Host Population Immunogenetics in a Changing World: Signatures of Selection in the Context of Endemic and Invasive Diseases in Wildlife

A dissertation to the Committee on Graduate Studies

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in the Faculty of Arts and Science

TRENT UNIVERSITY

Peterborough, Ontario, Canada

Environmental and Life Science Graduate Program

January 2025

ABSTRACT

Host Population Immunogenetics in a Changing World: Signatures of Selection in the Context of Endemic and Invasive Diseases in Wildlife

Tristan Mackenzie Baecklund

Heterogeneous environments impose discordant selective pressures on natural populations, where disparate biotic/abiotic factors and variable population connectivity, yield mosaic patterns of genetic variation on the landscape. The ability to maintain or change genetic mosaics of populations becomes key to persistence, as species increasingly need to adapt to rapidly changing environmental and human-associated selective pressures. Specifically, infectious diseases can impose strong and rapid selective pressures on populations, where anthropogenic disruptions of co-evolutionary patterns and altered distributions of hosts and pathogens exacerbate disease risk. Genomic tools provide means to evaluate disease-associated impacts on the genetic landscape of host populations and facilitate implementation of informed conservation efforts.

In this thesis, I evaluate disease dynamics in: 1) a long-standing arctic rabies/arctic fox (*Vulpes lagopus*) system affected by influxes of red fox (*V. vulpes*), and 2) an invasive bat pathogen system, where the North American introduction of *Pseudogymnoascus destructans* (*Pd*) has had variable impacts on bat species and populations. In these systems, signatures of host selection were estimated from temporal and spatial shifts in allelic diversity within genomic regions associated

with immune response, highlighting different host mechanisms to enzootic and invasive diseases. In the arctic rabies/fox system, pathogen variants did not influence red fox local disease responses, reflecting more recent expansions of this host to Arctic regions. In contrast, arctic fox revealed genomic patterns consistent with long-term, co-evolutionary processes. In *Pd*/bat systems, genomic evidence supported the hypothesis that eastern small-footed bats (*Myotis leibii*) were inherently resistant or tolerant to *Pd*, the causative agent of white-nose syndrome (WNS). In contrast, WNS-impacted little brown bat (*M. lucifugus*) populations had varied genomic impacts subsequent to strong selective sweeps from disease.

My research illustrates how immunogenetic profiling, in context of demographic processes inferred from neutral genetics, enhances understanding of the varied impacts of changing disease landscapes on host populations/species; insights relevant to other host-pathogen systems. Building on this thesis, future explorations of low coverage genomes, host-imposed reciprocal selection, and impacts on methylation, transcriptomic and proteomic patterns associated with shifts in genetic diversity, would enable more holistic understanding of the geographic mosaics within these disease systems.

KEYWORDS

Natural Selection, Disease Dynamics, Local Adaptation, Population Genetics, High-throughput Sequencing, Reduced Representation Sequencing, Immune System,

Arctic Rabies, White-nose Syndrome, Eastern Small-footed Bat, Little Brown Bat, Red Fox, Arctic Fox, *Vulpes*, *Myotis*

PREFACE

I have written my thesis in manuscript format. Chapter 2 and Chapter 3 have been published in *PLOS One*, Chapter 4 is currently under peer-review, and Chapter 5 is being prepared for submission. As such, each chapter is formatted in the style of the respective journal with independent References sections for each chapter. Throughout these manuscripts, I have used the collective 'we' where applicable reflecting the collaborative nature of the work. Supplementary materials are listed inappendices with permanent links for published materials or as a Google Drive link for unpublished materials.

ACKNOWLEDGEMENTS

I extend my deepest gratitude and appreciation to my supervisor, Dr. Christopher Kyle. Thank you for igniting my interest in wildlife genomics, and for the continuous stream of encouragement and support, guiding me to strive to improve my skills as a researcher. The opportunities you provided to me throughout my research that made this dissertation possible will not be forgotten. Secondly, I would like to acknowledge my appreciation and admiration to my thesis committee members: Dr. Michael Donaldson, Dr. Aaron Shafer, and Dr. Glen Brown, for providing invaluable guidance, insights, and perspectives throughout the completion of my dissertation. I would also like to thank all of my incredibly supportive collaborators that have made this research possible – thank you Dr. Christina Davy, Dr. Karsten Hueffer, Dr. Maarten Vonhof, and Blake Sasse.

I would not have reached this point without the steadfast support of my colleagues, family (Ellie, Kris, Lindsay), friends (notably, Mason, Xavier, Adrian, Sam, Sumiko, and Erin), and incredible partner Cassie. I may not have said it often enough, but no matter how large or small a part you played in creating this thesis, thank you! I am eternally grateful for the encouragement and time you invested in me throughout my Ph.D. I extend these sentiments to all members of the Kyle lab that I have had the pleasure of working with over the past several years.

Finally, I recognize Diane and Bodo Baecklund for their unwavering support and enthusiasm for my post-secondary education. I will forever be grateful for the opportunities and assistance you have always provided, without hesitation or

second thought. All I have, and will achieve, will always be indebted to the exceptional and selfless parents I have been fortunate enough to have, and honoured to call mum and dad.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Definition
µg	Microgram
µL	Microlitre
A	Adenine
AR	Arctic Rabies (Chapters 2 & 3); Arkansas (Chapter 4)
ARV1	Arctic Rabies Viral Variant 1
ARV2	Arctic Rabies Viral Variant 2
ARV3	Arctic Rabies Viral Variant 3
ARV4	Arctic Rabies Viral Variant 4
ATL-CAN	Atlantic Canada
Bd	*Batrachochytrium dendrobatidis*
bp	Base pair - a unit of length associated with sequences of nucleic acids
C	Cytosine
C2	Gene encoding for: Complement component 2
C3	The gene encoding for: Complement component 3
CCL2	The gene encoding for: C-C motif chemokine ligand 2
CCL5	The gene encoding for: C-C motif chemokine ligand 3
CCR8	The gene encoding for: C-C chemokine receptor type 8
cM	Centimorgan
CWD	Chronic wasting disease
DAPC	Discriminant analysis of principal components
DLA-12	The gene encoding for: MHC class I DLA-88 precursor
DLA-79	The gene encoding for: MHC class Ib precursor
DLA-88	The gene encoding for: Major histocompatibility complex, class I, B
DLA-DMB	The gene encoding for: Major histocompatibility complex, class II, DM beta precursor
DLA-DQA	The gene encoding for: Major histocompatibility complex, class II, DQ alpha precursor
DLA-DQBC1	The gene encoding for: HLA class II histocompatibility antigen, DQ beta 2 chain-like
DLA-DRB1	The gene encoding for: DLA class II histocompatibility antigen, DR-1 beta chain precursor
DNA	Deoxyribonucleic acid
DRB1-e2-like	DRB1-exon 2-like
dsDNA	Double-stranded deoxyribonucleic acid
D_{st}	Inter-population diversity
e.g.	For example
EDTA	Ethylenediaminetetraacetic acid

EMOR	Samples found dead, associated with the initial spread and sweep of white-nose syndrome
E-ON	eastern Ontario
F_{IS}	Inbreeding coefficient
F_{ST}	proportion (ranging from 0 to 1) of the total genetic variance contained in a subpopulation (S) relative to the total genetic variance (T)
G	Guanine
GATK	Genome Analysis Toolkit
GBS	Genotyping-by-sequencing
H_e	Expected heterozygosity
HLA-DPB1	Gene encoding for: HLA class II histocompatibility antigen, DP beta 1 chain-like
H_o	Observed heterozygosity
HRG	The gene encoding for: Histidine-rich glycoprotein
H_s	Intra-population diversity
ICAM1	The gene encoding for: Intercellular adhesion molecule 1 precursor
IFN-gamma	The gene encoding for: Interferon gamma precursor
iHS	Integrated haplotype homozygosity score
IL10	The gene encoding for: Interleukin-10
IL12RB1	The gene encoding for: Interleukin-12 receptor subunit beta-1
IL1A	The gene encoding for: Interleukin-1 alpha precursor
IL1B	The gene encoding for: Interleukin-1 beta
IL1R1	The gene encoding for: Interleukin-1 receptor type 1
IL23	The gene encoding for: Interleukin-23
IL23A	The gene encoding for: Interleukin-23 subunit alpha
IL6	The gene encoding for: Interleukin-6 precursor
INDEL	Insertion-deletion
ITGAM	The gene encoding for: Integrin alpha-M
JAGN1	The gene encoding for: Jagunal homolog 1
JAK2	The gene encoding for: Tyrosine-protein kinase JAK2
kbp	Kilobase pair
LBP	The gene encoding for: Lipopolysaccharide-binding protein
LD	Linkage disequilibrium
LM-PCR	Ligation-mediated polymerase chain reaction
M	Molar
M.	The genus *Myotis*
MAF	Minor allele frequency
MHC	Major histocompatibility complex
mM	Millimolar

mm	Millimeter
MSE	Mean squared error
MYD88	Gene encoding for: Myeloid differentiation primary response protein MyD88
n	Size of a subset of the total sample, N
N	Total sample size; Number of nonsynonymous sites (pN/pS calculations)
NC	North Carolina
NCBI	National Centre for Biotechnology Information
Nd	Total number of nonsynonymous sites
N_E	Effective population size
NFKB1	Gene encoding for: Nuclear factor NF-kappa-B p105 subunit
NH	New Hampshire
NOD1	Gene encoding for: Nucleotide-binding oligomerization domain-containing protein 1
NOS2	Gene encoding for: Nitric oxide synthase 2
NY	New York
ON	Ontario
PCA	Principal component analysis
PCR	Polymerase chain reaction
Pd	*Pseudogymnoascus destructans*
pN/pS	Relative ratio of non-synonymous substitutions per non-synonymous site to the number of synonymous substitutions per synonymous site
PRE	Samples obtained prior to the introduction white-nose syndrome in North America
RAG1	Gene encoding for: V(D)J recombination-activating protein 1
RNA	Ribonucleic acid
RNA-seq	Ribonucleic acid sequencing
S	Number of synonymous sites
Sd	Total number of synonymous sites
SNP	Single nucleotide polymorphism
STAT3	The gene encoding for: Signal transducer and activator of transcription 3
STAT4	The gene encoding for: Signal transducer and activator of transcription 4
STAT6	The gene encoding for: Signal transducer and activator of transcription 6
SURV	Samples obtained 6+ years after the introduction of white-nose syndrome in North America
T	Thymine
t	Time/generations

T_a	Annealing temperature
TAP1	The gene encoding for: Antigen peptide transporter 1 precursor
TBX21	The gene encoding for: T-box transcription factor 21
TE	Tris-EDTA
TICAM1	The gene encoding for: TIR Domain Containing Adaptor Molecule 1
TLR1	The gene encoding for: Toll-like receptor 1
TLR2	The gene encoding for: Toll-like receptor 2
TLR3	The gene encoding for: Toll-like receptor 3
TLR4	The gene encoding for: Toll-like receptor 4
TLR5	The gene encoding for: Toll-like receptor 5
TLR6	The gene encoding for: Toll-like receptor 6
TLR8	The gene encoding for: Toll-like receptor 8
TLR9	The gene encoding for: Toll-like receptor 9
TLRs	The gene encoding for: Toll-like receptors
U	Units of activity
V	Volts
V.	The genus *Vulpes*
VCF	Variant call format
VEP	Variant effect predictor
VT	Vermont
WGS	Whole genome sequencing
WNS	White-nose syndrome
WV	West Virginia
X^2	Chi-squared
XP-EHH	Cross-population extended haplotype homozygosity
α	alpha
Δ	Delta

CHAPTER 1

1.1 Conservation and Management

Natural populations exist across heterogeneous landscapes contemporarily experiencing drastic changes from continued globalization, human-mediated and/or facilitated translocations, and climate change. These changes threaten species and population persistance, where implementing informed conservation and management efforts are critical to mitigating biodiversity loss (Ancillotto et al. 2016; Albers et al. 2023; Baker et al. 2022; Malcolm & Markham 2000; Rose et al. 2019). Informed conservation actions require an understanding of the varied ways in which species and populations interact with heterogeneous landscapes within their respective environments, and how these interactions change given perturbation (Hohenlohe et al. 2021; Lozier & Zayed 2017; Yildirim et al. 2018).

1.2 Selection in Host-Pathogen Systems

Organisms inhabit landscapes where both biotic (e.g., predation, disease, diet) and abiotic (e.g., thermal, precipitation) factors vary through space and time; imposing selective pressures on species, to which populations must adapt. Over time, pressures exerted on organisms lead to relative increases in fitness for a particular environment through adaptation and processes of natural selection (Barton & Partridge 2000). Natural selection can act on populations in different ways such as through: i) balancing selection, where allele frequencies are maintained in a population at frequencies higher than expected by genetic drift through heterozygous advantages or frequency dependant selection; ii) directional

selection, where a genotype or phenotype is maintained at a higher frequency than other possible variants of the trait; iii) disruptive selection, that results in shifts in observed genotypes/phenotypes away from intermediates to extreme phenotypes; iv) stabilizing selection, that alters gene frequencies by selecting for more intermediate phenotypes rather than the extremes, and; v) negative selection, that removes deleterious adaptations (Jones et al. 2023; Latif et al. 2024; Kingsolver et al. 2012; Koene et al. 2024; Dillion et al. 2023; Liu et al. 2023). Understanding the type of selection imposed on populations provides insight into mechanisms responsible for the diversity and complexity of organisms, information directly relevant to the management and persistence of populations in context of shifting suites of selective pressures (Gregory 2009; Hohenlohe et al. 2021; Lozier & Zayed 2017; Yildirim et al. 2018).

Infectious diseases can impose important selective pressures on populations, often over short time frames. Wildlife diseases on the landscape are typically categorized into two phases, epizootic and enzootic. The epizootic phase occurs when a novel pathogen is introduced into a naïve population, or under certain environmental conditions, when a disease already existing on the landscape causes widespread or atypical virulency and/or mortality (Blehert et al. 2009; Briggs, Knapp, & Vrendenburg 2010; Buhnerkempe et al. 2011). Contemporary epizootic events threatening biodiversity are exemplified within several wildlife disease systems, including: 1) West Nile virus, where over two-thirds of crow (*Corvus brachyrhynchos*) populations initially succumbed to the disease (Yaremych et al

2004), 2) facial tumours in Tasmanian devils (*Sarcophilus harrisii*), where populations experienced 90% declines (Miller et al. 2011), and 3) white-nose syndrome in bats that led to extreme (>90%) population declines, extirpation, and projected extinction of some heavily-impacted populations/species (Cheng et al. 2021; Dzal et al. 2011; Francl et al. 2012). Eventually, disease cases subside due to dynamic density-dependant factors (e.g., population densities decreased by fragmentation or mortality) or the accumulation of adaptive traits conferring an increased tolerance or resistance to the disease (Maslo & Fefferman 2015; Newton et al. 2019). At such a time, the enzootic phase will begin, where disease is maintained within host populations or reservoir species without an influx from external sources or other populations producing relatively stable temporal patterns of morbidity (Artois et al. 2009). It is important to note that in both epizootic and enzootic phases, the pathogen remains present on the landscape, yet morbidity, or the amount of associated disease, differs between phases. For example, in rabies virus systems, a disease that can infect most mammals, there are often cycles between epizootic and enzootic phases based on the emergence of new hosts or density dependence (e.g., raccoon rabies; Elmore et al. 2017; Mørk & Prestrud 2004). Specifically, arctic rabies typically cycles between epizootic and enzootic phases that loosely follow population trends of lemming and subsequently arctic fox density (Mørk & Prestrud 2004). However, northern expansions of other susceptible hosts beyond arctic foxes (red foxes, skunks) can disrupt cyclical phases of arctic rabies by acting as reservoirs or altering disease distributions

(Nadin-Davis, Muldoon & Wandeler, 2006; Tabel et al. 1974). In contrast, in the raccoon rabies system it is hypothesized that human-mediated movement of animals disrupted regional co-evolutionary patterns of the host/pathogen, leading to a rapid epizootic spread of disease among North American populations (Elmore et al. 2017; Rupprecht & Smith 1994; Nettles et al. 1979). In rabies systems that can have widespread impacts among natural, domestic and human populations, targeted disease management efforts (culling or vaccine programs) are used mitigate large-scale epizootics (Hopken et al. 2024; Müller & Rupprecht 2023; Gilbert et al. 2018; Condori-Condori et al. 2013). Management of raccoon rabies through vaccine programs has largely mitigated large-scale epizootics or continued spread into rabies-free areas (Elmore et al. 2017; Mainguy et al. 2012; Mainguy et al. 2013). These examples illustrate how continued anthropogenetic impacts alter existing disease dynamics or introduce pathogens to naïve populations, thus posing major risks to the persistence of natural populations.

During epizootic events, infectious diseases have the potential to elicit rapid and marked changes to the genetic landscape of affected populations through natural selection (i.e., selective sweeps). These genetic changes occur within the framework of a co-evolutionary arms race, where pathogens attempt to circumvent host defence mechanisms that are adapting in turn to prevent infection (Berry et al. 1991; Gallana et al. 2013; Laine & Tylianakis 2024). During this arms race, typical population regulation cycles are interrupted by disease entering an epizootic phase, potentially leading to smaller population sizes. Changes in population size could

disrupt underlying allele and genotype frequencies leading to increased/decreased heterozygosity and subsequent departures from Hardy-Weinberg expectations in severe cases (Waples 2015). In hosts, frequencies of variants not beneficial to disease resistance/tolerance are expected to decrease (i.e., maladaptive). In contrast, the frequency of variants associated with positive disease outcomes will increase (i.e., adaptive), resulting in an evolutionary rescue of the population (Maslo & Fefferman 2015). Among natural populations that experience low fecundity and small population sizes, the potential for extirpation or extinction from disease remains when the imposed selective pressure is too strong, and populations are unable to recover from initial declines. In amphibian systems with chytridiomycosis, an infectious disease caused by the fungal pathogen *Batrachochytrium dendrobatidis* (*Bd*), the extinction of almost 100 species has been observed (Fisher & Garner, 2009; Oswald et al. 2020; Smith et al. 2022). Despite the extinction and extirpation of some species, other amphibians, such as the mountain yellow-legged frog (*Rana muscosa*) and Panamanian harlequin frogs *(Atelopus varius* and *Atelopus zeteki*) show evidence of evolutionary rescue, where increased frequencies of underlying advantageous genetic variants explain their persistence (Knapp et al. 2023; Byrne et al. 2021). In contrast, the growling grass frog (*Litoria raniformis*) has persisted in the environment despite the presence of *Bd*, not because of adaptation via natural selection and subsequent evolutionary rescue, but more likely because of an influx of individuals from connected populations (i.e., demographic rescue; Kanarek et al. 2015). Trumbo et al. (2023) also highlight signatures of evolutionary

rescue in boreal toads (*Anaxyrus [Bufo] boreas*), that could also facilitate demographic rescues for other impacted populations of the species. In cases of a demographic rescue, an influx of individuals from connected populations, or via human-mediated translocations, into a sink population is enough to promote growth and recovery. Across these amphibian systems and varied outcomes, it is clear that not all populations are similarly affected by disease and overcome strong selective pressure through a range of strategies. As such, understanding the relationship between pathogen and host, and the relative impacts infectious diseases have on host populations becomes increasingly important information for preserving and conserving natural populations and biodiversity.

1.3 Geographic Mosaic Theory of Coevolution

The geographic mosaic theory of coevolution (GMTC) offers a unique, holistic perspective toward understanding interactions between organisms across heterogeneous landscapes. The GMTC postulates that co-evolutionary dynamics are largely driven by three factors: trait remixing, co-evolutionary hot and cold spots, and varying selective regimes (Gomulkiewicz et al. 2000; Thompson 1994; Thompson 1997; Thompson 2019). Trait remixing concerns natural processes that affect the underlying genetic diversity of populations, such as gene flow, genetic drift, mutation, and population declines and increases following bottleneck events (Gomulkiewicz et al. 2007; Thompson 2019). Hot and cold spots of coevolution correspond to areas where both co-evolving organisms impact fitness of the other (hot spot), or areas where at least one organism is unaffected by the fitness of the

other (cold spot; Gomulkiewicz et al. 2000; Thompson 2019). Disparities in the source and strength of selective pressures on populations across heterogeneous landscapes can lead to differential impacts on genetic diversity, and thus population persistence. Carvalho et al. (2024) recently adopted a GMTC perspective to evaluate the impacts of generalist pathogens on natural environments, using *Bd* in amphibians as a case study. Carvalho et al. (2024) reviewed findings in *Bd* and amphibian systems to identify data supporting co-evolutionary mosaics, concluding that the global disparity of disease outcomes among amphibians combined with apparent reciprocal selection upon the fungus is consistent with expectations of geographic mosaics of co-evolution. In doing so, the findings of Carvalho et al. (2024) highlight the importance of host-pathogen research in predict disease outcomes among generalist pathogens, which likely pose the largest global threats. As such, the GMTC provides a framework to understand how gene flow, genetic drift, selection regimes, and demographics interact between host and pathogens enabling predictions for host population impacts as disease/host dynamics are altered.

1.4 Genomic Tools Provide Informative Metrics

Genomic investigations are increasingly relevant to conservation and management efforts. Genomic data obtained from high-throughput sequencing approaches can be applied and leveraged across all manner of systems, including rare and elusive species that are typically difficult to study (Hohenlohe, Funk, and Rajora 2021; Luikart et al. 2019; Rajora 2019; Supple & Shapiro 2018). Allozymes, microsatellite

genotyping, and mitochondrial DNA sequences have been used to investigate critical aspects of natural populations (Allendorf 2017), including population connectivity (Reusch 2002; Yang et al. 2019), effective population sizes (Thuillet et al. 2005), genetic diversity (Coart et al. 2003), and phylogeny (Barani-Beiranvand et al. 2017). In the context of infectious diseases, genetic stucies can investigate how populations respond to disease exposure, and mechanisms of recovery promoting their persistence (e.g., demographic, evolutionary, and genetic rescues). For example, assessments of major histocompatibility complex (MHC) genes have been used extensively to understand host disease responsesgiven their associations with antigen binding and the highly polymorphic nature of this gene complex (Ekblom et al 2007; Spurgin & Richardson 2010; Eizaguirre et al. 2011; Savage & Zamudio 2011). Specifically, exon-2 of the DRB gene is commonly used as an indicator of the genetic variation of MHC (Doxiadis et al. 2000; Gutierrez-Espeleta et al. 2001), yet studies have indicated that insights from other genomic areas associated with an immune response are needed to better understand host immunogenetic defences against infectious disease (Ottenhoff et al. 2005; Behnke et al. 2003; Hill 2001; Lam-Yuk-Tseung & Gros 2003). As such, more comprehensive genetic variation investigations are required to better understand how disease can elicit different patterns of selection across the entire genome.

The capability to explore thousands of loci across a large number of individuals, simultaneously, has grown to the point that full-genome analyses are becoming commonplace (Holiday et al. 2019; Hohenlohe, Funk, & Rajora 2021).

Genomic assessments are typically categorized into whole genome sequencing (WGS) or reduced representation techniques that provide subsets of data from across the genetic landscape of an organism or a subset of targeted regions, respectively (Hohenlohe, Funk, and Rajora 2021). While WGS approaches are becoming the gold standard for genomic investigations, targeted reduced representation approaches, such as genotyping-by-sequencing (GBS), offer a flexible approach to selectively and simultaneously sequence genomic regions of interest that encompass both neutral and functionally relevant loci for large sample sizes (Pina-Martins et al. 2019; Schweizer et al. 2016). For example, Schweizer et al. (2016) used GBS assays to identify shifts in genetic diversity of the immune responses in gray wolves, despite extensive population connectivity, demonstrating patterns indicative of strong impacts from local selective pressures. Using both neutral and functional loci becomes increasingly important in investigating impacts of infectious diseases on host populations, where the interplay between selection and demographic forces largely dictates population viability and persistence. For example, GBS approaches have been used to demonstrate shifts in genetic diversity in bats affected by white-nose syndrome (Donaldson et al. 2017), and in gopher tortoises (*Gopherus polyphemus*) facing upper respiratory tract disease (Elbers et al. 2018), as well as to monitor and support breeding efforts in Tasmanian devil (*Sarcophilus harrisii*) populations following the emergence of devil facial tumour disease that threatened the species with extinction (Wright et al. 2015). These studies also demonstrate the potential of GBS tools to directly assess signatures of

selection from immunogenetically relevant regions of the genome, while contextualizing these data with neutral markers that provide insights into population connectivity. As such, GBS approaches investigating neutral and functionally relevant regions of host genomes provide the opportunity to identify the overall consequences of selective sweeps on host populations and inform management efforts.

1.5 Study Systems and Approach

In my thesis, I study four host-pathogen systems from a host perspective, to investigate implications on genetic diversity, signatures of selection, and potential population recovery mechanisms. Specifically, I investigate: 1) the relative impact of pathogens in an enzootic system with an invasive host and a host with long-standing pathogen interactions, and 2) in an invasive enzootic system with a previously abundant and panmictic naïve host, and a second genetically structured, low-density host. To investigate the relative impact of these pathogens within these systems, I leverage genomic data generated from previously described assays investigating the immunome of foxes and bats (Donaldson et al. 2018; Donaldson et al. 2017), that selectively sequence conserved regions of the genome putitevly responsible for the immune system based on their annotation in other species. These immunogenetic assays enable the iinvestigation of shifts in genetic diversity indicative of signatures of selection, that can be contextualized by demographic patterns observed from neutral markers. It is important to acknowledge however, that in wildlife disease systems the opportunistic nature and difficulty sampling

natural populations makes it difficult to exclude other selective pressures as the cause of observed frequency shifts.

Arctic rabies is an enzootic circumpolar virus of the genus *Lyssavirus*, where in North America four viral variants occupy distinct distributions (Hueffer, Tryland, & Dresvyanikova, 2022; Hueffer & Murphy, 2018; Huettmann, Magnuson, & Hueffer, 2017; Goldsmith et al. 2016; Kuzmin et al. 2008; Nadin-Davis, Sheen, & Wandeler, 2012). While arctic rabies variant three is circumpolar along northern coasts of North America and Eurasia, variants two and four are largely restricted to Alaska, and variant one is primarily found in Ontario around and below Hudson's Bay (Huettmann, Magnuson, & Hueffer, 2017; Goldsmith et al. 2016; Kuzmin et al. 2008; Nadin-Davis, Sheen, & Wandeler, 2012). Arctic foxes (*Vulpes lagopus*), have a long-standing relationship with arctic rabies as the primary host (Mørk & Prestrud 2004). Range shifts, likely in response to resource availability, environmental changes, or continued human disruptions are enabling red foxes to expand into regions predominated by arctic fox, outcompeting them for resources, and acting as a more recent host/reservoir for arctic rabies (Hersteinsson & Macdonald 1992; Gallant et al. 2012). Past research suggests that the neutral genetic structure of arctic foxes aligns with distributions of disease variants (Goldsmith et al. 2016), and immunogenetic structure exists between red foxes inhabiting rabies endemic and absent regions (Donaldson et al. 2017). Beyond findings from these previous studies, several biological aspects of the arctic rabies system in Alaska remain unclear including: 1) the underlying mechanisms responsible for geographically

distinct distributions of arctic rabies variants between epizootics, 2) roles red foxes play in perpetuating arctic rabies, 3) co-evolutionary processes between rabies and arctic foxes, and 4) how continued environmental changes may impact the spread of rabies in the Arctic. Addressing these questions would provide a better understanding of host-pathogen interactions, enabling predictions on the relative impacts environmental changes may have on rabies and other disease systems in the Arctic.

