st} value for DBY7 Φ_{st} = -0.03475 with a p = 0.14. This suggests the same lack of structuring as F_{st} for this locus, but taking into account the evolutionary distance between haplotypes. For the SMCY1b

locus Φ_{st} = 0.10269 with p = 0.14. The AMOVA Φ_{st}, and the F_{st} both suggest a lack of male structuring in swarming sites, but the results are not significant. This data supports the lack of structuring seen with mtDNA and autosomal DN A in maternity colony that *M. lucifugus* is panmictic with high migration rates between swarming sites. This does not necessarily mean that there is no structuring present in males with swarming colonies due to the low resolution of the loci.

The geographic distribution of haplotypes suggests that they are fairly evenly dispersed among swarming sites, but small sample sizes and low variability make it difficult to be confident in this interpretation.

Lack of male structuring is important for conservation of *M. lucifugus* with white-nose syndrome affecting large proportions of the population. This suggests that white-nose syndrome could be spreading among swarming sites from migration of individuals. If a proportion of the population proves resistant to white-nose syndrome then the high migration rates will replenish bat populations to affected areas and spread the resistance to populations. The data present here suggests that there is high movement rates between swarming sites. Therefore, more research is needed to get a comprehensive understanding of *M. lucifugus* population dynamics.

Microsatellite markers are highly variable non-coding regions that can increase the resolution of male population structuring by increasing the number of haplotypes and ability to identify single paternal lines. This does not mean that these markers provide no benefit. With these markers it is a start to sequence *M. lucifugus* male genome. The female genome has been sequenced with 7x coverage using whole genome shotgun

sequencing of a female to move towards *M. lucifugus* being a model organism for genetic studies (Lindblad-Toh, K, et al, 2010). The sequences can also provide information on how to assemble some of the contigs if the male *M. lucifugus* genome is covered with shotgun sequencing.

The Y chromosome markers used in this study had low levels of variability, and therefore provided limited resolution for assessing population structure and gene flow. Analyses of these loci did not detect any significant signs of differentiation, suggesting a high degree of movement of males among swarming sites. However, these results should be interpretated with caution, and conclusions about male-mediated gene flow should await analyses of more variabile loci, which would be more informative.

Acknowledgements: I would like to thank my supervisors Dr. Broders and Dr. Frasier for their support and guidance, and members of the Frasier lab and Broders Lab. Thanks to my family and friends for their support in getting me through this. I would like to thank NSERC and Saint Mary's University for funding this research.

References:

Agnell, R., Butlin, R., & Altringham, J. (2013). Sexual segregation and flexible dating patterns in temperate bats. *PLOS one*, *8*(1), e54194.

Agulnik, A., Longepied, G., Ty, M., Bishop, C., & Mitchell, M. (1999). Mouse h-y encoding smcy gene and its x chromosomal homolog smcx. *Mammalian Genome*, *10*(9), 926-929.

Bandelt H-J, Forster P, Röhl A (1999). Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37-48

Blehert, D., Hicks, A., & Behr, M. (2009). Bat white-nose syndrome: An emerging fungal pathogen?. *Science*,*323*, 227. doi: 10.1126/science.1163874

Broders, H.G., G.M. Quinn, and G.J. Forbes. (2003). Species status, and the spatial and temporal patterns of activity of bats in southwest Nova Scotia, Canada. Northeastern Naturalist 10:383-398.

Broders H.G., G.J. Forbes, S. Woodley, and I.D. Thompson. (2006). Range extent and stand selection for forest-dwelling Northern Long-eared and Little Brown Bats in New Brunswick. Journal of Wildlife Management 70:1174-1184.

Brownlee-Bouboulis, S., & Reeder., D. (2013). White-nose syndrome-affected little brown myotis (myotis lucifugus) increase grooming and other active behaviors during arousals from hibernation. *Journal of Wildlife Diseases*, *49*(4), 850-859. doi: 10.7589/2012-10-242

Butler, J. (2012). *Advanced topics in forensic dna typing*. (pp. 371-403). Maryland, USA: National Institute of Standards and Technology.

Butler, J. (2005). *Forensic dna typing*. (2nd ed., pp. 201-239). New York, USA: Elsevier Academic Press.

Claire, E. (2011). Cryptic species? patterns of maternal and paternal gene flow in eight neotropical bats. *Legal Medicine*, *6*(7), e21460.

Croft, L., S. Schandorff, F. Clark, K. Burrage, P. Arctander, and J. Mattick. (2000). ISIS, the intron information system, reveals the high frequency of alternative splicing in the human genome. Nat. Genet. 24:340–341.

Cryan, P., C.U. Meteyer, J. Boyles, and D. S. Blehert. (2010). Wing pathology of white-nose syndrome in bats suggests life threatening disruption of physiology. BMC Biology 8:135

Excoffier, L. and H.E. L. Lischer (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources. 10: 564-567.

Excoffier, L., Smouse, P., & Quattro, J. (1992). Analysis of molecular variance inferred from metric distances among dna haplotypes: Application to human mitochondrial dna restriction data. *Genetics*,*131*(2), 479-491.

Faure, P. A., Re, D. E., & Clare, E. L. (2009). Wound healing in the flight membranes of big brown bats. *Journal of Mammology*, *90*, 1148-1156.

Felsenstein, J. (1981). Evolutionary trees from dna sequences: a maximum likelihood approach. *Journal of Molecular Evolution*, *17*(6), 368-376.