Pseudogymnoascus destructans (*Pd*) is a fungal pathogen and the causative agent of white-nose syndrome (WNS) introduced from Eurasia into North America in 2006 (Leopardi, Blake, & Puechmaille, 2015). White-nose syndrome epizootics devastated several species of Nearctic hibernating bats with > 90% declines in some populations (Cheng et al. 2021); a result of a cascade of physiological changes infected bats undergo leading to increased morbidity and mortality (Verant et al. 2014). Differences in environmental, behavioural, ecological, genetic and individual factors produced variable susceptibility to WNS among Nearctic bats (Cryan et al. 2013; Davy et al. 2020; Hayman et al. 2016). Eastern small-footed bats (*Myotis leibii*) exhibit population genetic structure based on a limited number of samples using amplified fragment length polymorphisms (N=10; Ammerman, Lee, & Pfau 2016). Where population declines following the introduction of *Pd* have not been observed despite contact with the pathogen across their range (Langwig et al. 2012; Frick et al. 2017; Anderson 2018; Reynolds et al. 2021; O'Keefe et al. 2019; Hooton et al. 2023; Sasse & Perry 2023; but see reported declines in Frick et al.

2015). In contrast, previously abundant little brown bats (*Myotis lucifugus*) experienced drastic population declines associated with the introduction of *Pd* (Cheng et al. 2021). Evidence of subsequent evolutionary rescue in some *M. lucifugus* populations has been observed, along with varied impacts among populations across eastern North America (Auteri & Knowles 2020; Davy et al. 2017; Donaldson et al. 2017; Lilley et al. 2020; Gignoux-Wolfsohn et al. 2021; Langwig et al. 2015; Maslo et al. 2015; Frick et al. 2017; Dobony and Johnson 2018; Ineson 2020; Hooton et al. 2023). Given these genomic investigations identified population declines and evolutionary rescues in *M. Lucifugus*, similar research should be undertaken to understand why, and by what means, *M. leibii* populations have not undergone declines as observed in species such as *M. lucifugus*. Furthermore, knowledge gaps remain on how varied population impacts result from the combined influences of life histories, demographics, and local selective regimes between populations, and the extent of those patterns among the panmictic metapopulation of eastern Canadian *M. lucifugus*. There is a need to further understand mechanisms underlying the apparent variability of WNS susceptibility among bat populations, and how geographic mosaics may influence the overall impacts imposed on these species by the disease. These data would provide meaningful insights for managing Nearctic bat populations as the disease continues to spread westward.

1.6 Thesis Hypothesis, Objectives, and Data Chapters

In this thesis, I investigate immunogenetic variation within host populations and assess impacts of selective pressures associated with infectious diseases by exploring the following hypotheses: i) as natural populations are exposed to spatially and temporally heterogeneous selective pressures associated with disease, in context of underlying variation in immunogenetic regions, then affected populations should demonstrate discordant shifts in immunogenetic variation throughout disease occurrence, ii) if an invasive pathogen is introduced to naïve populations that subsequently experience drastic population declines, then rapid shifts in underlying allele frequencies should be indicative of signatures of selection assocaited with disease outcome, and iii) if rapid shifts in allele frequencies are observed throughout disease occurrence, and genetic drift and gene flow are excluded as influencing these observed frequency changes, then the shifts in genetic variation are presumed to be attributed to the disease in question.

The overall objectives of my thesis are to address the aforementioned hypotheses in different host-pathogen systems to: i) identify signatures of selection across the range of study systems, ii) evaluate the use of reduced representation techniques in monitoring changing disease dynamics in wildlife populations, iii) elucidate the relative impacts of disease on host populations, iv) assess the consequences of continued changes in disease dynamics in these systems, and v) utilize a GMTC perspective to compare and contrast observed differences between endemic and epidemic systems. By investigating the hypotheses above, and

meeting the outlined research objectives, these data have the potential to inform management and conservation initiatives to mitigate the impacts of wildlife disease on natural populations.

In **Chapter 2**, I explore the role of red foxes (*Vulpes vulpes*) in maintaining unique distributions of arctic rabies viral variants in North America, as continued range expansions could potentially disrupt or alter disease dynamics in Alaska. In **Chapter 3**, I investigate the role of arctic foxes (*V. lagopus*), in maintaining spatial distributions of arctic rabies viral variants and the potential for patterns consistent with differential selection imposed on fox populations. This chapter provides the first immunogenetic investigation of the primary arctic rabies host increasing our understanding of this enzootic Arctic disease. **Chapter 2**, alongside **Chapter 3**, investigate the relative impacts of arctic rabies on two fox species with the overarching goal of increasing our understanding of how the distributions of viral variants are maintained, how species locally adapt to disease, and how invasive alternate hosts increase disease risk by potentially undermining co-evolutionary relationships.

In **Chapter 4**, I explore the relative impacts of *Pd* on the rare and elusive eastern small-footed bat (*Myotis leibii*). I specifically test the hypothesis that *M. leibii* was inherently resistant to the pathogen upon introduction to North America as demonstrated by a combination of demographic and immunogenetic evidence.

In **Chapter 5**, I leverage temporal sampling of little brown bats (*Myotis lucifugus*) across portions of the eastern Canadian range, including samples prior to the introduction of *Pd*, while WNS was epizootic, and now enzootic with high-throughput sequencing of immunogenetic regions. These data were used to elucidate patterns of selection between populations and identify underlying mechanisms responsible for the apparent recovery of some, but not all populations.

Together, these chapters illustrate how immunogenetic profiling, in the context of demographic processes inferred from neutral genetics, can be used to enhance understanding of the varied impacts of changing disease landscapes on host populations/species.

1.7 References

1. Ancillotto L, Santini L, Ranc N, Maiorano L, Russo D. Extraordinary range expansion in a common bat: the potential roles of climate change and urbanisation. The Science of Nature. 2016 Apr;103:1-8.
2. Albers HJ, Kroetz K, Sims C, Ando AW, Finnoff D, Horan RD, Liu R, Nelson E, Merkle J. Where, when, what, and which? Using characteristics of migratory species to inform conservation policy questions. Review of Environmental Economics and Policy. 2023 Jan
3. Allendorf, F.W., 2017. Genetics and the conservation of natural populations: allozymes to genomes.
4. Ammerman LK, Lee DN, Pfau RS. Patterns of genetic divergence among *Myotis californicus*, *M. ciliolabrum*, and *M. leibii* based on amplified fragment length polymorphism. Acta Chiropterologica. 2016 Dec 30;18(2):336-46.
5. Anderson, B. N. (2018). Immune Function and Metabolism of Hibernating North American Bats with White-Nose Syndrome.
6. Artois M, Bengis R, Delahay RJ, Duchêne MJ, Duff JP, Ferroglio E, Gortazar C, Hutchings MR, Kock RA, Leighton FA, Mörner T. Wildlife disease surveillance and monitoring. Management of disease in wild mammals. 2009:187-213.
7. Auteri, G. G., & Knowles, L. L. (2020). Decimated little brown bats show potential for adaptive change. Scientific reports, 10(1), 1-10.
8. Baker RE, Mahmud AS, Miller IF, Rajeev M, Rasambainarivo F, Rice BL, Takahashi S, Tatem AJ, Wagner CE, Wang LF, Wesolowski A. Infectious disease in an era of global change. Nature Reviews Microbiology. 2022
9. Barani-Beiranvand H, Aliabadian M, Irestedt M, Qu Y, Darvish J, Székely T, Van Dijk RE, Ericson PG. Phylogeny of penduline tits inferred from mitochondrial and microsatellite genotyping. Journal of Avian Biology. 2017 Jul;48(7):932-40
10. Barton N, Partridge L. Limits to natural selection. BioEssays. 2000 Dec;22(12):1075-84.
11. Behnke JM, Iraqi F, Menge D, Baker RL, Gibson J, Wakelin D. Chasing the genes that control resistance to gastrointestinal nematodes. Journal of Helminthology. 2003 Jun;77(2):99–109. pmid:12756063
12. Berry AJ, Ajioka JW, Kreitman M. Lack of polymorphism on the Drosophila fourth chromosome resulting from selection. Genetics. 1991 Dec 1;129(4):1111-7.
13. Blehert DS, Hicks AC, Behr M, Meteyer CU, Berlowski-Zier BM, Buckles EL, Coleman JT, Darling SR, Gargas A, Niver R, Okoniewski JC. Bat white-

nose syndrome: an emerging fungal pathogen?. Science. 2009 Jan 9;323(5911):227-.
14. Briggs CJ, Knapp RA, Vredenburg VT. Enzootic and epizootic dynamics of the chytrid fungal pathogen of amphibians. Proceedings of the National Academy of Sciences. 2010 May 25;107(21):9695-700.
15. Buhnerkempe MG, Eisen RJ, Goodell B, Gage KL, Antolin MF, Webb CT. Transmission shifts underlie variability in population responses to Yersinia pestis infection. PloS one. 2011 Jul 25;6(7):e22498.
16. Byrne AQ, Richards-Zawacki CL, Voyles J, Bi K, Ibáñez R, Rosenblum EB. Whole exome sequencing identifies the potential for genetic rescue in iconic and critically endangered Panamanian harlequin frogs. Global Change Biology. 2021 Jan;27(1):50-70.
17. Carvalho T, Belasen AM, Toledo LF, James TY. Coevolution of a generalist pathogen with many hosts: the case of the amphibian chytrid Batrachochytrium dendrobatidis. Current Opinion in Microbiology. 2024 Apr 1;78:102435.
18. Cheng TL, Reichard JD, Coleman JT, Weller TJ, Thogmartin WE, Reichert BE, Bennett AB, Broders HG, Campbell J, Etchison K, Feller DJ. The scope and severity of white-nose syndrome on hibernating bats in North America. Conservation Biology. 2021 Oct;35(5):1586-97.
19. Cheng TL, Reichard JD, Coleman JT, Weller TJ, Thogmartin WE, Reichert BE, Bennett AB, Broders HG, Campbell J, Etchison K, Feller DJ. The scope and severity of white-nose syndrome on hibernating bats in North America. Conservation Biology. 2021 Oct;35(5):1586-97.
20. Coart E, Vekemans X, Smulders MJ, Wagner I, Van Huylenbroeck J, Van Bockstaele E, Roldán-Ruiz I. Genetic variation in the endangered wild apple (Malus sylvestris (L.) Mill.) in Belgium as revealed by amplified fragment length polymorphism and microsatellite markers. Molecular ecology. 2003 Apr;12(4):845-57.
21. Condori-Condori RE, Streicker DG, Cabezas-Sanchez C, Velasco-Villa A. Enzootic and epizootic rabies associated with vampire bats, Peru. Emerging infectious diseases. 2013 Sep;19(9):1463.
22. Cryan, P. M., Meteyer, C. U., Boyles, J. G., & Blehert, D. S. (2013). White-nose syndrome in bats: illuminating the darkness. BMC biology, 11(1), 1-4.
23. Davy, C. M., Donaldson, M. E., Bandouchova, H., Breit, A. M., Dorville, N. A., Dzal, Y. A., ... & Kyle, C. J. (2020). Transcriptional host–pathogen responses of Pseudogymnoascus destructans and three species of bats with white-nose syndrome. Virulence, 11(1), 781-794.
24. Davy, C. M., Donaldson, M. E., Rico, Y., Lausen, C. L., Dogantzis, K., Ritchie, K., ... & Kyle, C. J. (2017). Prelude to a panzootic: Gene flow and

immunogenetic variation in northern little brown myotis vulnerable to bat white-nose syndrome. Facets, 2(2), 690-714.
25. Dillon MN, Thomas R, Mousseau TA, Betz JA, Kleiman NJ, Reiskind MO, Breen M. Population dynamics and genome-wide selection scan for dogs in Chernobyl. Canine Medicine and Genetics. 2023 Dec;10(1):1-4. Liu P, Li G, Zhao N, Song X, Wang J, Shi X, Wang B, Zhang L, Dong L, Li Q, Liu Q. Neutral forces and balancing selection interplay to shape the major histocompatibility complex spatial patterns in the striped Hamster in Inner Mongolia: suggestive of broad-scale local adaptation. Genes. 2023 Jul 22;14(7):1500.
26. Dobony CA, Johnson JB. Observed resiliency of little brown myotis to long-term white-nose syndrome exposure. Journal of Fish and Wildlife Management. 2018 Jun 1;9(1):168-79.
27. Donaldson ME, Davy CM, Willis CK, McBurney S, Park A, Kyle CJ. Profiling the immunome of little brown myotis provides a yardstick for measuring the genetic response to white-nose syndrome. Evolutionary applications. 2017 Dec;10(10):1076-90.
28. Donaldson ME, Rico Y, Hueffer K, Rando HM, Kukekova AV, Kyle CJ. Development of a genotype-by-sequencing immunogenetic assay as exemplified by screening for variation in red fox with and without endemic rabies exposure. Ecology and evolution. 2018 Jan;8(1):572-83.
29. Doxiadis GG, Otting N, de Groot NG, Noort R, Bontrop RE. Unprecedented polymorphism of Mhc-DRB region configurations in rhesus macaques. The Journal of Immunology. 2000 Mar 15;164(6):3193–9. pmid:10706710
30. Dzal Y, McGuire LP, Veselka N, Fenton MB. Going, going, gone: the impact of white-nose syndrome on the summer activity of the little brown bat (Myotis lucifugus). Biology letters. 2011 Jun 23;7(3):392-4.
31. Eizaguirre C, Lenz TL, Sommerfeld RD, Harrod C, Kalbe M, Milinski M. Parasite diversity, patterns of MHC II variation and olfactory based mate choice in diverging three-spined stickleback ecotypes. Evolutionary Ecology. 2011 May 1;25(3):605–22.
32. Ekblom R, Saether SA, Jacobsson P, Fiske P, Sahlman T, Grahn M, et al. Spatial pattern of MHC class II variation in the great snipe (Gallinago media). Molecular ecology. 2007 Apr;16(7):1439–51. pmid:17391268
33. Elbers JP, Brown MB, Taylor SS. Identifying genome-wide immune gene variation underlying infectious disease in wildlife populations–a next generation sequencing approach in the gopher tortoise. BMC genomics. 2018 Dec 1;19(1):64. pmid:29351737
34. Elmore SA, Chipman RB, Slate D, Huyvaert KP, VerCauteren KC, Gilbert AT. Management and modeling approaches for controlling raccoon

rabies: The road to elimination. PLoS neglected tropical diseases. 2017 Mar 16;11(3):e0005249.
35. Fisher MC, Garner TW, Walker SF. Global emergence of Batrachochytrium dendrobatidis and amphibian chytridiomycosis in space, time, and host. Annual review of microbiology. 2009 Oct 13;63:291-310.
36. Francl KE, Ford WM, Sparks DW, Brack Jr V. Capture and reproductive trends in summer bat communities in West Virginia: assessing the impact of white-nose syndrome. Journal of Fish and Wildlife Management. 2012 Jun;3(1):33-42.
37. Frick WF, Puechmaille SJ, Hoyt JR, Nickel BA, Langwig KE, Foster JT, Barlow KE, Bartonička T, Feller D, Haarsma AJ, Herzog C. Disease alters macroecological patterns of N orth A merican bats. Global Ecology and Biogeography. 2015 Jul;24(7):741-9.
38. Frick, W. F., Cheng, T. L., Langwig, K. E., Hoyt, J. R., Janicki, A. F., Parise, K. L., ... & Kilpatrick, A. M. (2017). Pathogen dynamics during invasion and establishment of white-nose syndrome explain mechanisms of host persistence.
39. Gallana M, Ryser-Degiorgis MP, Wahli T, Segner H. Climate change and infectious diseases of wildlife: altered interactions between pathogens, vectors and hosts. Current Zoology. 2013 Jun 1;59(3):427-37.
40. Gallant D, Slough BG, Reid DG, Berteaux D. Arctic fox versus red fox in the warming Arctic: four decades of den surveys in north Yukon. Polar biology. 2012 Sep;35:1421-31.
41. Gignoux-Wolfsohn, S. A., Pinsky, M. L., Kerwin, K., Herzog, C., Hall, M., Bennett, A. B., ... & Maslo, B. (2021). Genomic signatures of selection in bats surviving white-nose syndrome. Molecular Ecology, 30(22), 5643-5657.
42. Gilbert A, Johnson S, Walker N, Wickham C, Beath A, VerCauteren K. Efficacy of Ontario Rabies Vaccine Baits (ONRAB) against rabies infection in raccoons. Vaccine. 2018 Aug 6;36(32):4919-26. Goldsmith EW, Renshaw B, Clement CJ, Himschoot EA, Hundertmark KJ, Hueffer K. Population structure of two rabies hosts relative to the known distribution of rabies virus variants in Alaska. Molecularecology. 2016 Feb; 25(3):675–88. https://doi.org/10.1111/mec.13509 PMID: 26661691
43. Gomulkiewicz R, Drown DM, Dybdahl MF, Godsoe W, Nuismer SL, Pepin KM, Ridenhour BJ, Smith CI, Yoder JB. Dos and don'ts of testing the geographic mosaic theory of coevolution. Heredity. 2007 May;98(5):249-58.
44. Gomulkiewicz, R., J. N. Thompson, R. D. Holt, S. L. Nuismer, and M. E. Hochberg. 2000. Hot spots, cold spots, and the geographic mosaic theory of coevolution. Am. Nat 156–174.

45. Gregory TR. Understanding natural selection: essential concepts and common misconceptions. Evolution: Education and outreach. 2009 Jun;2:156-75.
46. Gutierrez-Espeleta GA, Hedrick PW, Kalinowski ST, Garrigan D, Boyce WM. Is the decline of desert bighorn sheep from infectious disease the result of low MHC variation?. Heredity. 2001 Apr;86(4):439. pmid:11520344
47. Hayman DT, Pulliam JR, Marshall JC, Cryan PM, Webb CT. Environment, host, and fungal traits predict continental-scale white-nose syndrome in bats. Science advances. 2016 Jan 29;2(1):e1500831.
48. Heard GW, Thomas CD, Hodgson JA, Scroggie MP, Ramsey DS, Clemann N. Refugia and connectivity sustain amphibian metapopulations afflicted by disease. Ecology Letters. 2015 Aug;18(8):853-63.
49. Hersteinsson P, Macdonald DW. Interspecific competition and the geographical distribution of red and arctic foxes Vulpes vulpes and Alopex lagopus. Oikos. 1992 Sep 1:505-15.
50. Hill AV. The genomics and genetics of human infectious disease susceptibility. Annual review of genomics and human genetics. 2001 Sep;2(1):373–400. pmid:11701655
51. Hohenlohe PA, Funk WC, Rajora OP. Population genomics for wildlife conservation and management. Molecular Ecology. 2021 Jan;30(1):62-82.
52. Holliday JA, Hallerman EM, Haak DC. Genotyping and sequencing technologies in population genetics and genomics. Population genomics: Concepts, approaches and applications. 2019:83-125.
53. Hooton LA, Adams AA, Cameron A, Fraser EE, Hale L, Kingston S, Fenton MB, McGuire LP, Stukenholtz EE, Davy CM. Effects of bat white-nose syndrome on hibernation and swarming aggregations of bats in Ontario. Canadian Journal of Zoology. 2023 May 29
54. Hopken MW, Gigante C, Gilbert AT, Chipman RB, Kirby JD, Condori RE, Mills S, Hartley C, Forbes J, Dettinger L, Xia D. Genetic Tracking of a Rabid Coyote (Canis latrans) Detected beyond a Rabies Enzootic Area in West Virginia, US. Journal of Wildlife Diseases. 2024 May 1.
55. Hueffer K, Murphy M. Rabies in Alaska, from the past to an uncertain future. International journal of circumpolar health. 2018 Jan 1; 77(1):1475185. https://doi.org/10.1080/22423982.2018.1475185 PMID:29764319
56. Hueffer K, Tryland M, Dresvyanikova S. Rabies in the Arctic. InArctic One Health: Challenges for Northern Animals and People 2022 Apr 24 (pp. 211-226). Cham: Springer International Publishing.
57. Huettmann F, Magnuson EE, Hueffer K. Ecological niche modeling of rabies in the changing Arctic of Alaska. Acta Veterinaria Scandinavica.

2017 Dec; 59(1):18. https://doi.org/10.1186/s13028-017-0285-0 PMID: 28320440

58. Ineson KM. Demography of a recovery: tracking the rebound of little brown bat (Myotis lucifugus) populations (Doctoral dissertation, University of New Hampshire).
59. Jones GM, Goldberg JF, Wilcox TM, Buckley LB, Parr CL, Linck EB, Fountain ED, Schwartz MK. Fire-driven animal evolution in the Pyrocene. Trends in Ecology & Evolution. 2023 Jul 13.
60. Kanarek AR, Webb CT, Barfield M, Holt RD. Overcoming Allee effects through evolutionary, genetic, and demographic rescue. Journal of biological dynamics. 2015 Jan 1;9(1):15-33
61. Kingsolver JG, Diamond SE, Siepielski AM, Carlson SM. Synthetic analyses of phenotypic selection in natural populations: lessons, limitations and future directions. Evolutionary Ecology. 2012 Sep;26:1101-18.
62. Knapp RA, Wilber MQ, Byrne AQ, Joseph MB, Smith TC, Rothstein AP, Grasso RL, Rosenblum EB. Reintroduction of resistant frogs facilitates landscape-scale recovery in the presence of a lethal fungal disease. bioRxiv. 2023 May 24:2023-05.
63. Koene JP, Bean CW, Kristjánsson BK, Skúlason S, Leblanc CA, Adams CE. Patterns, processes and conservation management consequences of intraspecific diversity, illustrated by fishes from recently glaciated lakes. Aquatic Conservation: Marine and Freshwater Ecosystems. 2024 Feb;34(2):e4076.
64. Kuzmin IV, Hughes GJ, Botvinkin AD, Gribencha SG, Rupprecht CE. Arctic and Arctic-like rabies viruses: distribution, phylogeny and evolutionary history. Epidemiology & Infection. 2008 Apr; 136(4):509–19.
65. Laine AL, Tylianakis JM. The coevolutionary consequences of biodiversity change. Trends in Ecology & Evolution. 2024 May 4.
66. Lam-Yuk-Tseung S, Gros P. Genetic control of susceptibility to bacterial infections in mouse models. Cellular microbiology. 2003 May;5(5):299–313. pmid:12713489
67. Langwig KE, Frick WF, Reynolds R, Parise KL, Drees KP, Hoyt JR, Cheng TL, Kunz TH, Foster JT, Kilpatrick AM. Host and pathogen ecology drive the seasonal dynamics of a fungal disease, white-nose syndrome. Proceedings of the Royal Society B: Biological Sciences. 2015 Jan 22;282(1799):20142335.
68. Langwig, K. E., Frick, W. F., Bried, J. T., Hicks, A. C., Kunz, T. H., & Marm Kilpatrick, A. (2012). Sociality, density-dependence and microclimates determine the persistence of populations suffering from a novel fungal disease, white-nose syndrome. Ecology letters, 15(9), 1050-1057.

69. Latif EN, Noordin NR, Shahari S, Amir A, Lau YL, Cheong FW, Abdullah ML, Fong MY. Genetic polymorphism and clustering of the Plasmodium cynomolgi Duffy binding protein 1 region II of recent macaque isolates from Peninsular Malaysia. Parasitology research. 2024 Jan;123(1):1-0.
70. Leopardi S, Blake D, Puechmaille SJ. White-nose syndrome fungus introduced from Europe to North America. Current Biology. 2015 Mar 16;25(6):R217-9.
71. Lilley, T. M., Wilson, I. W., Field, K. A., Reeder, D. M., Vodzak, M. E., Turner, G. G., ... & Paterson, S. (2020). Genome-wide changes in genetic diversity in a population of Myotis lucifugus affected by white-nose syndrome. G3: Genes, Genomes, Genetics, 10(6), 2007-2020.
72. Liu P, Li G, Zhao N, Song X, Wang J, Shi X, Wang B, Zhang L, Dong L, Li Q, Liu Q. Neutral forces and balancing selection interplay to shape the major histocompatibility complex spatial patterns in the striped Hamster in Inner Mongolia: suggestive of broad-scale local adaptation. Genes. 2023 Jul 22;14(7):1500.
73. Lozier, J. D., & Zayed, A. (2017). Bee conservation in the age of genomics. Conservation Genetics, 18(3), 713-729.
74. Luikart, G., Kardos, M., Hand, B. K., Rajora, O. P., Aiken, S. N., & Hohenlohe, P. A.(2019). Population genomics: Advancing understanding of nature. In Rajora, O. P. (Ed.), Population genomics: Concepts, approaches, and applications. (pp. 3–79). Springer Nature Switzerland AG.
75. Mainguy J, Rees EE, Canac-Marquis P, Bélanger D, Fehlner-Gardiner C, Séguin G, Larrat S, Lair S, Landry F, Côté N. Oral rabies vaccination of raccoons and striped skunks with ONRAB® baits: multiple factors influence field immunogenicity. Journal of Wildlife Diseases. 2012 Oct 1;48(4):979-90.
76. Mainguy J, Fehlner-Gardiner C, Slate D, Rudd RJ. Oral rabies vaccination in raccoons: comparison of ONRAB® and RABORAL V-RG® vaccine-bait field performance in Québec, Canada and Vermont, USA. Journal of Wildlife Diseases. 2013 Jan 1;49(1):190-3.
77. Malcolm JR, Markham A. Global warming and terrestrial biodiversity decline. 2000.
78. Maslo B, Valent M, Gumbs JF, Frick WF. Conservation implications of ameliorating survival of little brown bats with white-nose syndrome. Ecological Applications. 2015 Oct;25(7):1832-40.
79. Miller W, Hayes VM, Ratan A, Petersen DC, Wittekindt NE, Miller J, Walenz B, Knight J, Qi J, Zhao F, Wang Q. Genetic diversity and population structure of the endangered marsupial Sarcophilus harrisii (Tasmanian

devil). Proceedings of the National Academy of Sciences. 2011 Jul 26;108(30):12348-53.
80. Mørk T, Prestrud P. Arctic rabies–a review. Acta Veterinaria Scandinavica. 2004 Mar;45(1):1.
81. Müller T, Rupprecht CE. Rabies and Wildlife Conservation. Wildlife Disease and Health in Conservation. 2023 Sep 15:1867.
82. Nadin-Davis SA, Muldoon F, Wandeler AI. Persistence of genetic variants of the arctic fox strain of Rabies virus in southern Ontario. Canadian journal of veterinary research. 2006 Jan;70(1):11.
83. Nadin-Davis SA, Sheen M, Wandeler AI. Recent emergence of the Arctic rabies virus lineage. Virus research. 2012 Jan 1; 163(1):352–62. https://doi.org/10.1016/j.virusres.2011.10.026 PMID: 22100340
84. Nettles VF, Shaddock JH, Sikes RK, Reyes CR. Rabies in translocated raccoons. American Journal of Public Health. 1979 Jun;69(6):601-2.
85. Newton EJ, Pond BA, Tinline RR, Middel K, Bélanger D, Rees EE. Differential impacts of vaccination on wildlife disease spread during epizootic and enzootic phases. Journal of Applied Ecology. 2019 Mar;56(3):526-36.
86. O'Keefe JM, Pettit JL, Loeb SC, Stiver WH. White-nose syndrome dramatically altered the summer bat assemblage in a temperate Southern Appalachian forest. Mammalian Biology. 2019 Sep;98:146-53.
87. Oswald P, Rodríguez A, Bourke J, Wagner N, de Buhr N, Buschmann H, von Köckritz-Blickwede M, Pröhl H. Locality, time and heterozygosity affect chytrid infection in yellow-bellied toads. Diseases of aquatic organisms. 2020 Dec 17;142:225-37.
88. Ottenhoff TH, Verreck FA, Hoeve MA, van de Vosse E. Control of human host immunity to mycobacteria. Tuberculosis. 2005 Jan 1;85(1–2):53–64. pmid:15687028
89. Pina-Martins F, Baptista J, Pappas Jr G, Paulo OS. New insights into adaptation and population structure of cork oak using genotyping by sequencing. Global Change Biology. 2019 Jan;25(1):337-50.
90. Reusch TB. Microsatellites reveal high population connectivity in eelgrass (Zostera marina) in two contrasting coastal areas. Limnology and Oceanography. 2002 Jan;47(1):78-85.
91. Reynolds DS, Shoemaker K, von Oettingen S, Najjar S, Veilleux JP, Moosman PR. Integrating Multiple Survey Techniques to Document a Shifting Bat Community in the Wake of White-Nose Syndrome. Journal of Fish and Wildlife Management. 2021 Dec 1;12(2):395-411.
92. Rose DC, Amano T, González-Varo JP, Mukherjee N, Robertson RJ, Simmons BI, Wauchope HS, Sutherland WJ. Calling for a new agenda for

conservation science to create evidence-informed policy. Biological Conservation. 2019 Oct 1;238:108222.
93. Rupprecht CE, Smith JS. Raccoon rabies: the re-emergence of an epizootic in a densely populated area. InSeminars in virology 1994 Apr 1 (Vol. 5, No. 2, pp. 155-164). Academic Press.
94. Sasse DB, Perry RW. Apparent annual survival of female eastern small-footed bats (Myotis leibii) roosting in Arkansas bridges. Journal of Mammalogy. Journal of Mammalogy, 104(6):1257–1263, 2023.
95. Savage AE, Zamudio KR. MHC genotypes associate with resistance to a frog-killing fungus. Proceedings of the National Academy of Sciences. 2011 Oct 4;108(40):16705–10. pmid:21949385
96. Schweizer, R. M., Robinson, J., Harrigan, R., Silva, P., Galverni, M., Musiani, M., ... & Wayne, R. K. (2016). Targeted capture and resequencing of 1040 genes reveal environmentally driven functional variation in grey wolves. Molecular ecology, 25(1), 357-379.
97. Smith D, O'Brien D, Hall J, Sergeant C, Brookes LM, Harrison XA, Garner TW, Jehle R. Challenging a host–pathogen paradigm: Susceptibility to chytridiomycosis is decoupled from genetic erosion. Journal of Evolutionary Biology. 2022 Apr 1;35(4):589-98.
98. Spurgin LG, Richardson DS. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. Proceedings of the Royal Society B: Biological Sciences. 2010 Jan 13;277(1684):979–88. pmid:20071384
99. Supple, M. A., & Shapiro, B. (2018). Conservation of biodiversity in the genomics era. Genome Biology, 19, 131. https://doi.org/10.1186/ s13059-018-1520-3
100. Tabel H, Corner AH, Webster WA, Casey CA. History and epizootiology of rabies in Canada. The Canadian veterinary journal. 1974 Oct;15(10):271.
101. Thompson JN. Evaluating the dynamics of coevolution among geographically structured populations. Ecology. 1997 Sep;78(6):1619-23.
102. Thompson JN. The coevolutionary process. University of Chicago press; 1994 Nov 15.
103. Thompson JN. The geographic mosaic of coevolution. University of Chicago Press; 2019 Dec 31.
104. Thuillet AC, Bataillon T, Poirier S, Santoni S, David JL. Estimation of long-term effective population sizes through the history of durum wheat using microsatellite data. Genetics. 2005 Mar 1;169(3):1589-99.
105. Trumbo DR, Hardy BM, Crockett HJ, Muths E, Forester BR, Cheek RG, Zimmerman SJ, Corey-Rivas S, Bailey LL, Funk WC. Conservation genomics of an endangered montane amphibian reveals low population structure, low genomic diversity and selection pressure from disease. Molecular Ecology. 2023 Dec;32(24):6777-95.

106. Verant, M. L., Meteyer, C. U., Speakman, J. R., Cryan, P. M., Lorch, J. M., & Blehert, D. S. (2014). White-nose syndrome initiates a cascade of physiologic disturbances in the hibernating bat host. BMC physiology, 14(1), 1-11.
107. Waples RS. Testing for Hardy–Weinberg proportions: have we lost the plot?. Journal of heredity. 2015 Jan 1;106(1):1-9.
108. Wright B, Morris K, Grueber CE, Willet CE, Gooley R, Hogg CJ, O'Meally D, Hamede R, Jones M, Wade C, Belov K. Development of a SNP-based assay for measuring genetic diversity in the Tasmanian devil insurance population. BMC genomics. 2015 Dec;16:1-1.
109. Yang J, Zhu GF, Jiang J, Xiang CL, Gao FL, Bao WD. Non-invasive genetic analysis indicates low population connectivity in vulnerable Chinese gorals: concerns for segregated population management. Zoological Research. 2019 Sep 9;40(5):439.
110. Yaremych SA, Warner RE, Mankin PC, Brawn JD, Raim A, Novak R. West Nile virus and high death rate in American crows. Emerging infectious diseases. 2004 Apr;10(4):709. pmid:15200865
111. Yıldırım Y, Tinnert J, Forsman A. Contrasting patterns of neutral and functional genetic diversity in stable and disturbed environments. Ecology and Evolution. 2018 Dec;8(23):12073-89.

CHAPTER 2

The role of a mechanistic host in maintaining arctic rabies variant distributions: Assessment of functional genetic diversity in Alaskan red fox (*Vulpes vulpes*)

Preface

A version of this chapter is published in *PLOS One* (2021),16(4): e0258975
https://doi.org/10.1371/journal.pone.0249176

Authors: Tristan M. Baecklund, Jaycee Morrison, Michael E. Donaldson, Karsten Hueffer, Christopher J. Kyle

Reference: Baecklund TM, Morrison J, Donaldson ME, Hueffer K, Kyle CJ (2021) The role of a mechanistic host in maintaining arctic rabies variant distributions: Assessment of functional genetic diversity in Alaskan red fox (*Vulpes vulpes*). PLoS ONE 16(4): e0249176. https://doi.org/10.1371/journal.pone.0249176

Modifications: Figures and tables were renumbered to accommodate the incorporation into this thesis. Spelling/grammatical errors were also made for consistency throughout the thesis. Appendix III supplements iHS and XP-EHH calculations and assessments using Selscan to compare and contrast against the results obtained using REHH in the body of Chapter 2. An updated version of Table S2.6 has been included in Appendix I, reflecting the addition of the chi-square test to the table, as observed in Chapter 3. No other aspect of the manuscript has been modified from the original.

Contributions: MED and CJK conceived and designed the study; TMB and JM performed lab work; TMB analyzed and visualized the data; TMB wrote the manuscript; TMB, MED, KH, and CJK critically reviewed the manuscript.

Abstract

Populations are exposed to different types and strains of pathogens across heterogeneous landscapes, where local interactions between host and pathogen may present reciprocal selective forces leading to correlated patterns of spatial genetic structure. Understanding these coevolutionary patterns provides insight into mechanisms of disease spread and maintenance. Arctic rabies (AR) is a lethal disease with viral variants that occupy distinct geographic distributions across North America and Europe. Red foxes (*Vulpes vulpes*) are a highly susceptible AR host, whose range overlaps both geographically distinct AR strains and regions where AR is absent. It is unclear if genetic structure exists among red fox populations relative to the presence/absence of AR or the spatial distribution of AR variants. Acquiring these data may enhance our understanding of the role of red fox in AR maintenance/spread and inform disease control strategies. Using a genotyping-by-sequencing assay targeting 116 genomic regions of immunogenetic relevance, we screened for sequence variation among red fox populations from Alaska and an outgroup from Ontario, including areas with different AR variants, and regions where the disease was absent. Presumed neutral SNP data from the assay found negligible levels of neutral genetic structure among Alaskan populations. The immunogenetically-associated data identified 30 outlier SNPs supporting weak to moderate genetic structure between regions with and without AR in Alaska. The outliers included SNPs with the potential to cause missense mutations within several toll-like receptor genes that have been associated with AR outcome. In

contrast, there was a lack of genetic structure between regions with different AR variants. Combined, we interpret these data to suggest red fox populations respond differently to the presence of AR, but not AR variants. This research increases our understanding of AR dynamics in the Arctic, where host/disease patterns are undergoing flux in a rapidly changing Arctic landscape, including the continued northward expansion of red fox into regions previously predominated by arctic foxes (*Vulpes lagopus*).

2.1 Introduction

Understanding patterns of local adaptation is important to not only enhance insights into how species interact with their environment, but also in clarifying how rapid changes in the suite of selective pressures influence population fitness and persistence [1,2]. When populations undergo divergent selection across heterogeneous landscapes, the potential for local adaptation exists [1,3,4]. The process of local adaptation is not only influenced by selective pressures but also the interplay of gene flow and effective population size (genetic drift) relative to the strength of the selective pressure. Gene flow and genetic drift undermine a population's ability to locally adapt through either the homogenization of genetic variation or through the random loss of genetic variants in small populations [1,5]. However, if selection is both divergent in nature, and stronger than the combined force of gene flow and genetic drift, local adaptation is likely to occur [1].

Quantitative evidence for local adaptation can be difficult to decipher in natural populations and confounded by phenotypic plasticity that allows for the

expression of multiple phenotypes from a single genotype [6, 7]. As such, and short of common garden experiments which can be difficult to undertake in natural systems, genetic assessments of the interactions between selection and the demographic forces acting on different populations provide a means to detect patterns indicative of local adaptation. While challenges do exist, genetic signals of locally adapted populations have been identified across a wide range of systems including nonsynonymous gene changes among wolf populations that correlate with precipitation and vegetation patterns [8], and variation in salmonid immune genes associated with the thermal regimes of different waterbodies [9].

Among the array of selective forces that populations are exposed to, infectious diseases often present strong selective pressures capable of enacting rapid and marked population changes as demonstrated by: 1) white-nose syndrome, where disease emergence decimated populations of several species of bats [10,11]; 2) West Nile virus, where over two thirds of crows initially succumbed to the disease [12]; and 3) facial tumors in Tasmanian devils, where populations have experienced 90% declines [13]. In these systems, strong selective sweeps reshaped the genetic diversity of populations through the increased frequency of adaptive traits and a subsequent decrease in the frequency of maladaptive traits [14,15]. The importance of host adaptation in response to disease is further exemplified by chronic wasting disease (CWD) in mule deer, where a genotypic difference conveys resistance to CWD [16]. However, host response to disease is not solely based on the host's genotype, but also the genetic variants of the disease(s) to which they are exposed,

further complicating a population's interaction with disease based on these coevolutionary patterns [17].

Genetic assessments of local adaptation in response to infectious disease have typically focused on highly polymorphic regions of the major histocompatibility complex (MHC) given associations with antigen binding and overall population health [18–22]. Historically, these studies limit their assessments to a few, if not a single, region (e.g., DRB exon-2) [23–25]. These studies have provided reasonable assessments of the spatial genetic structure that can arise from local adaptation, yet the immune system is complex and includes adaptive-, innate-, and intrinsic immunity aspects such that more holistic analyses are required to understand immunogenetic interactions with disease. Several studies have started to use genotyping-by-sequencing (GBS) to explore larger subsets of genetic variation relative to surrounding selective pressures, including disease [e.g., 26–30]. Specifically, Miller et al. [13] developed several GBS arrays, targeting mtDNA and nuclear SNPs, where population structure was indicative of differential responses to the facial tumour disease plaguing Tasmanian devil populations [31].

Rabies is a *Lyssavirus* with several strains and subvariants that are normally maintained by a primary/maintenance mammalian host within the orders of Carnivora (e.g., foxes, coyotes, wolves, skunks, and raccoons) and Chiroptera (bats) [32–34]. Rabies is of concern given high mortality rates associated with this disease, the potential for primary hosts to infect domesticated animals and occasionally humans [35], and the fact that the disease can spill over into other reservoir hosts

when epizootic [36]. The arctic rabies (AR) strain has a circumpolar distribution that largely coincides with the distribution of its primary host, arctic fox (*Vulpes lagopus*). AR consists of four main viral variants occupying distinct geographical distributions including: variant 2 confined to the Seward Peninsula of Alaska, variant 4 found in southwestern portions of Alaska, and variant 3 which occurs throughout the northern coasts of North America and parts of northern Eurasia [33,34,37] (Fig 2.1). In contrast, AR variant 1 is isolated to southern Ontario [37], where arctic fox populations are absent and red fox (*Vulpes vulpes*) are presumed to be the maintenance host [38].

In Alaska, three AR variants (2, 3, and 4) circulate in geographically discrete regions where arctic fox and red fox are sympatric in coastal regions. While AR infects both arctic and red foxes, AR is largely absent from the interior of Alaska where only red fox exist [37]. The geographically distinct distribution patterns of AR presence/absence and AR strains in Alaska have led to questions with regards to: i) how this disease is maintained, ii) are there coevolutionary relationships between AR and its hosts that might be expected if patterns of local adaptation exist to explain the geographic restriction of AR variants, and iii) if arctic fox, red fox, or both, serve as maintenance hosts for AR in Alaska. Determining the host status of these fox species is pertinent in that maintenance hosts play different roles in long-term disease maintenance and spread, relative to spillover/reservoir hosts that do not perpetuate disease on broader timescales [36,39,40]. Categorizing disease hosts as either maintenance hosts or spillover hosts can be determined by variable

prevalence rates of disease in different host species, as seen with avian influenza in species of waterfowl [41]. However, these correlations sometimes do not provide a holistic understanding of underlying disease dynamics, as in the case of red and arctic fox populations in Alaska, where AR is commonly found in both species; therefore, these patterns may not distinctly differentiate their relative roles with respect to AR [37,41]. Previous research has attempted to use the host genetic structure of both red and arctic fox populations in Alaska, in context of AR strains, to assess the influence of these two hosts on AR disease dynamics [37]. Goldsmith et al. [37] found that population genetic structure of arctic fox correlated with the distribution of the three AR strains as would be expected of a maintenance host and long-term co-evolutionary patterns with AR. The role of red fox was less clear, and data within Goldsmith et al. [37] did not exclude red fox as a maintenance host. Goldsmith et al. [37] did find support for genetic structure among the sampled regions for red fox, where coastal tundra populations clustered together separately from those in the boreal interior [37], aligning with the geographical presence/absence of AR in Alaska. Goldsmith et al. [37] also found evidence of fine-scale geographically isolated genetic clusters, but the levels of admixture among these clusters undermined correlations of the host genetic structure of red fox with AR strains. These findings were not surprising given red fox are widely dispersed carnivores, native to much of the northern hemisphere as a matter of their high dispersal capabilities, their generalist nature, the capacity of the species to exhibit a phenotypic plasticity in response to changes in selective pressures, and historical

red fox translocations [42–47]. The lack of correlation of AR strains with low levels of red fox population genetic structure, may also be related to observations that red foxes have expanded their distribution northward, coinciding with Arctic warming, a factor postulated to continue to influence and alter AR dynamics [37,48–50].

While the neutral genetic structure of host species can be used to infer relationships with disease maintenance and spread [37,51,52], understanding deeper co-evolutionary patterns between hosts and pathogens requires insight into the variation that exists in genes that interact with the reciprocal selective pressures. To this end and building on the data from Goldsmith et al. [37], Donaldson et al. [53] developed a GBS assay to specifically target 116 immunogenetically relevant regions of the red fox genome. Donaldson et al. [53], tested the assay on a small sample size of red foxes from regions with different AR variants and regions without rabies and found 15 F_{ST}-based outlier SNPs that divided the samples into two genetic clusters corresponding to regions with and without AR, similar to the results by Goldsmith et al. [37]. In both studies, inferences on the relationship between host genetic clustering and rabies distributions were undermined by either small sample sizes or assessments on regions of the genome unlikely to show patterns of selection from infectious diseases. However, in these studies, red fox genetic structure was more pronounced in the data from the immunogenetic assay [53] relative to the microsatellite data [37], suggesting that gene flow was not solely responsible for the observed patterns of genetic structure.

Herein, we build on the work of Goldsmith et al. [37] and Donaldson et al. [53], using the same immunogenetic GBS assay, by increasing the number of red foxes sampled per location to better assess frequency differences of genetic variants among the sampled locations and to gain further insight into the role of red fox in AR maintenance in Alaska. We also aimed to put the interrelationships of AR and red foxes in Alaska into context by including red foxes from Ontario (Canada) as a potential outgroup. Of interest is the contrast between central Alaska and Ontario, as AR variant 1 is solely in Ontario and is maintained without the presence of arctic fox, where it is detected in red fox and skunk populations [54–56]. It is unclear how the distributions of AR variants are maintained, nor how rapid climate changes occurring in the Arctic may influence rabies disease dynamics such as through the continued northward expansion of red fox (a highly susceptible AR host) into ranges previously predominated by AR's natural host, the arctic fox [49,50,57–61]. We also aimed to further explore the data from the immunogenetic assay developed by Donaldson et al. [53], by using the well-annotated canine reference genome for enhanced assessments of SNP/gene associations, and by implementing additional SNP outlier tests to account for the inter-variability between methods. We hypothesized that red fox population genetic structure has been shaped by AR variants in Alaska despite high dispersal capability and gene flow found within the species. Therefore, we predicted that immunogenetically relevant genomic regions would demonstrate large shifts in allele frequencies indicative of genetic structure and local adaptation associated with the distribution of AR in Alaska; consistent

with the previous findings [53]. This research aims to increase our understanding of how AR is maintained in Alaska, the role of red fox as either a maintenance or spillover host of AR, and the potential role red fox may play as the Arctic continues to experience rapid warming trends which may affect the distribution of AR hosts and its variants.

2.2 Methods

2.2.1 Sampling, DNA extraction and quantification

Previously collected red fox muscle and spleen tissue samples from Alaska, originally obtained from a variety of independent trappers and organizations, were provided by the University of Alaska Museum of the North (Table S2.1). Red fox muscle tissue samples from Ontario (Canada) were obtained from the Ministry of Natural Resources and Forestry (Table S2.1). This study required no animal handling or direct sampling from animals (all samples were previously collected), as such animal care approval was not required. Tissue samples were stored in a -80°C freezer and DNA extraction was performed using the DNeasy Blood and Tissue Kit (Qiagen; File S2.1). We quantified DNA extractions using the Quant-iT PicoGreen dsDNA Assay Kit (ThermoFisher Scientific). Extracted DNA quality was assessed by ethidium bromide stained 0.8% agarose gel electrophoresis (90 V for 45 minutes) using the HighRanger 1 kbp DNA ladder (300 bp– 10,000 bp; Norgen Biotek) as a reference. After these quantity/quality assessments, a final set of 96 high molecular weight DNA samples suitable for sequencing were processed from four regions

across North America: Southwestern Alaska (n = 25), Seward Peninsula (n = 21), Central (South/Interior) Alaska (n = 30), and Renfrew County in Ontario (n = 20).

2.2.2 Library preparation, sequence capture and high-throughput sequencing

DNA libraries were prepared using the Kapa HyperPlus Kit (Roche) following the SeqCap-EZ HyperCap UGuide v1.0 (Roche) protocol with several modifications to the workflow (File S2.1). Pre-capture LM-PCR library quality was assessed utilizing an ethidium bromide stained 2% agarose gel electrophoresis (100 V for 45 minutes).

Equal-molar amounts of each library were combined to form a 1 µg DNA multiplex of the 96 libraries. Target enrichment was performed using the SeqCap EZ Developer Library probe set previously described by our lab [53]. Modifications to the target enrichment included: 2 µl xGen Universal Blockers—TS Mix (Integrated DNA Technologies) instead of the NimbleGen Multiplex Hybridization Enhancing Oligo Pool (Roche), NimbleGen SeqCap EZ Developer Reagent (Roche) was used instead of NimbleGen COT Human DNA (Roche) during hybridization sample preparation, and the hybridization was carried out at 47°C for 20 hours. The target-enriched multiplex was assessed on a bioanalyzer and sequenced on an Illumina MiSeq v3 run using 2x300 bp reads (Advanced Analysis Centre Genomics Facility, University of Guelph). We also obtained previously sequenced data for 29 individuals (Table S2.1) [53] from the NCBI Sequence Read Archive (SRP119314) [51] and included that data in subsequent analyses.

2.2.3 Sequence alignment and variant annotation

Paired-end reads from the 96 newly sequenced individuals, and the 29 previously sequenced samples (total of 125 libraries) were aligned to the canine reference genome (Southwestern Alaska n = 32; Seward Peninsula n = 33; South/Interior Alaska n = 40; Ontario n = 20; Fig 2.1; Table S2.1), utilizing the bwa-mem command in Burrows-Wheeler Aligner v0.7.12 [62]. Sequence alignment metrics were compiled using SAMTOOLS v1.5 [63]. We utilized the Genome Analysis Toolkit (GATK, v4.0.0.0) best practices pipeline and standard hard filtering parameters to perform duplicate sequence removal, SNP/INDEL variant annotation, genotyping, and variant recalibration [64–66]. After these steps, the GATK function SelectVariants was used to filter the obtained VCF file to only contain bi-allelic SNPs.

The original SeqCap EZ Developer Library probe was designed based on limited sequence information from a draft version of the red fox genome [53]; therefore, the positions of all probe targeted areas were first identified within the canine reference genome via BLASTn (Canfam 3.1; Table S2.2) which were then complied into a list of on-target intervals. Using these intervals, SNP variants were further categorized as being within coding regions or in intergenic regions (outside coding regions).

Throughout our analyses we addressed several recommendations outlined by other researchers [67] when attempting to identify loci under selection using F_{ST} outlier tests. Specifically, we accounted for possible linkage disequilibrium within

the datasets, we implemented a filter for minimum allele frequency (MAF), and we implemented the use of multiple outlier tests in our analyses.

2.2.4 Analyses of SNPs in intergenic regions

The sub dataset containing SNPs in intergenic regions was filtered using VCFtools, v0.1.13, to retain only biallelic variants with a MAF threshold of 2%, and a maximum missing genotype threshold (per site) of 20%. The remaining SNPs were analyzed using the Ensembl Variant Effect Predictor (VEP) tool [68,69] to remove any variants that were within 20 kbp from any known transcribed region (protein coding RNA or non-coding RNA) within the reference canine genome. Additionally, these SNPs were pruned for linkage disequilibrium as implemented by the SNPRelate package in R v3.5 [70], and further filtered for physical linkage (only SNPs that were ≥ 100 kbp from one another were retained). Variants that fulfilled these parameters were assumed to not be under selective pressure and were used to assess patterns of neutral population genetic structure. These filtering steps, and subsequent analyses, were also performed on a SNP dataset from intergenic regions that did not include red fox samples from Ontario to test for substructure within Alaska.

Principle component analysis (PCA) and discriminant analyses of principle components (DAPC) were performed in RStudio using the adegenet (v 2.1.1) [71] and ape (v 5.1) [72] packages. The components retained for PCA were those with eigenvalues ≥ 0.1 and cross validation were used to determine the number of retained components based on the root mean squared error (lowest MSE). The

optimal number of clusters identified through the data for the DAPC was determined using successive K-means.

Utilizing STRAUTO (v. 1.0) [73], we ran STRUCTURE over several processors concurrently. STRUCTURE analyses were performed using a burn-in length of 50,000 followed by 200,000 iterations for K = 1 through K = 6, with 20 iterations of each K. The ΔK statistic was calculated to determine the number of distinct genetic clusters that were inferred using structure harvester web v0.6.94 [74]. Utilizing CLUMPP 1.1.2 [75], and the LargeKGreedy algorithm (10,000 repeats) individuals were assigned to genetic clusters, the STRUCTURE analyses were combined and visualized using DISTRUCT v1.1 [76].

Power analyses on the presumed neutral SNP datasets were performed using POWSIM v. 4.1 [77]. Simulations were run with N_E = 500 and 5,000, over (t) generations = 0, 10, 100, 500, 1,000. 1,000 iterations were performed for each set of conditions, and each run sought to differentiate between the four sampled regions by evaluating the power to detect genetic homogeneity through chi-square and Fisher's exact tests. The Fisher's exact test was implemented within the program using a Monte Carlo Markov chain approach with the default parameters of 1,000 burn-ins, 100 batches, and 1,000 iterations.

2.2.5 Analyses of SNPs in coding regions

The sub dataset containing SNPs in coding regions was filtered using VCFtools to retain only biallelic variants with a MAF threshold of 2%, and a maximum missing genotype threshold (per site) of 20%. Outlier testing was performed on this sub

dataset using four different packages: PCAdapt [78], OutFLANK [79], Arlequin [80], and Bayescan [81]. Each of these tests identified SNP F_{ST} outliers using an adjusted p-value threshold of ≤ 0.05; detailed parameters implemented for each method are provided as supplement material (File S2.1). Outlier tests use different sets of assumptions and caveats, often leading to inconsistencies across packages [79], so we retained any outlier identified by at least one test and compiled these SNPs into a separate VCF file using VCFtools. That dataset was pruned for linkage disequilibrium using the SNPRelate package in R and further filtered for physical linkage by only keeping SNPs > 100 kbp from one another. That sub dataset containing filtered LD-pruned outlier SNPs from coding regions was used to assess the distribution of these variants across our sample design, achieved through the PCA, DAPCs and STRUCTURE analyses described above. Outlier analyses of SNPs from coding regions were also performed with a dataset that did not include red fox samples from Ontario to test for substructure within Alaska.