Gianni, C., G. Caretta, and C. Romano. (2003). Skin infection due to Geomyces pannorum var. pannorum. Mycoses 46:430–432

Greenwood, P. J. (1980). Mating systems, philopatry, and dispersal in birds and mammals. Animal Behaviour 28:1140–1162

Hare, M., & Palumbi, S. (2003). High intron sequence conservation across three mammalian orders suggests functional constraints. *Molecular Biology and Evolution*, *20*(6), 969-978. doi: 10.1093/molbev/msg111

Hellborg, L., & Ellegren, H. (2003). Y chromosome conserved anchored tagged sequences (ycats) for the analysis of mammalian male-specific dna. *Molecular Ecology*, *12*, 283-291.

Iacolina, L., Scandura, M., & Gazzola, A. (2010). Y chromosome microsatellite variation in italian wolves: A contribution to the study of wolf-dog hybridization patterns. *Mammalian Biology*, *75*, 341-347.

Igea, J., Juste, J., & Castresana, J. (2010). Novel intron markers to study the phylogeny of closely related mammalian species. *BMC Evolutionary Biology*, *10*, 369.

Kokko, H., and A. Lopez-Sepulcre. (2006). From individual dispersal to species ranges: perspectives for a changing world. Science 313:789–791.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins

DG. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, 23, 2947-2948.

Lim, B., Engstrom, M., Bickham, J., & Patton, J. (2008). Molecular phylogeny of new world sheath-tailed bats (emballonuridae: Diclidurini) based on loci from the four genetic transmission systems in mammals. *Biological Journal of the Linnean Society*, *93*, 189-209.

Lindblad-Toh, K., Chang, J., Gnerre, S., Clamp, M., & Lander, E. (2010). Myotis lucifugus, whole genome shotgun sequencing project. *Broad Institute of MIT and Harvard*,

Lindner, Daniel L. (2013). Phylogenetic evaluation of Geomyces and allies reveals no close relatives of Pseudogymnoascus destructans, comb. nov., in bat hibernacula of eastern North America. Fungal Biology, 117(9), 638-649

Marshall, W. (1998). Aerial transport of keratinaceous substrate and distribution of the fungus Geomyces pannorum in Antarctic soil. Microbial Ecology. 36, 212-219

Mendez, F., Krahn, T., & Schrack, B., et al. (2013). An african american paternal lineage adds an extremely ancient root to the human y chromosome phylogenetic tree. *The American Journal of Human Genetics*, *92*, 454-459.

Nakahori, Y., Tamura, T., & Nagafuchi, S., et al. (1991). Molecular cloning and mapping of 10 new probes on the human y chromosome. *Genomics*, *9*(4), 765-769

Nei, M. (1972). Genetic distance between populations.*American Naturalist*, *106*(949), 283-292.

Norquay, K., Martinez-Nunez, F., & Dubois, J. (2013). Long-distance movements of little brown bats (myotis lucifugus). *Journal of Mammology*, *94*(2), 506-515.

Papadatou, E., Butlin, R., Pradel, R., & Altringham, J. (2009). Sex-specific roost movements and population dynamics of the vulnerable long-fingered bat, myotis capaccinii. *Biological Conservation*,*142*, 280-289.

Petit, E., Balloux, F., & Excoffier, L. (2012). Mammalian population genetics: why not y?. *TRENDS in Ecology & Evolution*, *17*, 28-33.

Rozas, J. (2009). DNA Sequence Polymorphism Analysis using DnaSP. Pp. 337-350. *In* Posada, D. (ed.) *Bioinformatics for DNA Sequence Analysis; Methods in Molecular Biology Series* Vol. 537. Humana Press, NJ, USA.

Sambrook J, Russell D (2001) Molecular cloning: a laboratory manual, 3rd edn. Cold Spring Harbour Laboratory Press, Cold Spring Harbour New York

Senior, P., Butlin, R., & Altringham, J. (2005). Sex and segregation in temperate bats. *Proc. R. Soc. B*, *93*, 189-209.

Tamura K, Stecher G, Peterson D, Filipski A, and Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0.Molecular Biology and Evolution. 30: 2725-2729.

Thomas M Keane, Christopher J Creevey , Melissa M Pentony, Thomas J Naughton and James O McInerney (2006) Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. BMC Evolutionary Biology, 6:29 - See more at: http://bioinf.nuim.ie/modelgenerator/#sthash.TJROahgJ.dpuf

Villesen, P (2007), *FaBox: an online toolbox for fasta sequences*, Molecular Ecology Notes 7 (6), 965–968. doi:10.1111/j.1471-8286.2007.01821.x

Vogt, M., van den Muijsenberg, J., & Goulmy, E., et al. (2002). The dby gene codes for an hla-dq5−restricted human male-specific minor histocompatibility antigen involved in graft-versus-host disease. *blood*, *99*, 3027-3032. doi: doi:10.1182/blood.V99.8.3027

Wandeler, P., & Camenisch, G. (2011). Identifying y-chromosomal diversity by long-template pcr.*Molecular Ecoology Resources*, *11*, 835-841. doi: 10.1111/j.1755-0998.2011.03013.x

Weir, B., & Cockerham, C. (1984). Estimating f-statistics for the analysis of population structure. *Eolution*, *38*(6), 1358-1370.

Wilmer, W. J., & Barratt, E. (1996). A non-lethal method of tissue sampling for genetic studies of chiropterans. *Bat Research News*, *37*.

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114-138.