2.2.6 Signatures of selection; iHS, XP-EHH, pN/pS

A combined dataset of filtered (MAF and max-missing data) SNPs from both intergenic and coding regions were assessed to determine the integrated haplotype homozygosity score (iHS) and the cross-population extended haplotype homozygosity (XP-EHH) using the REHH package in RStudio (ihh2ihs and ies2xpehh functions respectively) [82,83]. Normalization of p-values for both iHS and XP-EHH are incorporated as part of the REHH package. Normalization is achieved following Gautier and Naves, where p-values are generated through the -log of the Gaussian

cumulative distribution function for each statistic [84]. These metrics facilitate a comparison of the integrated extended haplotype homozygosity within a population (iHS) and between populations (XP-EHH) [85,86].

Estimations of the relative ratio of nonsynonymous substitutions to synonymous substitutions (pN/pS ratio) were determined using the output of SnpEff (using the CanFam3.1.99 database) which annotated each SNP within coding regions as synonymous or non-synonymous polymorphisms [87]. Values were calculated per gene following Nei and Gojobori; $pS = \frac{Sd}{s}$ and $pN = \frac{Nd}{N}$ [88]. Where S and N are the number of synonymous and nonsynonymous sites and Sd and Nd are the total number of synonymous and nonsynonymous polymorphisms [88–90]. Per gene, DnaSP v6 was used to estimate the number of potential nonsynonymous (N) and synonymous sites (S) using the coding sequence for each gene [91]. Ratios > 1 can be indicative of positive selection, whereas ratios < 1 typically infer purifying selection [89].

2.3 Results

2.3.1 Raw sequence data

The combined dataset (N = 125) of newly (n = 96) and previously (n = 29) sequenced samples had an average of ~315,000 raw reads per library, of which 96.2% mapped to the canine reference genome, ~11.6% reads were filtered per library, and ~58% aligned to targeted regions (~ 65X coverage; Table S2.3).

2.3.2 SNPs in intergenic (off-target) regions

A dataset of 4,811,979 off-target SNPs was filtered to exclude SNPs < 100 kbp from a coding region and pruned to minimize linkage disequilibrium. This yielded a sub-dataset of 43 SNPs in intergenic regions with an average depth of coverage of 25X that were presumed not to be under selective pressure (Table S2.4). We visualized these data using PCA, DAPC and STRUCTURE. Power analyses of the 43 SNPs in intergenic regions indicated a power of ~98% to detect structure at an expected F_{ST} = 0.01 and a power of 100% to detect structure at an expected F_{ST} > 0.05, indicating a high likelihood that if there was population differentiation > 1%, the present dataset had the power to detect it. DAPC and STRUCTURE identified K = 2 as the most likely number of clusters with high levels of gene flow across all analyses (Fig 2.2). Specifically, Ontario and Alaska formed two distinct genetic clusters (pairwise F_{ST} = 0.035 among genetic clusters), with no clear patterns of genetic structure within the sampled Alaskan regions (Fig 2.2). Analyses containing only the Alaskan samples provided results similar to those obtained when we analyzed the subset of data containing red fox populations from both Alaska and Ontario; where PCA and DAPC analyses identified no population genetic structure within Alaska, however STRUCTURE results were suggestive of weak patterns of substructure (Fig S2.1). This subset of data contained 123 SNPs with an average coverage per site of ~30X, and an F_{ST} = 0.012 among Alaskan red fox genetic clusters (Table S2.4).

2.3.3 SNPs in protein-coding (on-target) regions

We detected 9,650 SNPs located within on-target intervals. After applying a MAF threshold = 2% a maximum missing data threshold = 20%, and discarding SNPs on the X-chromosome, 2,094 SNPs remained. Only Arlequin and PCAdapt identified outliers, producing a combined sub-dataset of 131 SNPs (Fig S2.2). We filtered these 131 SNPs for linkage disequilibrium, producing a sub-dataset of 30 outlier SNPs in protein-coding regions. The majority of these SNPs were found within interleukin, toll-like receptor, and MHC gene families (Table S2.5). Two of these SNPs, associated with TLR4 and IL12RB1 genes, were predicted to cause missense mutations that could alter the putative chemical characteristic of the substituent group (Table S2.5). We also noted an additional 16 SNPs (from the 131 SNP dataset) removed during linkage disequilibrium filtering were also predicted to cause a missense mutation, potentially altering the underlying chemical property of the respective amino acid. Notably, 12 of these 16 SNPs were found within TLR5 (Table S2.5). DAPC and STRUCTURE analyses of the 30 outlier SNPs in protein-coding regions identified K = 2 clusters, which was similar to the results obtained from the off-target SNP analysis, with detectable genetic structure between Alaskan and Ontario foxes (Fig 2.3). Pairwise assessments found F_{ST} = 0.135 among genetic clusters. Processing these data, obtained from Alaskan foxes exclusively, identified a larger number of outlier SNPs than the dataset that included Ontario foxes (221 versus 131, respectively). Of these 221 outliers, 22 resulted in a missense mutation with the potential to change protein function and one SNP resulted in a premature

stop codon (Table S2.5). We filtered the 221 SNPs for linkage disequilibrium and only four of the 22 SNPs with the potential to cause a missense mutation were retained in the sub-dataset that contained 16, filtered and linkage disequilibrium pruned, outlier SNPs (Table S2.5). STRUCTURE analyses without the Ontario red fox outgroup suggest weak substructure exists within Alaska, contrasting with both PCA and DAPC analyses that did not reveal distinct population genetic clusters (Fig 4). Pairwise assessments estimated F_{ST} = 0.090 among genetic clusters. Genome-wide pairwise F_{ST} of Alaskan red fox are provided as supplementary data (Fig S2.5).

2.3.4 Signatures of selection; iHS, XP-EHH, pN/pS

Plotting the p-value of iHS demonstrated few signals of selection for each population. Further analyses of the iHS scores identified one, two, and five outlier candidate regions for the Seward Peninsula, Southwest and Central red fox populations, respectively (Fig S2.3). Similarly, XP-EHH analyses demonstrated weak signals of selection between the populations, with the exception of chromosome 18, where the Southwest and Seward Peninsula populations appear to have much closer affinities relative to their comparisons with the Central red fox population (Fig S2.4).

From the SnpEff output, we were able to identify missense and synonymous substitutions at 85 of our targeted genes, however, pN/pS calculations were unable to be calculated for any of the three Alaskan red fox populations at 23 of these genes because there were no synonymous and/or missense substitutions at the locations in their respective datasets resulting in division by zero errors or pN values equal to

zero (Table S2.6). While the majority of the pN/pS ratios were indicative of purifying selection, six genes had pN/pS ratios > 1 in at least one of the populations, suggestive of positive selection (Table S2.6). Two of these genes, CCL5 and TLR4 only appeared to be under selection (pN/pS > 1) in the central Alaskan red fox population. The remaining four genes (DLA-12, DLA-88, DLA-DMB, and DLA-DQBC1), appeared to be under selection in all three of the populations.

2.4 Discussion

In this study, we sequenced immunogenetically associated regions of the red fox genome using a GBS assay in context of both variants of arctic rabies (AR) and the presence/absence of the disease. The goal was to further understand the role of red fox in AR maintenance/spread. The GBS assay generated both off- and on-target data. Analysis of these data provided relative assessments of genetic structure that could be attributed to gene flow or associated with local selective pressures. While additional samples, outgroups, and analyses (including the use of a different reference genome) were used in this study, our results were largely consistent with previous investigations [37,53], finding subtle genetic structure between regions with and without the presence of AR, but no evidence of differential selection associated with unique AR variants among the red fox populations. The latter finding was somewhat unexpected, as it is not clear how distinct AR lineages maintain stable geographic distributions if there is indeed extensive gene flow among red foxes, although some evidence implicates the primary AR host, arctic fox, in maintaining these distributions [37]. These data may also suggest that dispersal is

more limited among infected red foxes, yet latency of clinical AR symptoms can be weeks and sometimes months, suggesting that limited host dispersal may not be a factor [55,92]. That said, there is no indication that different AR strains show any differences in infectivity or pathogenicity, so not having distinct immunogenetic structure correlated with different AR strains in regions with the disease may then be less surprising. In contrast, the observed levels of neutral genetic structure from the off-target SNP analysis in this study, and previous research [37], are suggestive of high levels of gene flow within all sampled Alaskan regions. Based on the contrasting patterns of neutral and immunogenetic structure, we take these data to suggest that there is a subtle immunogenic response, and potentially a locally adaptive response to AR in red fox populations in Alaska. Further, of the outlier SNPs, several are associated with interleukins and toll-like receptors known to mediate responses to rabies infections [93–97]. Although these data provide further insight into potential mechanisms that control rabies maintenance and spread, additional analyses, such as those resulting from full genome sequencing could reveal different patterns of responses to disease to extend beyond those immunogenetically associated regions.

2.4.1 Use of off-target data

Targeted sequencing approaches, despite aiming to enrich for certain data, consistently generate undesired (off-target) reads [98]. Explorations of the usefulness of these previously discarded data suggest they possess adequate sequencing coverage and quality for downstream analyses [98–101]. Combining

these data with simulation-based assessments, provides a measure of the analytical power of the dataset, and enables confidence to be placed in the interpretation of these data [102–104]. The off-target datasets analyzed herein provide another example of the purposing of otherwise undesirable reads to discern presumed neutral genetic structure as a baseline for the on-target data. Previous red fox research using microsatellite and mtDNA markers [37] found weak genetic clusters that distinguished between the interior of Alaska and coastal regions (F_{ST} = 0.035 among populations), where both clusters displayed extensive admixture with foxes from the northern coast. Our analyses including Ontario red fox, based on presumed neutral off-target SNPs, found no evidence of genetic structuring within Alaska. This contrast in observed structure between studies might reflect the additional power of microsatellite markers, with many alleles per locus, to detect structure, or perhaps the benefit of using certain analytical assumptions which increase the likelihood of identifying subtle patterns of genetic structure as implemented in the work by Goldsmith et al. [37].

The lack of substructure within Alaska using the off-target data, when including Ontario red fox, may be due to stark differences in ancestry between those fox populations from Alaska and Ontario [105–107]. North American red fox occupied several glacial refugia during the last glacial maxima, currently recognized as a Holarctic clade (Western Canada, Alaska, Asia, and European origins) [106] and Nearctic clade (Eastern and Central Canada, the Rocky Mountains, and several montane regions throughout the US) [106]. The recent expansion of red foxes into

previously unsuitable habitats in the Arctic was thought to be due to the introduction of non-native red foxes with European origins, expanding from the Eastern coast of North America across Central US and Canada [105]. However, recent research has shown that expanding red fox populations in the North American arctic tundra are more closely related to the native, boreal, red fox [107]. Therefore, the lack of substructure observed in the off-target SNP dataset may in part be due to the prominent difference in ancestry between fox populations in Alaska and Ontario, thus masking signatures of differentiation between Alaskan populations. The lack of structure observed could also reflect recent population expansions of this species in Alaska, given the potential for genetically similar founders. Off-target analyses of only Alaskan red fox, suggested gene flow exists among the sampled regions despite physiographic features (i.e., mountain ranges) that might be expected to retard gene flow and supports the supposition that the inclusion of the Ontario outgroup of foxes appeared to mask weak signatures of genetic structure.

2.4.2 Analysis of SNPs in protein-coding regions

2.4.2.1 Interrelationship between AR and red fox across North America.

Methods that identify outliers are prone to varying amounts of type I and II error, which potentially result in inconsistent results between different tests [108]. Analyses of the on-target SNP dataset found only two (Arlequin and PCAdapt) of the four methods identified outliers. While both tests identified a similar number of SNPs, only seven of 131 SNPs were common between the two tests. Most SNPs

identified result in synonymous changes (n = 113 SNPs) and are unlikely to alter resulting protein structure and/or function [109]. Of the 18 identified non-synonymous SNPs, a large proportion were associated with TLR5 (n = 12 SNPs), but only the SNPs associated with TLR4 and IL12RB1 were retained in the sub-dataset given our filtering parameters (Table S2.5). TLR4 and TLR5 are associated with initiating an inflammatory response by recognizing molecular structures indicative of bacterial infiltration [110–113] whereas IL12RB1 encodes for the transmembrane protein responsible for regulating the response of both IL12 and IL23 [114]. Thus, these genes may have an important role in the species response to disease, especially in the context of rabies, as these gene families have previously been implicated in rabies resistance [93–96]. We also detected outlier SNPs within the protein-coding regions of MHC, interleukins, and toll-like receptors (Table S2.5). Genetic structure was detected when analyzing the 30 outlier SNPs in protein-coding regions between red fox populations in Alaska and Ontario (F_{ST} = 0.1353). This pattern, relative to the observed weak patterns from the putatively neutral data, is suggestive of local adaptation. The pattern of genetic structure between Alaska and Ontario red fox is interesting in the context of rabies where the AR variant circulating in Southern Ontario (AR variant 1), that is absent from the Arctic, persists without requiring reintroduction from arctic fox, and is phylogenetically distant from the other variants circulating in Alaska [33,34]. While red foxes were initially implicated in circulating the unique AR variant in Ontario, declining observations of AR variant 1 among red foxes, with subsequent increases among skunk populations,

led researchers to test the possibility of a host shift from foxes to skunks [115]. Nadin-Davis and Fehlner-Gardiner found that the variant had accumulated codon changes that coincide with the typical strain of rabies that is found in skunks supporting that this rabies variant may be shifting hosts [115]. The research presented here, combined with this previous work [115], suggest that AR variant 1 may have an increased capacity to locally adapt given the stark differences of the immune response between Ontario red fox populations and red fox populations in Alaska. These differences are made evident from the observed genetic structure (Fig 2.3), and the potential host shift that is occurring into skunk populations in Ontario.

We acknowledge that correlation does not equal causation, but the identified outliers associated with TLRs and interleukins, coupled with the unique distribution of arctic rabies in Alaska, point towards red fox populations locally adapting to the presence/absence of the disease, but not specific AR variants. Supported by past research demonstrating the involvement of these gene families in rabies mechanisms [93–97], the candidate genes identified herein provide an opportunity to further explore this potential coevolutionary relationship on a larger scale.

2.4.2.2 Interrelationship between AR and red fox in Alaska

When restricting on-target SNP analyses to include only Alaskan red foxes, three methods identified outliers (Arlequin, PCAdapt, and OutFLANK). Six of the outliers were identified by at least two programs, demonstrating the variability of program algorithms/assumptions to detect SNP outliers [108]. The outlier SNPs detected by multiple programs were associated with the protein-coding regions of toll-like

receptors (TLR5, TLR6) chemokine (C-C motif) ligand 2 (CCL2); however, only the SNPs associated with TLR6 and CCL2 were retained in the filtered sub-dataset. TLR6 forms a dimer with TLR2 associated with detecting gram-positive bacteria (Table S2.2) [116]. CCL2 is associated with the adaptive immune response by influencing monocyte activity during the inflammatory response (Table S2.2) [117]. A large number of the missense variants identified with the potential to change the underlying protein function were associated with the major histocompatibility complex (MHC) and different TLRs. The single nonsense mutation was associated with the protein-coding region of MHC locus DLA-DRB1 that is important in antigen binding. Thus, the respective protein may not be translated properly, which could lead to the loss of antigen recognition and potentially negatively impact the ability of the population to mount an immune response [24].

Estimates of pN/pS ratios of 62 genes offered interesting insights into genes that may be under positive selection between rabies absent areas and those areas where different arctic rabies variants circulate. Specifically, TLR4 appeared to be under positive selection only in the central interior Alaskan red fox population, where rabies is not endemic, indicating that this gene could play a role in preventing the virus from spreading into this region. Further, the four genes under selection in all 3 of the populations (DLA-12, DLA-88, DLA-DMB and DLA-DQBC1) are all components of the MHC. The MHC region plays important roles in antigen recognition and variation within this group of genes is often associated with healthy populations [18–22], as such, it is not surprising to see several MHC gene members

under positive selection. Despite these findings, it is important to note that some of the genes studied herein had very few synonymous/nonsynonymous sites which can inflate resulting pN/pS ratios and ultimately lead to potential biasing of these results [89].

We identified genetic structure between red fox from coastal Alaska, where rabies is endemic, and the central interior, where rabies is absent. There was no genetic structure observed in the context of different AR variants encompassed among the sampled regions consistent with previous studies [53]. Despite small sample sizes, a lack of linkage disequilibrium pruning, and the implementation of different outlier testing programs (i.e., LOSITAN), the data presented herein remain consistent with the findings of Donaldson et al. [53]. Additionally, both analyses presented in the current study and those of Donaldson et al. [53] identified SNPs in the protein-coding regions of C3 and ITGAM as outliers. C3 encodes for a protein of the same name, whose derivatives contribute to phagocytosis, an inflammatory response, and relaying signals to T cell-dependant antigens (Table S2.2) [118]. ITGAM encodes for integrin αM, one of two proteins that bind together to form macrophage antigen 1 which is involved in leukocyte adhesion and migration (Table S2.2) [119]. Since SNPs associated with these protein-coding regions have continuously been identified as outliers across studies using several different methods, there is increased support that these genes may be under selective pressure and warrants further investigation.

2.4.3 Maintenance of arctic rabies variant distributions in Alaska

It has been questioned whether the unique distributions of artic rabies variants in Alaska are influenced by the natural host (arctic fox) and maintained by the red fox. Using microsatellite marker data, Goldsmith et al. [37] demonstrated that the neutral genetic structure of arctic fox appeared consistent with the relative distributions of the rabies variants in the state of Alaska. Goldsmith et al. [37], also found that the neutral genetic structure of red foxes only distinguished between those foxes from rabies-endemic coastal regions and the rabies-absent interior of Alaska [37]. Together, these data were taken to suggest the distributions of AR variants are maintained solely by arctic foxes. Data presented in our current study further demonstrate that while the red fox may be a maintenance host for AR in Alaska, the species demonstrates no patterns of genetic structure correlating to the distributions of specific AR variants. In contrast to previous work, however, we find that red fox populations appear to exhibit a weak signature of local adaptation to the presence/absence of AR, as demonstrated by the genetic structure analyses of immunogenetically relevant loci relative to presumed neutral loci. Furthermore, TLR4 appears to be under positive selection within only the Central red fox population where rabies is not endemic. This gene family has previously been associated with rabies disease mechanisms [96], and the geographic structure of variants of these genes present a potential explanation as to why rabies is not able to reach an endemic status in the central interior of Alaska, although further research would be required to investigate this hypothesis in depth. Overall, these

findings are important in context of a warming Arctic, as they suggest that the Alaskan distribution of AR is likely to be unaffected by continued northward expansions of red fox, but rather by a northward retreat of arctic fox from the southern edge of their distribution in the state [50].

2.4.4 Future steps

Our data suggest red fox populations in Alaska have not undergone differential selection in response to different AR variants based on their unique distributions. It remains plausible, however, that this observation is due to a plastic response of the immune system or that there may be differences in the up/down-regulation of specific genes—hypotheses beyond the scope of the current study. Research utilizing RNA-seq to identify differences in gene expression among foxes exposed to the varying AR variants could have the potential to address some of these alternative hypotheses [120]. Additionally, while the phylogenetic relationship of the AR variants is well documented [33,34], our understanding of whether these genetic differences correspond to underlying differences in pathogenicity or virulence remain unknown. The observed spatial segregation of the AR variants may be caused by founder events with no subsequent gene flow; however, this seems unlikely given the gene flow present within two of the disease's main hosts in Alaska, red and arctic fox, that would homogenize AR variants if they did not have selective differences. Previous research indicates that only arctic fox populations appear to influence the distinct distribution of AR variants [37,53]. This finding, combined with data presented here, suggests selective differences between AR variants do not

exist for Alaskan red fox populations. Research should further explore this phenomenon in the natural host of AR, arctic fox, to provide an assessment of genetic differences suggestive of selective differences between AR variants. Given that historical records largely reflect arctic fox have been exposed to AR much longer than red foxes, the likelihood for coevolutionary forces that would result in patterns of differential selection is much greater [55]. Therefore, if such coevolutionary patterns were detected, it would be indicative of differential selection of the different AR variants on arctic fox populations.

2.5 Conclusions

Key to comprehending the ecology and evolution of a species are mechanisms of adaptation [1,3]. When selective pressures are differentially distributed across the landscape, there remains the potential for populations of species to become locally adapted to selective pressures, increasing their fitness within a unique environment [121]. By understanding these interactions between environments and populations, and how they have shaped the genetic structure of populations, we can better inform both wildlife disease- and species-management. Data presented herein suggested that the unique distributions of arctic rabies variants in Alaska have not led to locally adapted populations of red fox, indicating no differential selection between arctic rabies variants. This finding is relevant to wildlife disease management in Alaska and other northern regions as the Arctic continues to warm, likely resulting in range shifts of host species.

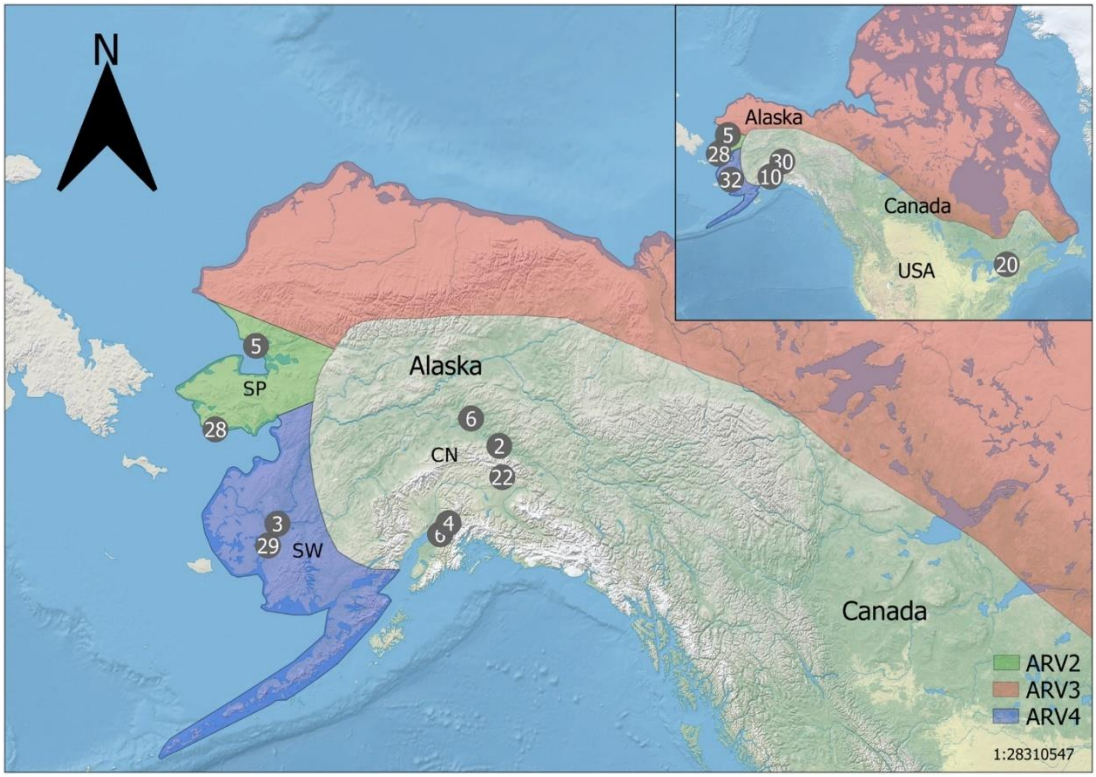

Fig 2.1. Schematic of red fox samples used in the genotype-by-sequencing assay.

Circles indicate sample locations, numbers within circles indicate number of samples within proximity to one another (5mm on the map). CN—Central (South/Interior) Alaska; SP—Seward Peninsula; SW—southwest Alaska; ON—Ontario. Approximate arctic rabies viral variant distributions (AR 2, 3, 4) are depicted by coloured regions (see in-figure legend). Insert (top right) shows relative positions of Alaskan red fox samples (n = 105) to those sampled from Ontario (n = 20). The schematic of arctic rabies variant distributions was adapted from Goldsmith et al. (2016) for illustrative purposes only. Made with Natural Earth.

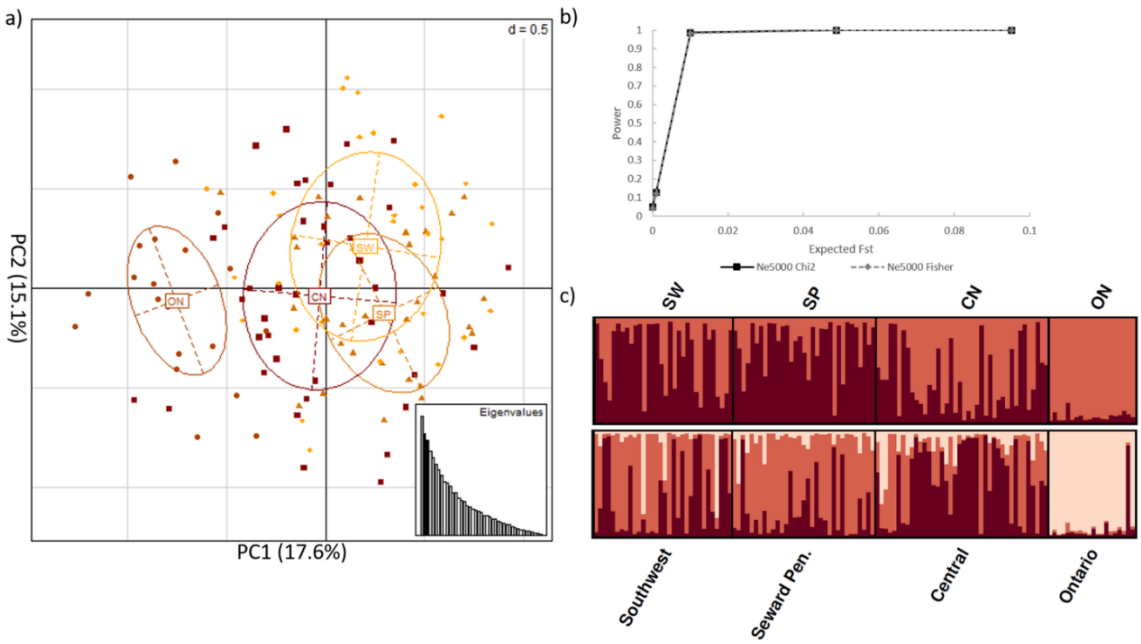

Fig 2.2. Neutral genetic structure between red fox from Alaska and Ontario. Analyses of the 43 SNPs in intergenic regions after filtering with *Variant Effect Predictor*, a MAF threshold = 2%, and pruning for Linkage disequilibrium; CN = central Alaska; SP = Seward Peninsula; SW = Southwest Alaska; ON = Ontario, Canada **a)** principle component analysis **b)** power analysis results for the estimated Fisher's exact F_{ST} and Chi^2 after t = 0, 10, 100, 500, and 1,000 generations and assuming an effective population size of 5,000 **c)** *STRUTURE* analysis of K = 2 and K = 3, where individuals are represented by each bar along the x-axis and assignment to clusters is represented by the y-axis and the different colours.

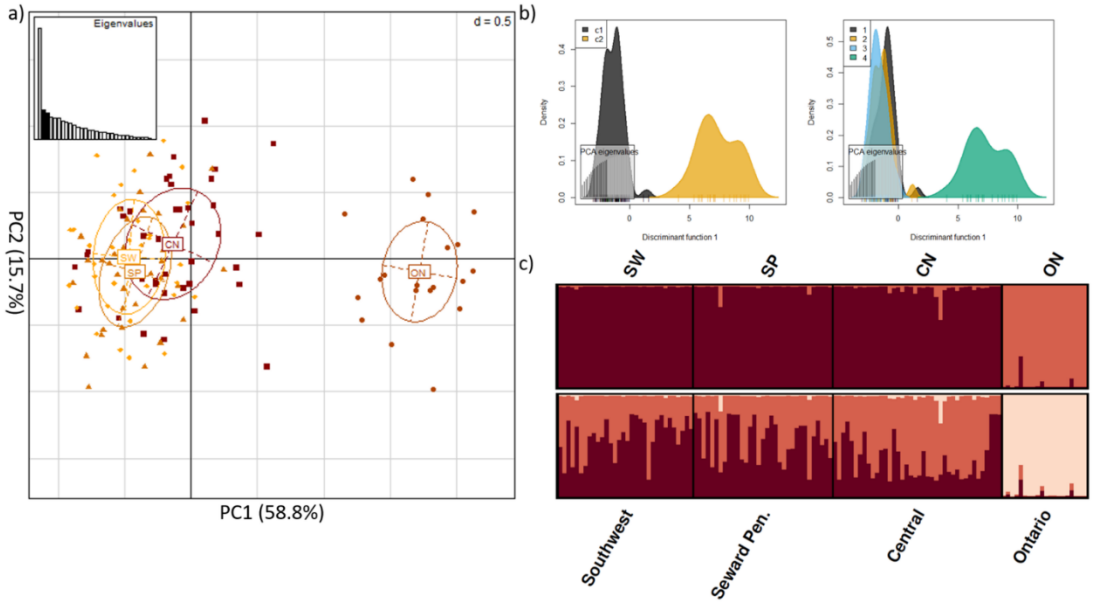

Fig 2.3. Immunogenetic structure distinguishes between red fox from Alaska and Ontario.

Analyses of the 30 outlier SNPs within protein-coding regions after filtering for a minor allele frequency threshold = 2% and pruning for linkage disequilibrium a) principle component analysis b) DAPC of the inferred clustering; c1 = inferred cluster 1, c2 = inferred cluster 2 (left) DAPC of Alaskan (1 = CN; 2 = SP; 3 = SW) vs Ontarian (4) red fox samples (right) c) STRUTURE analyses of K = 2 and K = 3, where individuals are represented by each bar along the x-axis and assignment to clusters is represented by the y-axis and the ` different colours; CN = central Alaska; SP = Seward Peninsula; SW = Southwest Alaska; ON = Ontario, Canada.

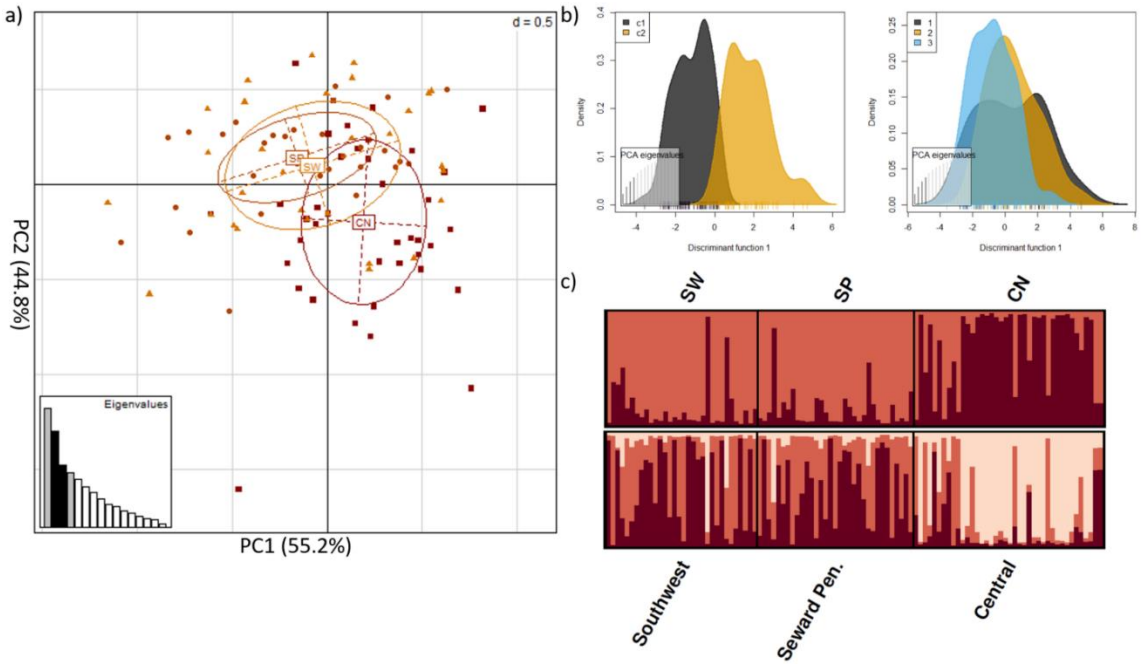

Fig 2.4. Weak signature of immunogenetic structure among red fox populations within Alaska (not including Ontario). Analyses of the 16 outlier SNPs within protein-coding regions, after filtering for a minor allele frequency threshold = 2% and pruning for linkage disequilibrium a) principle component analysis b) DAPC of the inferred clustering; c1 = inferred cluster 1, c2 = inferred cluster 2 (left) DAPC of Central (1), Seward Peninsula. (2) and Southwest (3) red fox samples in Alaska (right) c) STRUTURE analysis of K = 2, where individuals are represented by each bar along the x-axis and assignment to clusters is represented by the y-axis and the different colours; CN = central Alaska; SP = Seward Peninsula; SW = Southwest Alaska.

2.6 References

1. Kawecki TJ, Ebert D. Conceptual issues in local adaptation. Ecology letters. 2004 Dec 1;7(12):1225–41.
2. Pantel JH, Duvivier C, Meester LD. Rapid local adaptation mediates zooplankton community assembly in experimental mesocosms. Ecology letters. 2015 Oct;18(10):992–1000. pmid:26251339
3. Nuismer SL, Thompson JN, Gomulkiewicz R. Gene flow and geographically structured coevolution. Proceedings of the Royal Society of London. Series B: Biological Sciences. 1999 Mar 22;266(1419):605–9.
4. Thompson JN, Burdon JJ. Gene-for-gene coevolution between plants and parasites. Nature. 1992 Nov;360(6400):121.
5. Whitlock MC. Fixation probability and time in subdivided populations. Genetics. 2003 Jun 1;164(2):767–79. pmid:12807795
6. Van Tienderen PH. Generalists, specialists, and the evolution of phenotypic plasticity in sympatric populations of distinct species. Evolution. 1997 Oct;51(5):1372–80. pmid:28568610
7. Sultan SE. Phenotypic plasticity for fitness components in Polygonum species of contrasting ecological breadth. Ecology. 2001 Feb;82(2):328–43.
8. Schweizer RM, Robinson J, Harrigan R, Silva P, Galverni M, Musiani M, et al. Targeted capture and resequencing of 1040 genes reveal environmentally driven functional variation in grey wolves. Molecular ecology. 2016 Jan;25(1):357–79. pmid:26562361
9. Dionne M, Miller KM, Dodson JJ, Caron F, Bernatchez L. Clinal variation in MHC diversity with temperature: evidence for the role of host–pathogen interaction on local adaptation in Atlantic salmon. Evolution. 2007 Sep;61(9):2154–64. pmid:17767587
10. Frick WF, Pollock JF, Hicks AC, Langwig KE, Reynolds DS, Turner GG, et al. An emerging disease causes regional population collapse of a common North American bat species. Science. 2010 Aug 6;329(5992):679–82. pmid:20689016
11. Frick WF, Puechmaille SJ, Willis CK. White-nose syndrome in bats. In ‘Bats in the Anthropocene: Conservation of Bats in a Changing World’. (Eds Voigt CC and Kingston. T.) pp. 245–262.
12. Yaremych SA, Warner RE, Mankin PC, Brawn JD, Raim A, Novak R. West Nile virus and high death rate in American crows. Emerging infectious diseases. 2004 Apr;10(4):709. pmid:15200865
13. W, Hayes VM, Ratan A, Petersen DC, Wittekindt NE, Miller J, et al. Genetic diversity and population structure of the endangered marsupial Sarcophilus harrisii (Tasmanian devil). Proceedings of the National Academy of Sciences. 2011 Jul 26;108(30):12348–53.

14. Pritchard JK, Di Rienzo A. Adaptation–not by sweeps alone. Nature Reviews Genetics. 2010 Sep 14;11(10):665. pmid:20838407
15. Pritchard JK, Pickrell JK, Coop G. The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. Current biology. 2010 Feb 23;20(4):R208–15. pmid:20178769
16. Johnson C, Johnson J, Vanderloo JP, Keane D, Aiken JM, McKenzie D. Prion protein polymorphisms in white-tailed deer influence susceptibility to chronic wasting disease. Journal of General Virology. 2006 Jul 1;87(7):2109–14. pmid:16760415
17. Power AG. Competition between Viruses in a Complex Plant—Pathogen System. Ecology. 1996 Jun;77(4):1004–10.
18. Ekblom R, Saether SA, Jacobsson P, Fiske P, Sahlman T, Grahn M, et al. Spatial pattern of MHC class II variation in the great snipe (Gallinago media). Molecular ecology. 2007 Apr;16(7):1439–51. pmid:17391268
19. Spurgin LG, Richardson DS. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. Proceedings of the Royal Society B: Biological Sciences. 2010 Jan 13;277(1684):979–88. pmid:20071384
20. Eizaguirre C, Lenz TL, Sommerfeld RD, Harrod C, Kalbe M, Milinski M. Parasite diversity, patterns of MHC II variation and olfactory based mate choice in diverging three-spined stickleback ecotypes. Evolutionary Ecology. 2011 May 1;25(3):605–22.
21. Savage AE, Zamudio KR. MHC genotypes associate with resistance to a frog-killing fungus. Proceedings of the National Academy of Sciences. 2011 Oct 4;108(40):16705–10. pmid:21949385
22. Kyle CJ, Rico Y, Castillo S, Srithayakumar V, Cullingham CI, White BN, et al. Spatial patterns of neutral and functional genetic variations reveal patterns of local adaptation in raccoon (Procyon lotor) populations exposed to raccoon rabies. Molecular ecology. 2014 May;23(9):2287–98. pmid:24655158
23. Doxiadis GG, Otting N, de Groot NG, Noort R, Bontrop RE. Unprecedented polymorphism of Mhc-DRB region configurations in rhesus macaques. The Journal of Immunology. 2000 Mar 15;164(6):3193–9. pmid:10706710
24. Gutierrez-Espeleta GA, Hedrick PW, Kalinowski ST, Garrigan D, Boyce WM. Is the decline of desert bighorn sheep from infectious disease the result of low MHC variation?. Heredity. 2001 Apr;86(4):439. pmid:11520344
25. Rico Y, Morris-Pocock J, Zigouris J, Nocera JJ, Kyle CJ. Lack of spatial immunogenetic structure among wolverine (Gulo gulo) populations suggestive of broad scale balancing selection. PloS one. 2015 Oct 8;10(10):e0140170. pmid:26448462
26. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PloS one. 2011 May 4;6(5):e19379. pmid:21573248

27. Poland JA, Rife TW. Genotyping-by-sequencing for plant breeding and genetics. The Plant Genome. 2012 Nov 1;5(3):92–102.
28. Poland JA, Brown PJ, Sorrells ME, Jannink JL. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PloS one. 2012 Feb 28;7(2):e32253. pmid:22389690
29. Schweizer RM, Vonholdt BM, Harrigan R, Knowles JC, Musiani M, Coltman D, et al. Genetic subdivision and candidate genes under selection in North American grey wolves. Molecular ecology. 2016 Jan;25(1):380–402. pmid:26333947
30. Murchison EP, Schulz-Trieglaff OB, Ning Z, Alexandrov LB, Bauer MJ, Fu B, et al. Genome sequencing and analysis of the Tasmanian devil and its transmissible cancer. Cell. 2012 Feb 17;148(4):780–91. pmid:22341448
31. Siddle HV, Marzec J, Cheng Y, Jones M, Belov K. MHC gene copy number variation in Tasmanian devils: implications for the spread of a contagious cancer. Proceedings of the Royal Society B: Biological Sciences. 2010 Mar 10;277(1690):2001–6. pmid:20219742
32. Rupprecht CE, Hanlon CA, Hemachudha T. Rabies re-examined. The Lancet infectious diseases. 2002 Jun 1;2(6):327–43. pmid:12144896
33. Kuzmin IV, Hughes GJ, Botvinkin AD, Gribencha SG, Rupprecht CE. Arctic and Arctic-like rabies viruses: distribution, phylogeny and evolutionary history. Epidemiology & Infection. 2008 Apr;136(4):509–19.
34. Nadin-Davis SA, Sheen M, Wandeler AI. Recent emergence of the Arctic rabies virus lineage. Virus research. 2012 Jan 1;163(1):352–62. pmid:22100340
35. Dyer JL, Yager P, Orciari L, Greenberg L, Wallace R, Hanlon CA, et al. Rabies surveillance in the United States during 2013. Journal of the American Veterinary Medical Association. 2014 Nov 15;245(10):1111–23. pmid:25356711
36. Guerra MA, Curns AT, Rupprecht CE, Hanlon CA, Krebs JW, Childs JE. Skunk and raccoon rabies in the eastern United States: temporal and spatial analysis. Emerging Infectious Diseases. 2003 Sep;9(9):1143. pmid:14519253
37. Goldsmith EW, Renshaw B, Clement CJ, Himschoot EA, Hundertmark KJ, Hueffer K. Population structure of two rabies hosts relative to the known distribution of rabies virus variants in Alaska. Molecular ecology. 2016 Feb;25(3):675–88. pmid:26661691
38. Nadin-Davis SA, Muldoon F, Wandeler AI. Persistence of genetic variants of the arctic fox strain of Rabies virus in southern Ontario. Canadian journal of veterinary research. 2006 Jan;70(1):11. pmid:16548327
39. Haydon DT, Cleaveland S, Taylor LH, Laurenson MK. Identifying reservoirs of infection: a conceptual and practical challenge. Emerging infectious diseases. 2002 Dec;8(12):1468–73. pmid:12498665

40. Fenton A, Pedersen AB. Community epidemiology framework for classifying disease threats. Emerging infectious diseases. 2005 Dec;11(12):1815. pmid:16485464
41. Nishiura H, Hoye B, Klaassen M, Bauer S, Heesterbeek H. How to find natural reservoir hosts from endemic prevalence in a multi-host population: A case study of influenza in waterfowl. Epidemics. 2009 Jun 1;1(2):118–28. pmid:21352759
42. Larivière S, Pasitschniak-Arts M. *Vulpes vulpes*. Mammalian species. 1996 Dec 27(537):1–1.
43. Dell'Arte GL, Laaksonen T, Norrdahl K, Korpimäki E. Variation in the diet composition of a generalist predator, the red fox, in relation to season and density of main prey. Acta oecologica. 2007 May 1;31(3):276–81.
44. Kumar V, Kutschera VE, Nilsson MA, Janke A. Genetic signatures of adaptation revealed from transcriptome sequencing of Arctic and red foxes. BMC genomics. 2015 Dec;16(1):585. pmid:26250829
45. Edwards CJ, Soulsbury CD, Statham MJ, Ho SY, Wall D, Dolf G, et al. Temporal genetic variation of the red fox, *Vulpes vulpes*, across western Europe and the British Isles. Quaternary Science Reviews. 2012 Dec 4;57:95–104. pmid:24068852
46. Norén K, Angerbjörn A, Wallén J, Meijer T, Sacks BN. Red foxes colonizing the tundra: genetic analysis as a tool for population management. Conservation Genetics. 2017 Apr 1;18(2):359–70.
47. Zecchin B, De Nardi M, Nouvellet P, Vernesi C, Babbucci M, Crestanello B, et al. Genetic and spatial characterization of the red fox (*Vulpes vulpes*) population in the area stretching between the Eastern and Dinaric Alps and its relationship with rabies and canine distemper dynamics. PloS one. 2019 Mar 12;14(3):e0213515. pmid:30861028
48. Hueffer KO'Hara TM, Follmann EH. Adaptation of mammalian host-pathogen interactions in a changing arctic environment. Acta Veterinaria Scandinavica. 2011 Dec;53(1):17.
49. Huettmann F, Magnuson EE, Hueffer K. Ecological niche modeling of rabies in the changing Arctic of Alaska. Acta Veterinaria Scandinavica. 2017 Dec;59(1):18. pmid:28320440
50. Hueffer K, Murphy M. Rabies in Alaska, from the past to an uncertain future. International journal of circumpolar health. 2018 Jan 1;77(1):1475185. pmid:29764319
51. Cullingham CI, Kyle CJ, Pond BA, Rees EE, White BN. Differential permeability of rivers to raccoon gene flow corresponds to rabies incidence in Ontario, Canada. Molecular Ecology. 2009 Jan;18(1):43–53. pmid:19140963

52. Blanchong JA, Samuel MD, Scribner KT, Weckworth BV, Langenberg JA, Filcek KB. Landscape genetics and the spatial distribution of chronic wasting disease. Biology Letters. 2007 Dec 11;4(1):130–3.
53. Donaldson ME, Rico Y, Hueffer K, Rando HM, Kukekova AV, Kyle CJ. Development of a genotype-by-sequencing immunogenetic assay as exemplified by screening for variation in red fox with and without endemic rabies exposure. Ecology and evolution. 2018 Jan;8(1):572–83. pmid:29321894
54. MacInnes CD, Smith SM, Tinline RR, Ayers NR, Bachmann P, Ball DG, et al. Elimination of rabies from red foxes in eastern Ontario. Journal of wildlife diseases. 2001 Jan;37(1):119–32. pmid:11272485
55. Mørk T, Prestrud P. Arctic rabies–a review. Acta Veterinaria Scandinavica. 2004 Mar;45(1):1.
56. Rosatte RC, Power MJ, Donovan D, Davies JC, Allan M, Bachmann P, et al. Elimination of arctic variant rabies in red foxes, metropolitan Toronto. Emerging Infectious Diseases. 2007 Jan;13(1):25. pmid:17370512
57. Bradley MJ, Kutz SJ, Jenkins E, O'hara TM. The potential impact of climate change on infectious diseases of Arctic fauna. International Journal of Circumpolar Health. 2005 Dec 1;64(5):468–77. pmid:16440609
58. Parkinson AJ, Butler JC. Potential impacts of climate change on infectious diseases in the Arctic. International Journal of Circumpolar Health. 2005 Dec 1;64(5):478–86. pmid:16440610
59. Sokolov AA, Sokolova NA, Ims RA, Brucker L, Ehrich D. Emergent rainy winter warm spells may promote boreal predator expansion into the Arctic. Arctic. 2016 Jun 6;69(2):121–9.
60. Elmhagen B, Kindberg J, Hellström P, Angerbjörn A. A boreal invasion in response to climate change? Range shifts and community effects in the borderland between forest and tundra. Ambio. 2015 Jan 1;44(1):39–50. pmid:25576279
61. Elmhagen B, Tannerfeldt M, Angerbjörn A. Food-niche overlap between arctic and red foxes. Canadian Journal of Zoology. 2002 Jul 1;80(7):1274–85.
62. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.3997. 2013 Mar 16.
63. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009 Aug 15;25(16):2078–9. pmid:19505943
64. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research. 2010 Sep 1;20(9):1297–303. pmid:20644199

65. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature genetics. 2011 May;43(5):491. pmid:21478889
66. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. Current protocols in bioinformatics. 2013 Oct;43(1):11–0. pmid:25431634
67. Ahrens CW, Rymer PD, Stow A, Bragg J, Dillon S, Umbers KD, et al. The search for loci under selection: trends, biases and progress. Molecular ecology. 2018 Mar;27(6):1342–56. pmid:29524276
68. Donaldson ME, Davy CM, Willis CK, McBurney S, Park A, Kyle CJ. Profiling the immunome of little brown myotis provides a yardstick for measuring the genetic response to white-nose syndrome. Evolutionary applications. 2017 Dec;10(10):1076–90. pmid:29151862
69. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, et al. The ensembl variant effect predictor. Genome biology. 2016 Dec;17(1):122. pmid:27268795
70. Team RC. R: A language and environment for statistical computing.
71. Jombart T. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics. 2008 Apr 8;24(11):1403–5. pmid:18397895
72. Paradis E, Claude J, Strimmer K. APE: analyses of phylogenetics and evolution in R language. Bioinformatics. 2004 Jan 22;20(2):289–90. pmid:14734327
73. Chhatre VE, Emerson KJ. StrAuto: automation and parallelization of STRUCTURE analysis. BMC bioinformatics. 2017 Dec;18(1):192. pmid:28340552
74. Earl DA. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation genetics resources. 2012 Jun 1;4(2):359–61.
75. Jakobsson M, Rosenberg NA. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics. 2007 May 7;23(14):1801–6. pmid:17485429
76. Rosenberg NA. DISTRUCT: a program for the graphical display of population structure. Molecular ecology notes. 2004 Mar;4(1):137–8.
77. Ryman N, Palm S. POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. Molecular Ecology Notes. 2006 Sep;6(3):600–2.
78. Luu K, Bazin E, Blum MG. pcadapt: an R package to perform genome scans for selection based on principal component analysis. Molecular ecology resources. 2017 Jan;17(1):67–77. pmid:27601374

79. Whitlock MC, Lotterhos KE. Reliable detection of loci responsible for local adaptation: Inference of a null model through trimming the distribution of F ST. The American Naturalist. 2015 Oct 1;186(S1):S24–36.
80. Excoffier L, Lischer HE. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular ecology resources. 2010 May;10(3):564–7. pmid:21565059
81. Foll M, Gaggiotti O. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics. 2008 Oct 1;180(2):977–93. pmid:18780740
82. Gautier M, Klassmann A, Vitalis R. rehh 2.0: a reimplementation of the R package rehh to detect positive selection from haplotype structure. Molecular ecology resources. 2017 Jan;17(1):78–90. pmid:27863062
83. Gautier M, Vitalis R. rehh: an R package to detect footprints of selection in genome-wide SNP data from haplotype structure. Bioinformatics. 2012 Apr 15;28(8):1176–7. pmid:22402612
84. Gautier M, Naves M. Footprints of selection in the ancestral admixture of a New World Creole cattle breed. Molecular ecology. 2011 Aug;20(15):3128–43. pmid:21689193
85. Voight B.F. et al. (2006) A Map of Recent Positive Selection in the Human Genome Hurst L., ed. PLoS Biology, 4(3), p.e72. [R.S4.4]. pmid:16494531
86. Sabeti P.C. et al. (2007) Genome-wide detection and characterization of positive selection in human populations. Nature, 449(7164), 913–918. pmid:17943131
87. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly. 2012 Apr 1;6(2):80–92. pmid:22728672
88. Nei M, Gojobori T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Molecular biology and evolution. 1986 Sep 1;3(5):418–26. pmid:3444411
89. Fuller ZL, Niño EL, Patch HM, Bedoya-Reina OC, Baumgarten T, Muli E, et al. Genome-wide analysis of signatures of selection in populations of African honey bees (Apis mellifera) using new web-based tools. BMC genomics. 2015 Dec;16(1):1–8. pmid:26159619
90. Huguet G, Nava C, Lemiere N, Patin E, Laval G, Ey E, et al. Heterogeneous pattern of selective pressure for PRRT2 in human populations, but no association with autism spectrum disorders. PloS one. 2014 Mar 3;9(3):e88600. pmid:24594579
91. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. DnaSP 6: DNA sequence polymorphism analysis of

large data sets. Molecular biology and evolution. 2017 Dec 1;34(12):3299–302. pmid:29029172
92. Rausch RL. Observations on some natural-focal zoonoses in Alaska. Archives of Environmental Health: An International Journal. 1972 Oct 1;25(4):246–52. pmid:5055675
93. Srithayakumar V, Sribalachandran H, Rosatte R, Nadin-Davis SA, Kyle CJ. Innate immune responses in raccoons after raccoon rabies virus infection. Journal of General Virology. 2014 Jan 1;95(1):16–25. pmid:24085257
94. Madhu BP, Singh KP, Saminathan M, Singh R, Tiwari AK, Manjunatha V, et al. Correlation of inducible nitric oxide synthase (iNOS) inhibition with TNF-α, caspase-1, FasL and TLR-3 in pathogenesis of rabies in mouse model. Virus genes. 2016 Feb 1;52(1):61–70. pmid:26690069
95. Ito N, Moseley GW, Sugiyama M. The importance of immune evasion in the pathogenesis of rabies virus. Journal of Veterinary Medical Science. 2016:16–0092.
96. Li J, Faber M, Dietzschold B, Hooper DC. The role of toll-like receptors in the induction of immune responses during rabies virus infection. InAdvances in virus research 2011 Jan 1 (Vol. 79, pp. 115–126). Academic Press. pmid:21601045
97. Komastu T, Ireland DD, Reiss CS. IL-12 and viral infections. Cytokine & growth factor reviews. 1998 Dec 1;9(3–4):277–85. pmid:9918125
98. Guo Y, Long J, He J, Li CI, Cai Q, Shu XO, et al. Exome sequencing generates high quality data in non-target regions. BMC genomics. 2012 Dec;13(1):194. pmid:22607156
99. Ellegren H, Smeds L, Burri R, Olason PI, Backström N, Kawakami T, et al. The genomic landscape of species divergence in Ficedula flycatchers. Nature. 2012 Nov;491(7426):756. pmid:23103876
100. Varshney RK, Song C, Saxena RK, Azam S, Yu S, Sharpe AG, et al. Draft genome sequence of chickpea (Cicer arietinum) provides a resource for trait improvement. Nature biotechnology. 2013 Mar;31(3):240. pmid:23354103
101. Zhao S, Zheng P, Dong S, Zhan X, Wu Q, Guo X, et al. Whole-genome sequencing of giant pandas provides insights into demographic history and local adaptation. Nature Genetics. 2013 Jan;45(1):67. pmid:23242367
102. Attard CR, Beheregaray LB, Sandoval-Castillo J, Jenner KC, Gill PC, Jenner MN, et al. From conservation genetics to conservation genomics: a genome-wide assessment of blue whales (Balaenoptera musculus) in Australian feeding aggregations. Royal Society open science. 2018 Jan 31;5(1):170925. pmid:29410806
103. Morin PA, Martien KK, Taylor BL. Assessing statistical power of SNPs for population structure and conservation studies. Molecular Ecology Resources. 2009 Jan;9(1):66–73. pmid:21564568

104. Attard CR, Beheregaray LB, Möller LM. Genotyping-by-sequencing for estimating relatedness in nonmodel organisms: Avoiding the trap of precise bias. Molecular ecology resources. 2018 May;18(3):381–90. pmid:29160928
105. Kamler JF, Ballard WB. A review of native and nonnative red foxes in North America. Wildlife Society Bulletin. 2002 Jul 1:370–9.
106. Aubry KB, Statham MJ, Sacks BN, Perrine JD, Wisely SM. Phylogeography of the North American red fox: vicariance in Pleistocene forest refugia. Molecular Ecology. 2009 Jun;18(12):2668–86. pmid:19457180
107. Berteaux D, Gallant D, Sacks BN, Statham MJ. Red foxes (*Vulpes vulpes*) at their expanding front in the Canadian Arctic have indigenous maternal ancestry. Polar Biology. 2015 Jun 1;38(6):913–7.
108. Narum SR, Hess JE. Comparison of FST outlier tests for SNP loci under selection. Molecular ecology resources. 2011 Mar;11:184–94. pmid:21429174
109. Schaefer C, Rost B. Predict impact of single amino acid change upon protein structure. InBMC genomics 2012 Jun (Vol. 13, No. 4, p. S4). BioMed Central. pmid:22759652
110. Li S, Walters W, Chassaing B, Zhang B, Shi Q, Waters J, et al. Gut microbiota influence B cell function in a TLR5-dependent manner. bioRxiv. 2019 Jan 1:537894.
111. Paulos CM, Wrzesinski C, Kaiser A, Hinrichs CS, Chieppa M, Cassard L, et al. Microbial translocation augments the function of adoptively transferred self/tumor-specific CD8+ T cells via TLR4 signaling. The Journal of clinical investigation. 2007 Aug 1;117(8):2197–204. pmid:17657310
112. Semnani RT, Venugopal PG, Leifer CA, Mostböck S, Sabzevari H, Nutman TB. Inhibition of TLR3 and TLR4 function and expression in human dendritic cells by helminth parasites. Blood. 2008 Aug 15;112(4):1290–8. pmid:18541719
113. Miao EA, Andersen-Nissen E, Warren SE, Aderem A. TLR5 and Ipaf: dual sensors of bacterial flagellin in the innate immune system. InSeminars in immunopathology 2007 Sep 1 (Vol. 29, No. 3, pp. 275–288). Springer-Verlag. pmid:17690885
114. Turner AJ, Aggarwal P, Miller HE, Waukau J, Routes JM, Broeckel U, et al. The introduction of RNA-DNA differences underlies interindividual variation in the human IL12RB1 mRNA repertoire. Proceedings of the National Academy of Sciences. 2015 Dec 15;112(50):15414–9. pmid:26621740
115. Nadin-Davis SA, Fehlner-Gardiner C. Origins of the arctic fox variant rabies viruses responsible for recent cases of the disease in southern Ontario. PLoS neglected tropical diseases. 2019 Sep 6;13(9):e0007699. pmid:31490919

116. Medvedev AE. Toll-like receptor polymorphisms, inflammatory and infectious diseases, allergies, and cancer. Journal of Interferon & Cytokine Research. 2013 Sep 1;33(9):467–84.
117. Zhang J, Lu Y, Pienta KJ. Multiple roles of chemokine (CC motif) ligand 2 in promoting prostate cancer growth. Journal of the National Cancer Institute. 2010 Apr 21;102(8):522–8. pmid:20233997
118. Sunyer JO, Boshra H, Lorenzo G, Parra D, Freedman B, Bosch N. Evolution of complement as an effector system in innate and adaptive immunity. Immunologic research. 2003 Jun 1;27(2–3):549–64. pmid:12857998
119. Crispín JC, Hedrich CM, Tsokos GC. Gene-function studies in systemic lupus erythematosus. Nature reviews rheumatology. 2013 Aug;9(8):476. pmid:23732569
120. Kukurba KR, Montgomery SB. RNA sequencing and analysis. Cold Spring Harbor Protocols. 2015 Nov 1;2015(11):pdb-top084970. pmid:25870306
121. Waser NM, Price MV. Reciprocal transplant experiments with Delphinium nelsonii (Ranunculaceae): evidence for local adaptation. American Journal of Botany. 1985 Nov;72(11):1726–32.

Chapter 3

Genetic structure of immunologically associated candidate genes suggests arctic rabies variants exert differential selection in arctic fox populations

Preface

A version of this chapter is published in *PLOS One* (2021),16(10): e0258975
https://doi.org/10.1371/journal.pone.0258975

Authors: Tristan M. Baecklund, Michael E. Donaldson, Karsten Hueffer, Christopher J. Kyle

Reference: Baecklund TM, Donaldson ME, Hueffer K, Kyle CJ (2021) Genetic structure of immunologically associated candidate genes suggests arctic rabies variants exert differential selection in arctic fox populations. PLoS ONE 16(10): e0258975. https://doi.org/10.1371/journal. pone.0258975

Modifications: Figures and tables were renumbered to accommodate the incorporation into this thesis. Spelling/grammatical errors were also made for consistency throughout the thesis. Additional sentences were included to clarify methodological approaches outlined in the paper (e.g., the significance test for the pN/pS ratios). No other aspect of the manuscript has been modified from the original.

Contributions: MED and CJK conceived and designed the study; TMB performed all lab work; TMB analyzed and visualized the data; TMB wrote the manuscript; TMB, MED, KH, and CJK critically reviewed the manuscript.

Abstract

Patterns of local adaptation can emerge in response to the selective pressures diseases exert on host populations as reflected in increased frequencies of respective, advantageous genotypes. Elucidating patterns of local adaptation enhance our understanding of mechanisms of disease spread and the capacity for species to adapt in context of rapidly changing environments such as the Arctic. Arctic rabies is a lethal disease that largely persists in northern climates and overlaps with the distribution of its natural host, arctic fox. Arctic fox populations display little neutral genetic structure across their North American range, whereas phylogenetically unique arctic rabies variants are restricted in their geographic distributions. It remains unknown if arctic rabies variants impose differential selection upon host populations, nor what role different rabies variants play in the maintenance and spread of this disease. Using a targeted, genotyping-by-sequencing assay, we assessed correlations of arctic fox immunogenetic variation with arctic rabies variants to gain further insight into the epidemiology of this disease. Corroborating past research, we found no neutral genetic structure between sampled regions, but did find moderate immunogenetic structuring between foxes predominated by different arctic rabies variants. F_{ST} outliers associated with host immunogenetic structure included SNPs within interleukin and Toll-like receptor coding regions (IL12B, IL5, TLR3 and NFKB1); genes known to mediate host responses to rabies. While these data do not necessarily reflect causation, nor a direct link to arctic rabies, the contrasting genetic structure of

immunologically associated candidate genes with neutral loci is suggestive of differential selection and patterns of local adaptation in this system. These data are somewhat unexpected given the long-lived nature and dispersal capacities of arctic fox; traits expected to undermine local adaptation. Overall, these data contribute to our understanding of the co-evolutionary relationships between arctic rabies and their primary host and provide data relevant to the management of this disease.

3.1 Introduction

Hosts and pathogens are in a continual co-evolutionary arms race, where patterns of local adaptation can emerge in response to the selective pressures diseases exert on host populations, and thus influence disease spread and maintenance [1–3]. Elucidating where adaptations have occurred throughout the genomes of host populations can provide better understanding of disease mechanisms in general, including the impacts of pathogens in shaping host population diversity, and how perturbations to these systems may influence disease distributions and outcomes in host populations.

Divergent selection can lead to locally adapted populations across heterogeneous landscapes where selective pressures differ [1–3]. When divergent selection occurs without interference from other forces, local populations evolve traits best suited to local pressures providing increased fitness within a specific environment regardless of the consequences of the trait in different environments [1–3]. In natural populations, the process of local adaptation is largely influenced by

three factors: gene flow, effective population size/genetic drift, and force of the selective pressure [1]. In these natural systems, homogenization of variation through gene flow and stochastic loss of variants via genetic drift can undermine increases in adaptive trait frequencies that are suggestive of local adaptation [1–3]. Thus, when trying to elucidate patterns of local adaptation in natural populations is it necessary to evaluate the effects of gene flow and genetic drift in context of the distribution and frequencies of traits under natural selection.

Patterns of local adaptation are often assessed through common garden experiments, where populations are exposed to a series of different environmental variables, and changes in fitness are observed over time [1]. In natural populations common garden experiments are not always feasible, and they are further complicated by plastic responses of populations, where single genotypes can elicit multiple phenotypes that mask measurable changes in fitness [4, 5]. Genetic assessments of the interplay between selection and demographic forces, such as by contrasting patterns of neutral genetic structure of populations relative to the genetic structure of loci under selection, can provide an alternate means to detect genetic signatures indicative of local adaptation in lieu of common garden experiments. Examples of this approach include assessments of the variation of salmonid immune responses within environments with different aquatic thermal regimes [6] and variation in genes associated with vision and hearing in wolf populations in context of specific environmental variables [7].

While species are exposed to a myriad of selective pressures, infectious diseases can exert strong selective pressures over short periods of time, where the genetic composition of a population can change considerably within a few generations from these selective sweeps [8, 9]. The emergence of white-nose syndrome in bats [10, 11] and the development of facial tumors in Tasmanian devil populations [12], where both diseases led to drastic population declines, are exemplar of the rapid effects infectious diseases can have on natural populations. In some patho-systems, such as chronic wasting disease (CWD) in mule deer, population differences in the frequency of genotypes responsible for susceptibility to CWD have been observed, providing strong evidence for local adaption to this disease over time [13]. Patterns of host-pathogen interactions can also occur over longer time scales and lead to coevolutionary interactions when diseases are endemic and multiple variants of disease circulate in the environment. In these systems, host populations adapt to persist against pathogens and pathogens evolve to circumvent host immune defenses [2, 3]. Thus, a population's response to disease can be dependent upon both genetic variants circulating within host populations, but also the genetic variants of the pathogen(s) host populations are exposed to [14].

Explorations of population responses to disease have typically undertaken genetic assessments of the major histocompatibility complex (MHC) due to its association with antigen binding and its highly polymorphic nature [15–18]. Specifically, MHC DRB exon-2 has been used extensively as an indicator of the

genetic variation of MHC [19, 20], yet studies have indicated that variation in other genomic areas associated with an immune response also play important roles in defenses against infectious disease [21–24]. Genotyping-by-sequencing (GBS) assays enable assessments of genetic variation from a larger number of loci simultaneously, thus providing an opportunity to explore the variation of genes associated with an immune response more holistically [7, 25, 26]. For example, GBS has been used to explore 138 genes associated with the immune response in little brown bats in context of white-nose syndrome exposure [27], and Elbers et al. [28] found immunologically relevant variants associated with macromolecule and protein modifications in gopher tortoises that influenced upper respiratory tract disease severity. In the absence of feasible/pragmatic common garden experiments, GBS techniques provide a means to study the genetic impacts disease can have on host populations.

Arctic rabies (AR) is a lethal lyssavirus that circulates in northern climates through its natural host, the arctic fox (*Vulpes lagopus*), where epizootic cycles of the disease occur every 3–6 years [29, 30]. Arctic rabies is comprised of four phylogenetically distinct subvariants, where all four variants circulate in unique geographically maintained distributions in North America (Fig 3.1) [30–33]. Arctic rabies variant 2 (ARV2) is restricted to the Seward Peninsula of Alaska, AR variant 4 (ARV4) is restricted to Southwestern Alaska, and AR variant 3 (ARV3) circulates along northern coasts across North America and Eurasia [31]. However, AR variant 1 (ARV1) circulates only in Southern Ontario and is maintained in the absence of

arctic fox, presumably by red fox [34]. Despite these phylogenetic and distribution differences, it remains unknown whether geographically restricted AR variants have differences in pathogenicity that may impose divergent selection between arctic fox populations, and potentially reveal signatures of locally adapted host populations.

In Alaska, three variants of AR circulate predominantly within red and arctic fox populations [31]. Previous red fox studies found that patterns of neutral genetic structure [31] and genetic variants associated with an immune response [35] demonstrated correlations of red fox genetic structure with the presence/absence of AR. However, no signatures indicative of differential selection were observed between AR variants in the red fox system. In contrast, the observed neutral genetic structure of arctic fox was noted to closely parallel the distribution of AR variants [31]. These data were corroborated by the fact that where AR variant 3 persists throughout northern coasts, arctic fox exist as a largely panmictic population as a matter of high gene flow facilitated by sea ice connectivity [36–38]. While inferences of disease spread/maintenance through observations of patterns of host gene flow based on neutral genetic markers are feasible, neutral markers alone do not provide insight into coevolutionary patterns that may exist between arctic fox and AR. Therefore, assessments of genetic variation associated with responses to selective pressures exerted by AR, such as genes related to an immune response, have the potential to further our understanding AR spread and maintenance in arctic fox populations.

Herein, we build upon previous research [31, 35] using an immunogenetic assay targeting 116 regions of the arctic fox genome associated with an immune response. We aimed to: 1) determine if genetic variants associated with an immune response give rise to patterns of genetic structure in arctic fox, and 2) determine if patterns of differential selection exist in artic fox relative to AR variant distributions that may be indicative of local adaption to this disease; data that also provides insight into the maintenance and spread of AR. The Arctic continues to experience rapid warming, thus understanding host population responses to different disease variants, and the potential for local adaptation in hosts, becomes increasingly important as climatic changes are expected to cause range shifts in both pathogens and their hosts [29, 30, 39–43]. Overall, this research aims to enhance our understanding of AR dynamics in arctic fox where unique distributions of AR variants are maintained in North America.

3.2 Methods

3.2.1 Sampling, DNA extractions and quantification

Arctic fox muscle tissue samples were obtained from various independent trappers and organizations as part of the University of Alaska Museum of the North tissue collections or as part of previous research [36; Table S3.1]. No direct handling/sampling of animals took place for this study. Samples were stored at -80°C until processed. Samples were digested in 200 µL 1X lysis buffer (4 M Urea, 0.2 M NaCl, 0.5% n-lauroyl sarcosine, 10 mM ethylenediaminetetraacetic acid (EDTA), 0.1 M Tris HCl pH 8.0) with the addition of 20 µL proteinase K and incubated at 56°C

for two hours. During digestion, samples were vortexed and briefly spun down every 30 minutes. DNA was extracted from the resulting lysate utilizing the DNeasy Blood and Tissue Kit (Qiagen) following manufacturer protocols with the exception that DNA was eluted in a total volume of 60 µL, using two 30 µL aliquots of TE buffer (10 mM Tris, 0.1 mM EDTA). Isolated DNA was quantified using the Quant-iT PicoGreen dsDNA Assay Kit (ThermoFisher Scientific) and quality assessed by ethidium stained 0.8% agarose gel electrophoresis (90 V for 45 minutes) where DNA fragment size was assessed in context of HighRanger 1 kbp DNA ladder (Norgen Biotek). A subset of 96 high molecular weight DNA samples suitable for sequencing were selected from three regions across the arctic fox's distribution in North America (Fig 3.1; Table S3.1). Samples from Hooper Bay and Chevak are referred to as a single region 'Southwestern Alaska' based on their geographic proximity to one another but were left ungrouped for neutral genetic analyses (~ 30 kilometers).

3.2.2 Library preparation, sequence capture and high-throughput sequencing

DNA libraries were prepared using Kapa HyperPlus Kit (Roche) following the SeqCap-EZ HyperCap UGuide V1.0 (Roche) protocol. Seven cycles were implemented as part of the pre-LM PCR as recommended by the manufacturer with the following modifications to the workflow: i) PCR-grade water was used for dilutions and elution, ii) samples were treated with 5 µL of conditioning solution during fragmentation, iii) TruSeq HT Dual-Index Adapters (Integrated DNA Technologies) were used in place of SeqCap Adapter Kits A and B (Roche), and iv) Illumina P5 and P7 primers (Integrated DNA Technologies) were substituted in place

of the Pre LM-PCR Oligos 1 & 2 (Roche). At the end of the Pre-capture LM-PCR step, DNA library quality was assessed using ethidium bromide-stained gel electrophoresis as per above.

A 1 µg DNA multiplex was created from equal-molar amounts of each of the 96 libraries. Target enrichment was performed as previously described [35], using the designed SeqCap EZ Developer Library probe. Modifications to the enrichment protocol implemented in this study included: i) replacement of the NimbleGen Multiplex Hybridization Enhancing Oligo Pool (Roche) with 2 µL xGen Universal Blockers–TS Mix (Integrated DNA Technologies), ii) NimbleGen SeqCap EZ Developer Reagent (Roche) was used in place of the NimbleGen COT Human DNA (Roche) during hybridization sample preparation, and iii) hybridization was carried out at 47°C for 20 hours. A final product assessment was conducted with a bioanalyzer on the target-enriched multiplex before sequencing on an Illumina MiSeq V3 run using 2x300 bp reads (Advanced Analysis Centre Genomics Facility, University of Guelph).

3.2.3 Sequence alignment and variant annotation

Utilizing the bwa-mem command in Burrows-Wheeler Aligner v0.7.12 [44], paired-end reads for each of the 96 samples were aligned to the canine reference genome (CanFam3.1; Fig 3.1; Table S3.1). After sequence metrics were obtained using SAMTOOLS v1.5 [45], the Genome Analysis Toolkit (GATK, V4.0.0.0) best practices pipeline and standard hard filtering parameters were used to perform duplicate

sequence removal, SNP/INDEL variant annotation, genotyping, and variant recalibration [46–48]. The SelectVariants function was then used to compile a VCF file containing only bi-allelic SNPs.

The SeqCap EZ Developer Library probe was originally designed from a draft version of the red fox genome [33], thus positions of the probe-baited targets needed to be determined and converted into positions in the canine reference genome (CanFam3.1; Accession: PRJNA12384) using BLASTn (Table S3.2). These targeted regions were compiled into a list of on-target intervals. On-target intervals were used to further categorize SNPs as being within coding regions (including 1,500 bp upstream from the start codon) or within intergenic (off-target; outside targeted coding regions and promoters) regions. We attempted to mitigate biases to identify loci under selection using F_{ST} outlier tests as recommended in the literature [49] by accounting for: linkage disequilibrium within datasets, method variation by implementing several F_{ST} outlier tests, and filtering variants for minimum allele frequency (MAF).

3.2.4 SNP filtering and analyses

Both sub datasets of SNPs from within coding regions, and those from intergenic regions were filtered using VCFtools v0.1.13 to retain only biallelic variants with a MAF threshold of 2.5%, a maximum missing genotype threshold (per site) of 20% and excluding variants on the X-chromosome. Additionally, both filtered sub-datasets were pruned for linkage disequilibrium as implemented by the SNPRelate

package in R v.3.5 [50, 51] and further pruned for physical linkage to only retain SNPs ≥ 100kbp from one another using bcftools v1.9 [52] (File S3.1 for further details). SNPs from within intergenic regions underwent further filtering using the Ensembl Variant Effect Predictor [53] to retain only those SNPs ≥ 40 kbp from an annotated coding region. Disequilibrium and physical linkage pruning occurred after F_{ST} outlier tests for the sub-dataset composed of SNPs from within coding regions. Mantel tests (as implemented in R) [50] were performed on both the intergenic SNP-dataset and the SNP-dataset from coding regions to identify patterns of isolation-by-distance.

3.2.5 Analyses of SNPs in intergenic regions

Variants passing filtering parameters set for the intergenic regions were assumed to not be under selective influence and were therefore used to estimate patterns of neutral population genetic structure. The filtered sub-dataset was analyzed using principal component analysis (PCA) and discriminant analyses of principle components (DAPC) as implemented in RStudio using the adegenet (v2.1.1) [54] and ape (v5.1) [55] packages. Principle components with eigenvalues ≥ 0.1 were retained for the PCA, and cross validation was used to determine the number of retained components based on the root mean squared error (lowest MSE) for the DAPC. Discriminant analyses implemented successive K-means to determine the optimal number of identified clusters of the data.

Implementing STRAUTO (v.1.0) [56], STRUCTURE analyses were performed with a burn-in length of 50,000 followed by 200,000 iterations for K = 1 through K = 6, with 20 iterations of each K. Using structure harvester web (v0.6.94) [57], the ΔK statistic was calculated to determine the number of distinct genetic clusters inferred from the data. CLUMPP v1.1.2 [58], and the LargeKGreedy algorithm (10,000 repeats) were used to assign individuals to genetic clusters, followed by the implementation of DISTRUCT v1.1 [59] to combine and visualize results.

POWSIM v.4.1 [60] was used to estimate the effective power of the presumed neutral SNP sub-dataset to detect genetic structure. Simulations were run with N_E = 500 and 5,000, over (t) generations = 0, 10, 100, 500, and 1,000. Each set of conditions was performed 1,000 times to differentiate between the three sampled regions and evaluate the power to detect genetic homogeneity with both chi-square and Fisher's exact tests. The Fisher's exact test was implemented within the program using a Monte Carlo Markov chain approach with default parameters of 1,000 burn-ins, 100 batches, and 1,000 iterations.

3.2.6 Analyses of SNPs in coding regions

Variants within coding regions that passed initial MAF and missing data filtering parameters were assessed with PCAdapt [61], OutFLANK [62], Arlequin [63], and Bayescan [64] to identify F_{ST} outliers within the sub-dataset. Each of these tests identified outliers using an adjusted p-value threshold of ≤ 0.05; more detailed parameters for each method can be found as a supplement (File S3.1).

Inconsistencies in identified outliers can occur between different methods due to differences in underlying assumptions and caveats used by each method [62]. For the purposes of this study, we retained any outlier identified by at least one of the methods in the final sub-set of SNPs from within coding regions. The final sub-set of outliers was then pruned for linkage disequilibrium and physical linkage prior to PCA, DAPC, and structure analyses, as described above. Further, to provide an adequate control for these data, we attempted to repeat F_{ST} outlier detection on the off-target dataset utilizing the same methods and parameters as the on-target data. F_{ST} estimates were generated using VCFtools (based on Weir & Cockerham, 1984), and 97.5% F_{ST} confidence intervals were determined in R.

The program SnpEff was used to annotate synonymous and non-synonymous polymorphisms within the dataset of SNPs from coding regions (using the CanFam3.1.99 database) [65]. Using these annotations, we calculated the relative ratio of non-synonymous substitutions per non-synonymous site to the number of synonymous substitutions per synonymous site (pN/pS) as highlighted by Nei and Gojobori; $pN = \frac{Nd}{N}$ and $pS = \frac{Sd}{S}$ [66]. Where Nd/Sd is the number of nonsynonymous or synonymous polymorphisms and N or S is the total number of nonsynonymous or synonymous sites [66, 67]. Following previous research, we determined the potential number of nonsynonymous/synonymous sites using DnaSP v6 and the coding sequence for each gene as input [67, 68] where positive selection can be inferred from pN/pS ratios > 1, and purifying selection from ratios <

1 [69]. A chi-square test was implemented to evaluate if the observed pN/pS ratios were statistically different from the expected values of neutral selection.

3.3 Results

3.3.1 Raw sequence data

We obtained an average of ~ 563,000 raw reads per library, 99.5% of which mapped to the canine reference genome. After processing the raw data through the GATK SNP calling pipeline, an average of ~ 109,000 (19.35%) reads were filtered from each sample library leaving ~ 452,000 reads per library. An average of 56% reads per library mapped to probe targeted regions with an average of 54 X coverage across all 96 libraries (Table S3.3).

3.3.2 SNPs in intergenic (off-target) regions

A dataset of 5,490,704 intergenic SNPs was filtered using Variant Effect Predictor (VEP), which was then pruned to minimize linkage disequilibrium. This yielded a dataset of 29 intergenic SNPs presumed to be not under selective pressure with an average depth of coverage of 15 X across all SNPs (average depth of coverage of 5 X when excluding 4 SNPs with coverage exceeding 15X), and a MAF of 2.5% (Table S3.4). These data were visualized using PCA, DAPC and STRUCTURE. PCA did not identify genetic structuring as all clusters had extensive overlap with each other (S3.9 Fig). The DAPC identified K = 6 as the most likely number of clusters; however, STRUCTURE analyses of K = 2–6 showed high levels of admixture across all analyses and indicated an optimal K = 2 clusters (Fig 3.2). Power analyses of these

29 SNPs indicated a power of ~86% at an expected F_{ST} of 0.01 and a power of 100% at an expected F_{ST} of 0.05 or above, indicating a high likelihood that if population differentiation was > 1%, our dataset had the power to detect that difference. Isolation-by-distance was not observed within intergenic SNPs based on Mantel tests. We observed a total of 4 outlier SNPs (13 prior to linkage pruning) within the off-target dataset, all of which were identified by PCAdapt, and visualization of these data demonstrate no genetic clustering (S3.8 Fig). Overall, these analyses indicated no apparent genetic structure across the sample design for the off-target, and presumed neutral, SNP dataset.

3.3.3 SNPs in coding (on-target) regions

We found 9,467 SNPs located within target intervals before filtering. After applying search criteria for 2.5% MAF, a maximum of 20% missing data, and discarding SNPs on the X-chromosome, 2,277 SNPs remained. F_{ST} outlier analyses on these 2,277 SNPs produced a sub-dataset containing 107 SNPs (Table S3.5). After accounting for linkage disequilibrium, a final sub-dataset containing 22 F_{ST} outlier SNPs remained, several of which were associated with interleukins, Toll-like receptors, and the MHC (Table S3.5). Based on VEP results, four SNPs retained in the final sub-dataset associated with DLA-DQA, NOD1, RAG1 and TLR5 genes, were likely to convey a change in chemical characteristic of the encoded amino acid during translation (e.g., acidic to basic amino acid). A further 9 SNPs, that conveyed such missense changes, were removed from the final sub-dataset when filtering for linkage disequilibrium (Table S3.5). DAPC and STRUCTURE analyses of the

remaining 22 F_{ST} outlier SNPs were assessed identifying K = 2 clusters. We found arctic fox sampled from Alaska appeared genetically distinct from foxes from Northern Canada (Fig 3.3) with F_{ST} estimates of 0.127. F_{ST} confidence intervals indicated that F_{ST} between Southwestern Alaska and the two regions in Canada were significantly different from zero; however, F_{ST} between the two Canadian arctic fox populations was not significantly different from zero (Table S3.6). It is important to note however, that we could not exclude isolation-by-distance as a potential factor contributing to these patterns (could not be calculated being only 2 geographic points–Canadian vs Alaskan samples).

Estimates of pN/pS were determined for 90 of the initial 116 targeted genes (Table S3.7) although for 17 of these genes, pN/pS ratios could not be calculated due to a lack of polymorphic substitutions (either nonsynonymous or synonymous) resulting in pN values equal to zero or division by zero errors (occurring when pS = 0). The pN/pS ratios for most genes were indicative of purifying selection; however, 10 genes (C2, DLA-12, DLA-79, DLA-DQA, DLA-DBQBC1, IL1B, IL23A, MYD88, STAT3, and STAT6) appeared to be under positive selection (pN/pS ratio ≥ 1) in all three arctic fox populations sampled (Table S3.7). One gene appeared to be under positive selection only in Southwestern Alaska (CD8A), whereas both CCL2 and CCR8 were under positive selection within foxes sampled from Canada (Arviat and Victoria Island; Table S3.7). Based on an assessment of Chi^2 p-values (Table S3.7) only the STAT3 gene had a p value of $p < 0.05$ (0.000137), indicating a significantly different pN/pS ratio from expected values.

3.4 Discussion

Herein, we contribute to the understanding of host/pathogen evolutionary systems by examining the genetic structure of immunologically associated molecular markers of arctic fox in context of arctic rabies variants. Specifically, we used a genotyping-by-sequencing (GBS) assay to explore immunogenetic regions of the arctic fox genome that also yielded off-target sequence data from intergenic regions presumed to be neutral. We take these presumed neutral data, combined with immunogenetically relevant data, to suggest patterns of local adaptation exist in this system as expected if AR variants displayed differential selection upon arctic fox populations across their range. While these correlations do not necessarily reflect causation, AR is known to present strong selective forces on arctic fox, combined with the observed structure at several identified F_{ST} outliers associated with genes known to mediate response to rabies [70–73] in a presumably panmictic population (37–39; and data presented herein), is notable. Further, pN/pS ratios identified genes under positive selection in arctic fox across North America, but also between sampled regions where different AR variants circulate. As only one gene provided a significantly different pN/pS ratio from expected values (Table S3.7), these data should be interpreted with caution. Specifically, previous research highlights that these analyses should not be interpreted as sufficient evidence of selection on their own [74, 75] and are subject to several biasing factors, such as few polymorphisms identified within genes [69, 76]. We take these data to imply that patterns of local adaptation exist in the arctic fox/rabies system reflecting the

strong selective pressure AR likely exerts on arctic fox populations despite the impressive dispersal abilities and panmictic neutral genetic structure of arctic fox [31, 36–38]. While these data provide insight into how unique distributions of AR variants are maintained, potential differences in pathogenicity between AR variants have not been established and further research is required to ascertain more definitive insight into the interrelationship of AR variants and their main host. Specifically, research encompassing more arctic foxes from across their range that include all AR variant distributions accompanied by further evidence of underlying pathogenic differences between variants would benefit our understanding of AR maintenance and spread.

3.4.1 Analyses of intergenic SNPs (off-target)

The generation of secondary (off-target) sequences using targeted GBS approaches is a common feature observed in other studies [77]. These untargeted products are often consistently sequenced allowing these data to be used in downstream analyses if they are of sufficient coverage and quality [77–79]. Studies have noted that simulation-based assessments should be combined with off-target datasets to determine the power of these data to discern differences between populations and therefore form meaningful conclusions [80, 81]. Although the observed neutral genetic structure presented here is devoid of the subtle structure previously found among arctic fox populations in Alaska (F_{ST} = 0.02 with microsatellites) [31], it remains consistent with coarse geographic scale studies of the species [36–38]. This is unsurprising to a limited extent given foxes from only the northern coast of

Alaska were used as representatives for the whole state in these coarse-scale assessments as previously noted by Goldsmith et al. [31]. It is worth noting that the inability to detect similar patterns of neutral population genetic structure in this study were likely due to the limited number of variants passing the filtering parameters, combined with differences in the power of discrimination between microsatellites and biallelic SNPs. Furthermore, when we attempt to identify F_{ST} outliers within the off-target dataset, there were insufficient numbers of outliers (4 after linkage pruning) to draw any meaningful result. The lack of identified outliers likely results from prominent gene flow documented among arctic fox, and as such, in conjunction with data presented in these previous studies [31, 36–38], we take the presumed neutral data presented herein to provide further evidence of the extensive gene flow among North American arctic fox populations.

While we acknowledge that filtering parameters employed herein were rigorous and likely removed informative SNPs form the final subsets used in analyses, determining neutral genetic structure of arctic fox across North America was not a direct objective of this study having been investigated thoroughly elsewhere [31, 36–38]. Additionally, recent research suggests recombination distances vary across chromosomes in red foxes (ranging from 0.07 cM – 5 cM between pericentromeric regions and chromosome ends) [82], meaning that the filtering step to ensure these SNPs were not in proximity (≥ 40 kbp) of any annotated coding regions provides only a coarse approximation of neutrality at these sites.

With these potential limitations acknowledged, we used off-target data to provide a contextual baseline for the on-target data.

3.4.2 Analysis of SNPs in protein-coding regions

Multiple F_{ST} outlier identification programs were implemented to identify SNPs within genomic regions of interest in order to mitigate discordance between methods [62]. Of the 107 identified F_{ST} outlier SNPs within protein-coding regions, only 32 were identified by multiple programs, and only two were identified by all four methods used. The number of identified F_{ST} outliers between programs ranged from 9–61, where Bayescan identified the least and PCAdapt identified the largest number of F_{ST} outliers. Most identified F_{ST} outlier SNPs resulted in synonymous changes at their respective positions in the genome and thus unlikely to affect function of synthesized proteins. In contrast, 13 identified F_{ST} outlier SNPs conveyed a missense mutation that also changed chemical characteristics of translated amino acids, increasing the likelihood that these mutations could affect subsequent protein function [83]. These 13 missense F_{ST} outlier SNPs were associated with gene sequences of DLA-DQA, DLA-DQBC1, NOD1, RAG1 and TLR5, although only SNPs associated with the latter four genes were retained in the final filtered, linkage pruned, sub-dataset of 22 F_{ST} outlier SNPs. Both DLA-DQA and DLA-DQBC1 are class II components of the dog leukocyte antigen, responsible for initiating the immune response through antigen presentation and recognition [84]. NOD1 recognizes gram-negative bacteria and initiates a pro-inflammatory response [85], and RAG1 is a factor initiating immunoglobulin V(D)J recombination [86].

Finally, TLR5 detects bacteria with flagellin and induces a pro-inflammatory response [87]. Of interest are the large number of F_{ST} outlier SNP associations to interleukin and Toll-like receptor gene families, especially in context of AR, as members of these gene families are implicated in mediating a response to rabies [70–73]. Further, three missense F_{ST} outlier SNPs were associated with TLR5 demonstrating that this gene may play a prominent role in mediating responses to AR variants, although specific mechanisms of this response require further investigation.

Analyses of signatures of selection through assessments of pN/pS ratios demonstrate several genes that may be related to the spatially distinct distributions of AR variants. The CD8A gene is uniquely under positive selection in Southwestern Alaska, where AR variant 4 circulates. However, genes CCL2 and CCR8 appear to be under positive selection in both sampled regions in Canada where different AR variant 3 circulates. These data suggest that these three genes may be under positive selection, indicative of differential selection, relative to the AR variants in those locations and may play a role in the mechanisms that maintain unique distributions of AR variants. Of the ten genes under positive selection in all three populations, a large portion are involved in the MHC, which is expected given this gene family's involvement in antigen binding and overall health of organisms [15–18]. Furthermore, two genes under selection in all three populations are associated with interleukins, a gene family already implicated in mediating a response to rabies [70, 71]. It is important to note that estimates of pN/pS are subject to potential

biases in systems where there is prominent gene flow and migration [67], as is the case of arctic fox across their range, and by genes where there are few polymorphisms identified [69, 76]. To mitigate these potential biases, some researchers implement a threshold of pN/pS ratios > 2 to be indicative of positive selection [67]. In this context, only the STAT3 gene would appear to be under positive selection across all three of our sampled populations (pN/pS ratio of ~3.4). This gene was also the only gene to have a statistically significant pN/pS ratio with a p value of $p < 0.05$ (0.000137). However, in combination with the potentially inflated ratio of STAT3 due to those biases mentioned above, there were no outlier SNPs detected within the STAT3 gene. As such, although this gene may be under directional selection, it is unlikely to be contributing to the patterns of genetic structure observed within the on-target SNP-dataset. This is contrasted by outlier SNPs with the potential to change protein function identified within the genes DLQ-DQA and DLQ-DQBC1, however, these genes lacked significance in their pN/pS ratios that were suggestive of positive selection (pN/pS > 1; Table S3.7). Given these results and potential biases, we present these pN/pS ratios only as an initial estimate requiring further testing and that these pN/pS results should be interpreted with caution.

3.4.3 Arctic rabies variant distributions and differential selection

Analyses of F_{ST} outlier SNPs demonstrate that genetic differentiation between arctic fox populations inhabiting an AR variant region were not significantly different from zero, however, genetic differentiation between arctic fox populations inhabiting

regions where a different AR variant circulates were significantly different from zero. Given that arctic fox populations have been exposed to AR over a large time frame [88], there remains the potential for coevolutionary forces to have shaped patterns of differential selection between AR variants where analyses of pN/pS ratios and identified F_{ST} outlier SNPs within immunogenetically relevant genes tentatively support these observations.

Previous data [31, 35–38], combined with those presented herein, demonstrate that arctic fox populations are largely panmictic with prominent gene flow, yet still appear to be locally adapted to the different AR variants. It had been anticipated that elevated levels of gene flow among arctic foxes [31, 35–38] would homogenize AR variants across the landscape, precluding local adaptation. In addition, rabies incubation periods range from several days to several months [85], which would prevent movement of AR variants, and further undermine the spatially distinct AR distributions observed. Some authors have theorized that dispersal capabilities of rabid foxes are reduced [31], and thus maintain AR variants spatial distributions where populations then undergo differential selection and subsequent local adaptation. However, there are no data to suggest phylogenetically distinct AR variants elicit differential responses, making it unclear how the weak genetic structure of immunogenic variants observed elicit differential selection. It is of interest that despite differences in biogeography and geographic distances between arctic fox populations from Arviat and Victoria Island, where AR variant 3 circulates, these two populations of arctic fox show no genetic structure of either presumed

neutral or functional markers associated with an immunogenetic response. Further, in arctic fox from Southwest Alaska, where AR variant 4 circulates, there appears to be no neutral genetic structure in context of the Arviat or Victoria Island populations. This pattern contrasts the genetic clustering of functional, immunogenetically relevant, markers distinguishing between those arctic fox populations from Canada relative to those in Alaska; suggestive of differential selection between arctic rabies variants 3 and 4. Overall, we take these data to suggest that AR variants may impose differential selective pressures on populations despite the impressive dispersal capabilities and gene flow within the primary host, arctic fox.

There remains potential that the observed patterns indicative of local adaptation between arctic fox and AR variants to have arisen from purifying selection or genetic drift, rather than directional selection as interpreted here [89], with drift being unlikely in a panmictic system with extensive gene flow. We attempted to account for purifying selection biasing our interpretation by providing pN/pS ratios for genes where calculations were possible. Given the pitfalls of targeted sequencing approaches, such as the difficulty of interpreting signatures of selection from genomic data and successfully targeting the most informative loci [89], we present only candidate genes suggestive of patterns of differential selection. Continued research, such as whole genome analyses or sequencing of identified candidate genes presented herein, will be required to provide more support for the identified patterns of local adaptation. Whole genome analyses

would benefit these data, as it would facilitate the testing of whether the observed patterns of genetic structure documented herein are consistent among other regions of the genome or an artifact of sequencing a handful of genes from the genome. Further, future research that aims to investigate the interrelationship between arctic fox and AR should sample from more populations of arctic fox from each AR variant distribution as well as include multiple populations for each AR variant region where possible. Additionally, the inclusion of foxes that have survived exposure to each AR variant, as well as those succumbing to the disease would greatly enhance the study by allowing for more in-depth analyses between AR variant regions, and direct consequences/benefits of specific SNPs. Lastly, there remains the potential for unknown factors, beyond the scope of the objectives and methods implemented here to have identified them, that may better explain the observed patterns of genetic structure.

3.5 Conclusions

Infectious diseases can pose strong selective pressures on populations. Important to our understanding of the spread and maintenance of such diseases are the underlying interactions between host and pathogen [1–3]. By studying genetic variation within host populations associated with the immune response, we increase our understanding of how infectious diseases shape populations over time through patterns of local adaptation. Data from this study are also relevant to wildlife disease management efforts for arctic rabies where range shifts are occurring for both arctic and red fox, key arctic rabies vectors, in a rapidly warming

Arctic. It remains unknown how these range shifts will affect the distributions of arctic rabies variants in North America.

Fig 3.1. Approximate arctic rabies variant distributions in North America and schematic of the 96 arctic fox samples used in the genotype-by-sequencing assay. Circles indicate sample locations; numbers within circles indicate sample size. Samples were obtained from Arviat, NU (n = 36); Victoria Island, NWT (n = 36); and two regions of Southwest Alaska, (Chevak (n = 12) and Hooper Bay (n = 12). Approximate arctic rabies viral variant distributions (ARV 2,3,4) are depicted by coloured regions (see in-figure legend). The schematic of viral variant distributions was adapted from Goldsmith et al. (2016) for illustrative purposes only. Map created using Natural Earth (naturalearthdata.com).

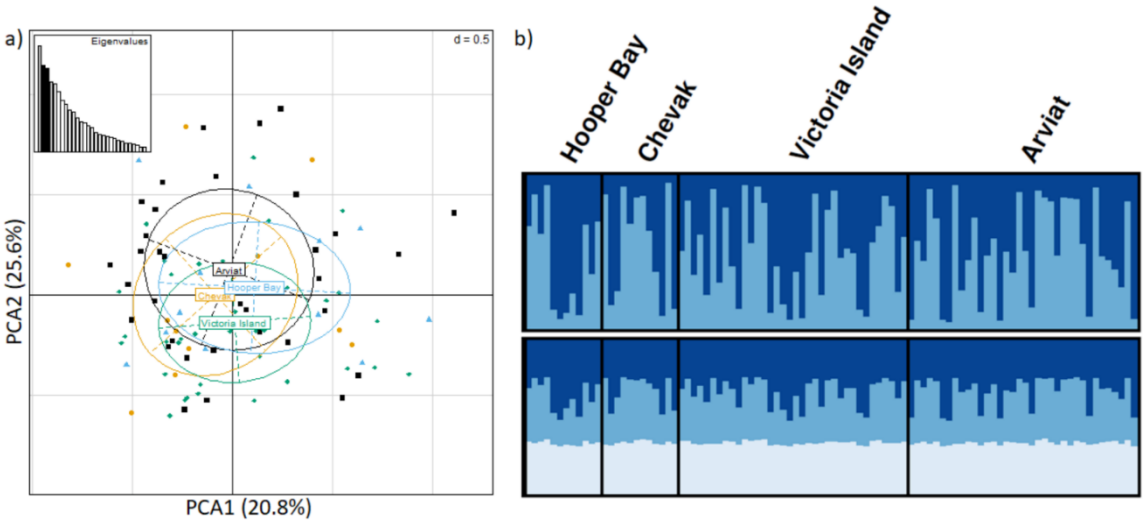

Fig 3.2. Neutral genetic homogeneity of arctic fox across North America. Analyses of the 29 presumed neutral SNPs after filtering with Variant Effect Predictor, MAF = 2.5%, and pruning for Linkage disequilibrium **A)** principle component analysis **B)** STRUCTURE analysis of K = 2 and 3 (top and bottom respectively), where each bar across horizontal axis indicates an individual, vertical axis depicts cluster assignment, and different colours depict each genetic cluster; Arviat = black square; Chevak = yellow circle; Hooper Bay = blue triangle; Victoria Island = green diamond.

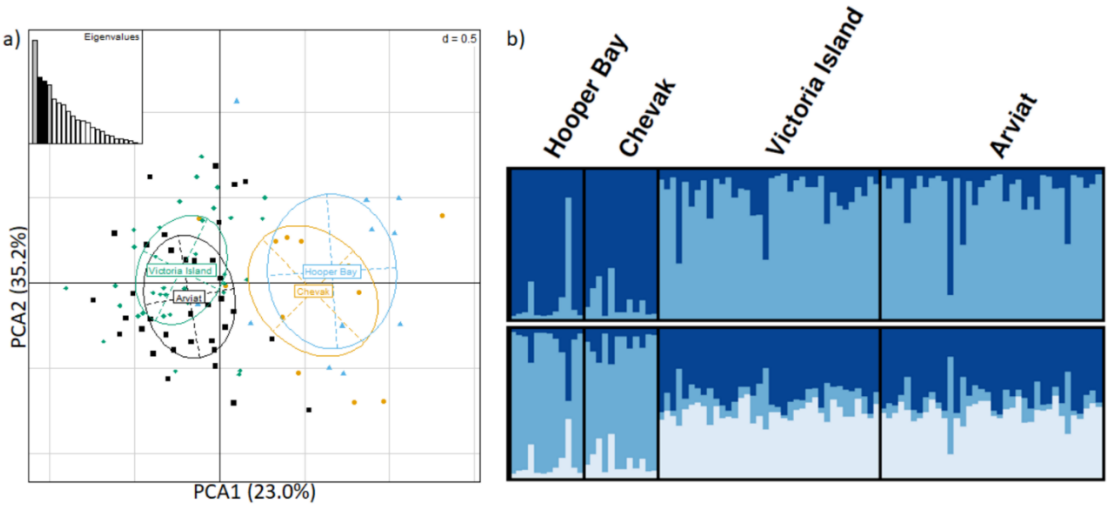

Fig 3.3. Arctic fox immunogenetic structure differentiates sampled regions in Southwestern Alaska and Northern Canada. Analyses of 22 protein-coding F_{ST} outlier SNPs after filtering for MAF = 2.5% and pruning for Linkage disequilibrium **A)** principle component analysis and **B)** STRUTURE analysis of K = 2 and 3 (top and bottom respectively), where each bar across horizontal axis indicates an individual, vertical axis depicts cluster assignment, and different colors depict each genetic clusters; Arviat = black square; Chevak = yellow circle; Hooper Bay = blue triangle; Victoria Island = green diamond (note, Chevak and Hooper Bay were pooled for this analysis).

3.6 References

1. Kawecki TJ, Ebert D. Conceptual issues in local adaptation. Ecology letters. 2004 Dec 1;7(12):1225–41.
2. Nuismer SL, Thompson JN, Gomulkiewicz R. Gene flow and geographically structured coevolution. Proceedings of the Royal Society of London. Series B: Biological Sciences. 1999 Mar 22;266(1419):605–9.
3. Thompson JN, Burdon JJ. Gene-for-gene coevolution between plants and parasites. Nature. 1992 Nov;360(6400):121.
4. Van Tienderen PH. Generalists, specialists, and the evolution of phenotypic plasticity in sympatric populations of distinct species. Evolution. 1997 Oct;51(5):1372–80. pmid:28568610
5. Sultan SE. Phenotypic plasticity for fitness components in Polygonum species of contrasting ecological breadth. Ecology. 2001 Feb;82(2):328–43.
6. Dionne M, Miller KM, Dodson JJ, Caron F, Bernatchez L. Clinal variation in MHC diversity with temperature: evidence for the role of host–pathogen interaction on local adaptation in Atlantic salmon. Evolution. 2007 Sep;61(9):2154–64. pmid:17767587
7. Schweizer RM, Vonholdt BM, Harrigan R, Knowles JC, Musiani M, Coltman D, et al. Genetic subdivision and candidate genes under selection in North American grey wolves. Molecular ecology. 2016 Jan;25(1):380–402. pmid:26333947
8. O'Brien SJ, Evermann JF. Interactive influence of infectious disease and genetic diversity in natural populations. Trends in Ecology & Evolution. 1988 Oct 1;3(10):254–9. pmid:21227241
9. Kundu S, Faulkes CG, Greenwood AG, Jones CG, Kaiser P, Lyne OD, et al. Tracking viral evolution during a disease outbreak: the rapid and complete selective sweep of a circovirus in the endangered Echo parakeet. Journal of virology. 2012 May 1;86(9):5221–9. pmid:22345474
10. Frick WF, Pollock JF, Hicks AC, Langwig KE, Reynolds DS, Turner GG, et al. An emerging disease causes regional population collapse of a common North American bat species. Science. 2010 Aug 6;329(5992):679–82. pmid:20689016
11. Frick WF, Puechmaille SJ, Willis CK. White-nose syndrome in bats. In 'Bats in the Anthropocene: Conservation of Bats in a Changing World'. (Eds CC Voigt and T. Kingston.) 2016:245–262.
12. Miller W, Hayes VM, Ratan A, Petersen DC, Wittekindt NE, Miller J, et al. Genetic diversity and population structure of the endangered marsupial Sarcophilus harrisii (Tasmanian devil). Proceedings of the National Academy of Sciences. 2011 Jul 26;108(30):12348–53.
13. Johnson C, Johnson J, Vanderloo JP, Keane D, Aiken JM, McKenzie D. Prion protein polymorphisms in white-tailed deer influence susceptibility to chronic wasting disease. Journal of General Virology. 2006 Jul 1;87(7):2109–14. pmid:16760415

14. Power AG. Competition between Viruses in a Complex Plant--Pathogen System. Ecology. 1996 Jun;77(4):1004–10.
15. Ekblom R, Saether SA, Jacobsson P, Fiske P, Sahlman T, Grahn M, et al. Spatial pattern of MHC class II variation in the great snipe (Gallinago media). Molecular ecology. 2007 Apr;16(7):1439–51. pmid:17391268
16. Spurgin LG, Richardson DS. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. Proceedings of the Royal Society B: Biological Sciences. 2010 Jan 13;277(1684):979–88. pmid:20071384
17. Eizaguirre C, Lenz TL, Sommerfeld RD, Harrod C, Kalbe M, Milinski M. Parasite diversity, patterns of MHC II variation and olfactory based mate choice in diverging three-spined stickleback ecotypes. Evolutionary Ecology. 2011 May 1;25(3):605–22.
18. Savage AE, Zamudio KR. MHC genotypes associate with resistance to a frog-killing fungus. Proceedings of the National Academy of Sciences. 2011 Oct 4;108(40):16705–10. pmid:21949385
19. Doxiadis GG, Otting N, de Groot NG, Noort R, Bontrop RE. Unprecedented polymorphism of Mhc-DRB region configurations in rhesus macaques. The Journal of Immunology. 2000 Mar 15;164(6):3193–9. pmid:10706710
20. Gutierrez-Espeleta GA, Hedrick PW, Kalinowski ST, Garrigan D, Boyce WM. Is the decline of desert bighorn sheep from infectious disease the result of low MHC variation?. Heredity. 2001 Apr;86(4):439. pmid:11520344
21. Ottenhoff TH, Verreck FA, Hoeve MA, van de Vosse E. Control of human host immunity to mycobacteria. Tuberculosis. 2005 Jan 1;85(1–2):53–64. pmid:15687028
22. Behnke JM, Iraqi F, Menge D, Baker RL, Gibson J, Wakelin D. Chasing the genes that control resistance to gastrointestinal nematodes. Journal of Helminthology. 2003 Jun;77(2):99–109. pmid:12756063
23. Hill AV. The genomics and genetics of human infectious disease susceptibility. Annual review of genomics and human genetics. 2001 Sep;2(1):373–400. pmid:11701655
24. Lam-Yuk-Tseung S, Gros P. Genetic control of susceptibility to bacterial infections in mouse models. Cellular microbiology. 2003 May;5(5):299–313. pmid:12713489
25. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PloS one. 2011 May 4;6(5):e19379. pmid:21573248
26. Poland JA, Rife TW. Genotyping-by-sequencing for plant breeding and genetics. The Plant Genome. 2012 Nov 1;5(3):92–102.
27. Donaldson ME, Davy CM, Willis CK, McBurney S, Park A, Kyle CJ. Profiling the immunome of little brown myotis provides a yardstick for measuring the genetic response to white-nose syndrome. Evolutionary applications. 2017 Dec;10(10):1076–90. pmid:29151862
28. Elbers JP, Brown MB, Taylor SS. Identifying genome-wide immune gene variation underlying infectious disease in wildlife populations–a next

generation sequencing approach in the gopher tortoise. BMC genomics. 2018 Dec 1;19(1):64. pmid:29351737
29. Hueffer K, Murphy M. Rabies in Alaska, from the past to an uncertain future. International journal of circumpolar health. 2018 Jan 1;77(1):1475185. pmid:29764319
30. Huettmann F, Magnuson EE, Hueffer K. Ecological niche modeling of rabies in the changing Arctic of Alaska. Acta Veterinaria Scandinavica. 2017 Dec;59(1):18. pmid:28320440
31. Goldsmith EW, Renshaw B, Clement CJ, Himschoot EA, Hundertmark KJ, Hueffer K. Population structure of two rabies hosts relative to the known distribution of rabies virus variants in Alaska. Molecular ecology. 2016 Feb;25(3):675–88. pmid:26661691
32. Kuzmin IV, Hughes GJ, Botvinkin AD, Gribencha SG, Rupprecht CE. Arctic and Arctic-like rabies viruses: distribution, phylogeny and evolutionary history. Epidemiology & Infection. 2008 Apr;136(4):509–19.
33. Nadin-Davis SA, Sheen M, Wandeler AI. Recent emergence of the Arctic rabies virus lineage. Virus research. 2012 Jan 1;163(1):352–62. pmid:22100340
34. Nadin-Davis SA, Muldoon F, Wandeler AI. Persistence of genetic variants of the arctic fox strain of Rabies virus in southern Ontario. Canadian journal of veterinary research. 2006 Jan;70(1):11. pmid:16548327
35. Donaldson ME, Rico Y, Hueffer K, Rando HM, Kukekova AV, Kyle CJ. Development of a genotype-by-sequencing immunogenetic assay as exemplified by screening for variation in red fox with and without endemic rabies exposure. Ecology and evolution. 2018 Jan;8(1):572–83. pmid:29321894
36. Carmichael LE, Krizan J, Nagy JA, Fuglei E, Dumond M, Johnson D, et al. Historical and ecological determinants of genetic structure in arctic canids. Molecular Ecology. 2007 Aug;16(16):3466–83. pmid:17688546
37. Dalén L, Fuglei EV, Hersteinsson P, Kapel CM, Roth JD, Samelius G, et al. Population history and genetic structure of a circumpolar species: the arctic fox. Biological Journal of the Linnean Society. 2005 Jan 1;84(1):79–89.
38. Norén K, Carmichael L, Dalén L, Hersteinsson P, Samelius G, Fuglei E, et al. Arctic fox *Vulpes lagopus* population structure: circumpolar patterns and processes. Oikos. 2011 Jun;120(6):873–85.
39. Hueffer K, O'Hara TM, Follmann EH. Adaptation of mammalian host-pathogen interactions in a changing arctic environment. Acta Veterinaria Scandinavica. 2011 Dec;53(1):17.
40. Elmhagen B., Tannerfeldt M., & Angerbjörn A. (2002). Food-niche overlap between arctic and red foxes. Canadian Journal of Zoology, 80(7), 1274–1285.
41. Hersteinsson P., & Macdonald D. W. (1992). Interspecific competition and the geographical distribution of red and arctic foxes *Vulpes vulpes* and Alopex lagopus. Oikos, 505–515.

42. Pamperin N. J., Follmann E. H., & Petersen B. (2006). Interspecific killing of an arctic fox by a red fox at Prudhoe Bay, Alaska. Arctic, 361–364.
43. Tannerfeldt M., Elmhagen B., & Angerbjörn A. (2002). Exclusion by interference competition? The relationship between red and arctic foxes. Oecologia, 132(2), 213–220. pmid:28547354
44. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.3997. 2013 Mar 16.
45. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009 Aug 15;25(16):2078–9. pmid:19505943
46. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research. 2010 Sep 1;20(9):1297–303. pmid:20644199
47. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature genetics. 2011 May;43(5):491. pmid:21478889
48. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. Current protocols in bioinformatics. 2013 Oct;43(1):11–0. pmid:25431634
49. Ahrens CW, Rymer PD, Stow A, Bragg J, Dillon S, Umbers KD, et al. The search for loci under selection: trends, biases and progress. Molecular ecology. 2018 Mar;27(6):1342–56. pmid:29524276
50. Team RC. R: A language and environment for statistical computing.
51. Zheng X, Levine D, Shen J, Gogarten S, Laurie C, Weir B. A High-performance Computing Toolset for Relatedness and Principal Component Analysis of SNP Data. Bioinformatics. 2012; 28(24):3326–3328. pmid:23060615
52. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. GigaScience. 2021; 10(2): gib008. pmid:33590861
53. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, et al. The ensembl variant effect predictor. Genome biology. 2016 Dec;17(1):122. pmid:27268795
54. Jombart T. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics. 2008 Apr 8;24(11):1403–5. pmid:18397895
55. Paradis E, Claude J, Strimmer K. APE: analyses of phylogenetics and evolution in R language. Bioinformatics. 2004 Jan 22;20(2):289–90. pmid:14734327
56. Chhatre VE, Emerson KJ. StrAuto: automation and parallelization of STRUCTURE analysis. BMC bioinformatics. 2017 Dec;18(1):192. pmid:28340552

57. Earl DA. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation genetics resources. 2012 Jun 1;4(2):359–61.
58. Jakobsson M, Rosenberg NA. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics. 2007 May 7;23(14):1801–6. pmid:17485429
59. Rosenberg NA. DISTRUCT: a program for the graphical display of population structure. Molecular ecology notes. 2004 Mar;4(1):137–8.
60. Ryman N, Palm S. POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. Molecular Ecology Notes. 2006 Sep;6(3):600–2.
61. Luu K, Bazin E, Blum MG. pcadapt: an R package to perform genome scans for selection based on principal component analysis. Molecular ecology resources. 2017 Jan;17(1):67–77. pmid:27601374
62. Whitlock MC, Lotterhos KE. Reliable detection of loci responsible for local adaptation: Inference of a null model through trimming the distribution of F ST. The American Naturalist. 2015 Oct 1;186(S1):S24–36.
63. Excoffier L, Lischer HE. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular ecology resources. 2010 May;10(3):564–7. pmid:21565059
64. Foll M, Gaggiotti O. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics. 2008 Oct 1;180(2):977–93. pmid:18780740
65. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly. 2012 Apr 1;6(2):80–92. pmid:22728672
66. Nei M, Gojobori T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Molecular biology and evolution. 1986 Sep 1;3(5):418–26. pmid:3444411
67. Huguet G, Nava C, Lemiere N, Patin E, Laval G, Ey E, et al. Heterogeneous pattern of selective pressure for PRRT2 in human populations, but no association with autism spectrum disorders. PloS one. 2014 Mar 3;9(3):e88600. pmid:24594579
68. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Molecular biology and evolution. 2017 Dec 1;34(12):3299–302. pmid:29029172
69. Fuller ZL, Niño EL, Patch HM, Bedoya-Reina OC, Baumgarten T, Muli E, et al. Genome-wide analysis of signatures of selection in populations of African honey bees (Apis mellifera) using new web-based tools. BMC genomics. 2015 Dec;16(1):1–8 pmid:26159619

70. Ito N, Moseley GW, Sugiyama M. The importance of immune evasion in the pathogenesis of rabies virus. Journal of Veterinary Medical Science. 2016:16–0092.
71. Srithayakumar V, Sribalachandran H, Rosatte R, Nadin-Davis SA, Kyle CJ. Innate immune responses in raccoons after raccoon rabies virus infection. Journal of General Virology. 2014 Jan 1;95(1):16–25. pmid:24085257
72. Li J, Faber M, Dietzschold B, Hooper DC. The role of toll-like receptors in the induction of immune responses during rabies virus infection. InAdvances in virus research 2011 Jan 1 (Vol. 79, pp. 115–126). Academic Press. https://doi.org/10.1016/B978-0-12-387040-7.00007-X pmid:21601045
73. Madhu BP, Singh KP, Saminathan M, Singh R, Tiwari AK, Manjunatha V, et al. Correlation of inducible nitric oxide synthase (iNOS) inhibition with TNF-α, caspase-1, FasL and TLR-3 in pathogenesis of rabies in mouse model. Virus genes. 2016 Feb 1;52(1):61–70. pmid:26690069
74. Renaut S, Nolte AW, Bernatchez L. Mining transcriptome sequences towards identifying adaptive single nucleotide polymorphisms in lake whitefish species pairs (Coregonus spp. Salmonidae). Molecular ecology. 2010 Mar;19:115–31. pmid:20331775
75. Moulana A, Anderson RE, Fortunato CS, Huber JA. Selection is a significant driver of gene gain and loss in the pangenome of the bacterial genus Sulfurovum in geographically distinct deep-sea hydrothermal vents. Msystems. 2020 Apr 14;5(2):e00673–19. pmid:32291353
76. Barnard-Kubow KB, Sloan DB, Galloway LF. Correlation between sequence divergence and polymorphism reveals similar evolutionary mechanisms acting across multiple timescales in a rapidly evolving plastid genome. BMC evolutionary biology. 2014 Dec;14(1):1–0.
77. Guo Y, Long J, He J, Li CI, Cai Q, Shu XO, et al. Exome sequencing generates high quality data in non-target regions. BMC genomics. 2012 Dec;13(1):194. pmid:22607156
78. Ellegren H, Smeds L, Burri R, Olason PI, Backström N, Kawakami T, et al. The genomic landscape of species divergence in Ficedula flycatchers. Nature. 2012 Nov;491(7426):756. pmid:23103876
79. Zhao S, Zheng P, Dong S, Zhan X, Wu Q, Guo X, et al. Whole-genome sequencing of giant pandas provides insights into demographic history and local adaptation. Nature Genetics. 2013 Jan;45(1):67. pmid:23242367
80. Attard CR, Beheregaray LB, Sandoval-Castillo J, Jenner KC, Gill PC, Jenner MN, et al. From conservation genetics to conservation genomics: a genome-wide assessment of blue whales (Balaenoptera musculus) in Australian feeding aggregations. Royal Society open science. 2018 Jan 31;5(1):170925. pmid:29410806
81. Morin PA, Martien KK, Taylor BL. Assessing statistical power of SNPs for population structure and conservation studies. Molecular Ecology Resources. 2009 Jan;9(1):66–73. pmid:21564568

82. Kukekova AV, Johnson JL, Xiang X, Feng S, Liu S, Rando HM, et al. Red fox genome assembly identifies genomic regions associated with tame and aggressive behaviours. Nature ecology & evolution. 2018 Sep;2(9):1479–91.
83. Schaefer C, Rost B. Predict impact of single amino acid change upon protein structure. InBMC genomics 2012 Jun (Vol. 13, No. 4, p. S4). BioMed Central.
84. Wagner JL, Hayes-Lattin B, Works JD, Storb R. Molecular analysis and polymorphism of the DLA-DQB genes. Tissue Antigens. 1998 Sep;52(3):242–50. pmid:9802604
85. Scurrell E., Stanley R., & Schöniger S. (2009). Immunohistochemical detection of NOD1 and NOD2 in the healthy murine and canine eye. Veterinary ophthalmology, 12(4), 269–275. pmid:19604345
86. Verfuurden B., Wempe F., Reinink P., van Kooten P. J., Martens E., Gerritsen R., et al. (2011). Severe combined immunodeficiency in Frisian Water Dogs caused by a RAG1 mutation. Genes and immunity, 12(4), 310. pmid:21293384
87. Miao E. A., Andersen-Nissen E., Warren S. E., & Aderem A. (2007, September). TLR5 and Ipaf: dual sensors of bacterial flagellin in the innate immune system. In Seminars in immunopathology (Vol. 29, No. 3, pp. 275–288). Springer-Verlag. https://doi.org/10.1007/s00281-007-0078-z pmid:17690885
88. Mørk T, Prestrud P. Arctic rabies–a review. Acta Veterinaria Scandinavica. 2004 Mar;45(1):1.
89. Kardos M, Shafer AB. The peril of gene-targeted conservation. Trends in ecology & evolution. 2018 Nov 1;33(11):827–39. pmid:30241778

Chapter 4

Minimal signatures of selection in immunogenetic data from eastern small-footed bats (*Myotis leibii*), pre- and post-white-nose syndrome occurrence, implies inherent resistance or tolerance

Preface

A version of this chapter is currently under peer-review

Authors: Tristan M. Baecklund, Michael E. Donaldson, Maarten J. Vonhof, D. Blake Sasse, Christina M. Davy, and Christopher J. Kyle

Contributions: MED, CMD, and CJK conceptualized and designed the study; TMB did all the immunogenetic lab work; MJV provided microsatellite data; TMB and MED did initial processing of raw data; TMB did the data analysis and visualization; TMB wrote the manuscript; all authors critically reviewed the manuscript.

Abstract

White-nose syndrome (WNS) is an infectious disease of hibernating bats caused by the pathogenic fungus *Pseudogymnoascus destructans*. When epizootic, WNS-associated declines were observed in populations of several Nearctic bat species, while other species appeared resistant or tolerant suggesting disparate responses are associated with differing life histories and immunogenetic factors. *Myotis leibii* (eastern small-footed bat) is a species of conservation concern across much of its range, whose rarity and unclear dispersal patterns make it difficult to assess population connectivity and relative impacts associated with WNS. To address these knowledge gaps, *M. leibii* were sampled across their range and profiled at neutral and immunogenetic loci (n = 147 and 160, respectively) to assess levels of genetic diversity and spatial genetic structure relative to pre- and post-WNS occurrence.

Microsatellite marker and amplicon sequence data from immune-related genes yielded three distinct and concordant genetic clusters across the range of *M. leibii*. Analyses within genetic clusters sampled pre- and post-WNS occurrence, identified allele frequency shifts in a limited number of candidate genes, including missense SNPs associated with immune responses to fungal pathogens, but did not show evidence of genetic bottlenecks from selective sweeps.

Given catastrophic demographic declines and genetic bottlenecks associated with WNS in other Nearctic bat populations, in contrast to stable

population trends and weak patterns of selection between pre- and post-WNS samples from this study, we cannot exclude *M. leibii* as being inherently tolerant or resistant to WNS. These data contribute to our understanding of the genetic diversity and structure of this elusive species, and further illustrate the varied responses of Nearctic bat species to WNS.

4.1 Introduction

Species' intra- and interspecific interactions, coupled with usage and dispersal within their environment, are key ecological and evolutionary parameters necessary for understanding how rapidly changing selective pressures threaten evolutionary potential and species persistence (Hohenlohe et al. 2021; Lozier & Zayed 2017; Yildirim et al. 2018). Of the selective pressures exerted on organisms, infectious diseases can quickly elicit changes to the genetic composition and diversity of populations (e.g., McCallum 2008; Epstein et al. 2016). These changes can occur via bottleneck events and selective sweeps, where host/pathogen dynamics are in recurring co-evolutionary cycles of responding to infection and circumventing host immune systems (DeCandia et al. 2018; Santillan et al. 2021).

The spread of infectious disease is exacerbated by globalization, such as human-mediated translocations, which undermine barriers to dispersal and thus expose naïve populations to an increasing array of novel pathogens (Baker et al., 2022; Davy et al. 2017; Fischer et al. 2022; Longo et al. 2014; Ricciardi et al. 2013; Vander Wal et al. 2014). For example, the fungal pathogen *Batrachochytrium*

dendrobatidis, the causative agent of chytridiomycosis, has spread worldwide; leading to extreme population declines in 124 species and the presumed extinction of an additional 90 species (Fisher & Garner, 2020). In some systems, chytridiomycosis susceptibility is influenced by host immunogenetic variants, where conservation efforts aimed to maintain genetic diversity and resistant individuals are valuable in minimizing disease impacts (Oswald et al., 2020). Similar patterns of immunogenetic selection have been observed in mammals, including Tasmanian devils (*Sarcophilus harrisii*) prone to a highly transmissible infectious facial tumour disease leading to population declines of over 60% (see McCallum 2008). In Tasmanian devils, strong selection on immunogenetic variants associated with disease risk were observed, despite limited standing genetic variation (Epstein et al. 2016). These examples demonstrate how immunogenetic data can contextualize the evolutionary potential of a population or species in applying conservation efforts in response to emerging and changing pathogen dynamics.

Integrating data from neutral loci and candidate loci, such as immune-related genes, can clarify how gene flow, genetic drift, and selection shape genetic diversity across the landscape (Kawecki & Ebert 2004; Neaves et al., 2009; Peakall et al., 2003). These data thus provide insights into movement and historical/contemporary selective pressures of species (Davy et al. 2017; Supple & Shapiro 2018; Yi & Latch 2022). Whole genome sequencing (WGS) approaches are increasingly being used to address ecological and evolutionary significant questions (e.g., immune system adaptations unique to bats, Santillan et al. 2021). More

specifically, targeted alternatives to WGS, genotype-by-sequencing (GBS), can provide effective and simplified means to simultaneously study patterns of genetic diversity from neutral and candidate loci (Pina-Martins et al. 2019; Schweizer et al. 2016). For example, Schweizer et al. (2016) found differential signatures of selection on immunogenetic loci between ecoregions of gray wolves using GBS, despite extensive gene flow among populations. These approaches demonstrate the potential of GBS as a tool to directly assess the relative impacts of selection and gene flow on populations in the context of rapidly changing selective pressures on the landscape.

Pseudogymnoascus destructans (*Pd*), the fungal pathogen that causes white-nose syndrome (WNS), was introduced from Eurasia and first documented in North America in 2006 (Blehert et al. 2009; Leopardi, Blake, & Puechmaille, 2015). The introduction of this pathogen resulted in drastic population declines for several hibernating bat species in eastern and central North America (Blehert et al. 2009; Frick et al. 2016; Hoyt et al. 2016; Kovacova et al. 2018; Puechmaille et al. 2011; Wilder et al. 2015). *Pseudogymnoascus destructans* is a persistent environmental fungus that thrives in cold temperatures, thus facilitating disease transmission in winter months when susceptible bats congregate in hibernacula, body temperatures fall within an optimal range for fungal growth, and immune functions are downregulated (Fischer et al. 2022; Langwig et al. 2012; Verant et al. 2012; Whiting-Fawcett et al. 2021). Once infected, bats undergo a cascade of physiological changes (Verant et al. 2014), including increased bouts of arousal

from torpor (Cryan et al. 2010; Moore et al. 2013). Associated loss of fat stores from these disruptions before spring emergence can lead to starvation or death because of exposure (Anderson 2018; Auteri & Knowles 2020; Blehert et al. 2009; Reeder et al. 2012; Verant et al. 2014).

Susceptibility to WNS is influenced by varied environmental, behavioural, ecological, genetic, and individual factors across habitats and species (Cryan et al. 2013; Davy et al. 2020; Hayman et al. 2016); differences that result in the disparate effects of WNS among bat species. For example, eastern populations of little brown bats (*Myotis lucifugus*) experienced drastic population declines (Cheng et al. 2021; Dzal et al. 2011; Francl et al. 2012) and became a focal species for host response studies (Bohn et al. 2016; Warnecke et al. 2013; Wilder et al. 2011; Wilder et al. 2015). Genetic studies of *M. lucifugus* highlight unimpeded host-mediated dispersal of *P. destructans* in relation to potential environmental barriers and revealed underlying genetic adaptations associated with disease outcome (Auteri & Knowles 2020; Davy et al. 2015; Davy et al. 2017; Donaldson et al. 2017; Lilley et al. 2020 Miller-Butterworth et al. 2014; Vonhof et al. 2015; Wilder et al. 2015). In contrast, another Nearctic species, the big brown bat (*Eptesicus fuscus*), demonstrates a resistance or tolerance to the disease (Frank et al. 2014; Moore et al. 2018), which has been attributed to an extremely localized immune response relative to the systemic immune response mounted by *M. lucifugus* and *M. myotis* (Davy et al. 2020). Transcriptomic work has also revealed differing responses to infection between *Myotis* congeners, where clear distension between up and down-regulated

genes between *M. lucifugus* and *M. myotis* were observed (Lilley et al. 2019). These data provide information relevant to understanding patterns of susceptibility to *P. destructans* in other bat species.

The eastern small-footed bat (*Myotis leibii*) is a rare Nearctic species classified as endangered by IUCN (Solari 2018) and provincially endangered in Ontario, Canada (Humphrey 2017; noting there is no such designation federally within the United States or Canada). *Myotis leibii* tolerate colder temperatures than other species, starting hibernation later and emerging earlier (Fenton 1972; Turner, Reeder & Coleman 2011). *Myotis leibii* typically hibernate and roost individually or in small groups, and often roost in crevices or under rocks making them difficult to monitor using standard bat survey techniques (Erdle and Hobson 2001; Johnson et al. 2008; Johnson et al. 2011; Moosman, Anderson & Fraiser, 2017; Moosman et al. 2015;). Analysis of amplified fragment length polymorphisms in *M. leibii* has suggested population genetic structure (i.e., restricted patterns of gene flow across their range), although small sample sizes limit interpretations from these data (N = 10 in Ammerman, Lee & Pfau, 2016). Surveys of *M. leibii* did not detect substantial population declines following the initial occurrence of *P. destructans* (Langwig et al. 2012; Frick et al. 2017; Anderson 2018; Reynolds et al. 2021; O'Keefe et al. 2019; Hooton et al. 2023; Sasse & Perry 2023; but see reported declines in Frick et al. 2015). These observed stable *M. leibii* population trends imply that the species may be resistant or tolerant to WNS supported by reduced pathogen loads and

physiological impacts, respectively (Medzhitov, Schneider & Soares 2012; Råberg, Graham & Read 2009).

The relative rarity of *M. leibii* makes it difficult to monitor population trends because small sample sizes and low recapture rates limit the use of classical mark-recapture methods across large geographic areas. Genomic signatures of population bottlenecks associated with changes in major selective pressures have the potential to complement field data and improve confidence in estimated population trends, such as those inferred following the arrival of WNS in other species (Auteri & Knowles 2020; Donaldson et al. 2017; Gignoux-Wolfsohn et al. 2021; Lilley et al. 2020). A robust genomic assessment of *M. leibii* across its range could quantify population connectivity, evaluate genetic variation and diversity pre- and post-WNS, and clarify if signatures of selection associated with WNS are present in this species.

Herein, we explore neutral and functionally relevant genetic variation to address several knowledge gaps in our understanding of *M. leibii*. Specifically, we use genomic data to test competing hypotheses: 1) that *M. leibii* populations were already resistant or tolerant to WNS on first exposure, predicting similar immunogenetic diversity within populations pre- and post-WNS, or 2) that *M. leibii* populations have undergone selective sweeps, similar to those observed in *M. lucifugus*, predicting that signatures of selection (i.e., shifting immunogenetic diversity) would be observed in populations post-WNS occurrence. We use

microsatellite genotypes to identify neutral genetic structure across the species' range and investigate factors such as gene flow and genetic drift that may undermine/mask signatures of selection imposed by exposure to WNS. Further, we assess sequence data from genomic regions associated with immune responses to screen for signatures of selection pre- and post-WNS exposure. These data provide further insight into the status of *M. leibii* and the consequences of WNS on Nearctic bats as the disease continues to spread westward toward naïve populations.

4.2 Materials and Methods

4.2.1 Sample collection, DNA extraction, DNA quality and quantity assessments

We collected tissue samples from 5 representative locations across the range of *M. leibii* (Fig 4.1; Table S4.1). All tissue samples were collected non-harmfully by taking a 3mm wing or tail membrane biopsy punch following animal care protocols from the institutions of the respective sample contributors. Upon receipt, tissue and DNA samples were stored in a -80 °C freezer prior to DNA extraction. DNA for microsatellite analyses (n=147) and the immunogenetic assay (n=160) were extracted from wing punches using the DNeasy Blood and Tissue Kit (Qiagen), noting 60 samples overlap between the microsatellite and immunogenetic assays. DNA for the immunogenetic assay was assessed for quality using gel electrophoresis (90 V, 45 minutes, 0.8% agarose gel) stained with ethidium bromide, HighRanger 1 kbp DNA ladder as reference (300 – 10,000 bp; Norgen Biotek) and quantified using a Quant-iT PicoGreen dsDNA Assay Kit (ThermoFischer Scientific).

4.2.2 Microsatellite genotyping

We genotyped 147 *M. leibii* samples at ten microsatellite loci using previously described primers developed for other vespertilionids (Paur05 originally from Burland et al. 1998 but using redesigned primers; Table S4.2). Microsatellite amplification was carried out in six reactions (Table S4.2), with cycling conditions consisting of 1 min at 94 °C, three cycles of 30 sec at 94 °C, 20 sec at Ta (54 or 60 °C), and 5 sec at 72 °C, 33 cycles of 15 sec at 94 °C, 20 sec at Ta, and 2 sec at 72 °C, followed by a final extension at 72 °C for 30 min. Annealing temperatures and cycling numbers varied depending on the multiplex (Table S4.2). Fragments were analyzed and scored using GeneMarker software (SoftGenetics LLC, State College, PA).

4.2.3 Analysis of microsatellite data

Basic summary statistics were compiled using the package hierfstat (Goudet 2005) under R v4.1 (R Core Team 2021). Hardy-Weinberg equilibrium (HWE) was tested using both the X^2 approach and the Monte Carlo permutations exact test. Using STRAUTO v.1.0 (Chhatre & Emerson 2017), STRUCTURE analyses were performed with a burn-in length of 50,000 followed by 200,000 iterations for K = 1 through K = 12, with 20 iterations of each K (Pritchard, Stephens, & Donnelly, 2000). Under R v4.1 (R Core Team 2021), packages pophelper and pophelpershiny (v2.1.1; Francis 2017) were used to estimate the most likely value of K using the Evanno method (Evanno, Regnaut, & Goudet, 2005; Fig S4.8-12) and to visualize and combine

results. Discriminant Analyses of Principal Components (DAPC) were also performed on the dataset as implemented using adegenet v2.1.7 (Jombart 2008) and ape v5.6-2 (Paradis, Claude, & Strimmer 2004). Here, K was estimated using sussesive K-means but also investigated using K equal to the number of representative geographical sampling locations to evaluate which number of identified clusters seemed most plausible in context of the ecology of the species and hypotheses being tested. The function *xvalDapc* was used to identify the number of principle components achieving lowest mean-squared error and retained for discriminant analysis in an attempt to not overfit these data. Identified genetic clusters from these analyses were then used as *a priori* groupings to inform subsequent immunogenetic data analyses.

4.2.4 Library preparation, sequence capture, and high-throughput sequencing of immunogenetic assay

DNA libraries (n = 160) were created following the SeqCap-EZ HyperCap UGuide V1.0 (Roche) protocol using the Kapa HyperPlus Kit (Roche) with the following specifications: i) samples were treated with 5 µL conditioning solution during fragmentation, ii) TruSeq HT Dual-Index Adapters (Integrated DNA Technologies) were used in place of SeqCap Adapter Kits A and B (Roche), iii) Illumina P5 and P7 primers (Integrated DNA Technologies) were substituted in place of pre-LM-PCR Oligos 1 & 2 (Roche), and iv) seven cycles were used for pre-LM-PCR. Before pooling and target enrichment, individual DNA libraries were assessed for quality and

concentration using ethidium bromide-stained gel electrophoresis as outlined above.

Equimolar amounts of each DNA library were pooled into a 1 µg DNA multiplex prior to target enrichment of immunogenetic loci as described by Donaldson et al. (2017) using the designed SeqCap EZ Developer Library probe (see Table S4.3 for targeted genes). Modifications to the enrichment protocol included: i) replacement of NimbleGen Multiplex Hybridization Enhancing Oligo Pool (Roche) with 2 µL xGen Universal Blockers – TS Mix (Integrated DNA Technologies), ii) NimbleGen SeqCap EZ Developer Reagent (Roche) was used in place of NimbleGen COT Human DNA (Roche) during hybridization sample preparation, and iii) hybridizations were carried out at 47°C for 20 hours. Final product assessments were conducted with a bioanalyzer for the target-enriched multiplex before sequencing on an Illumina MiSeq V3 run using 2x300 bp reads (Advanced Analysis Centre Genomics Facility, University of Guelph).

4.2.5 Sequence alignment, variant annotation, and filtering

Raw data were first assessed using FastQC (Andrews 2010) to inspect for initial issues with these data that would warrant further investigation, such as per base/sequence quality scores, sequence duplication, and adapter content. These raw data were then processed with Trimmomatic (Bolger, Lohse, & Usadel 2014) to trim reads for quality and length using the following parameters: i) removing bases at the start and end of reads not surpassing a quality of 3 as determined by

Trimmomatic, ii) SLIDINGWINDOW of 4 bp and a quality of 15, and iii) a minimum sequence length of 36 bp. Paired-end reads for each of the 160 samples were then aligned to the *M. lucifugus* reference genome (Myoluc2.0) using the bwa-mem command in Burrows-Wheeler Aligner (Li 2013). Sequence metrics were assembled using SAMTOOLS v1.9 (Danecek et al. 2021) and GATK v4.1.0.0 best practices pipeline was used to remove duplicate sequences, identification and removal of secondary alignments (potential paralogs or copy variants), variant discovery, genotyping, and variant recalibration (DePristo et al. 2011; McKenna et al. 2010; Poplin et al. 2017; Van der Auwera & O'Connor 2020). We expanded the GATK standard variant discovery by implementing the following hard filters: i) QualByDepth < 2.0, ii) FisherStrand > 60.0, iii) RootMeanSquareMappingQuality < 40.0, iv) MappingQualityRankSum < -12.5, v) ReadPosRankSum < -8.0, and vi) to only retain biallelic SNPs. We further filtered the data to ensure less than 20% missing data (max-missing) and a minor allele frequency (MAF) ≥ 2.5% using VCFtools (Danecek et al. 2011) and removed any loci presumably on the X-chromosome. When required during subsequent analyses of data subsets, SNPs were pruned for linkage disequilibrium utilizing BCFtools v1.9 (Danecek et al. 2021) with max-LD = 0.2 and retaining a single locus within a 500 kbp window.

4.2.6 Analyses of SNPs from the immunogenetic assay

Variants passing initial filtering parameters (MAF, max-missing, biallelic) were used as input for F_{ST} outlier identification using PCAdapt (Luu, Bazin, and Blum 2017), OutFLANK (Whitlock & Lotterhos 2015), Arlequin (Excoffier & Lischer 2010), and

Bayescan (Foll & Gaggiotti 2008), with an adjusted p-value of ≤ 0.05. Filtering followed methods described in Baecklund et al. (2021), detailed further in File S4.1. F_{ST} outlier identification was performed on the dataset consisting of all individuals (N=160) and then subsets of bats from Arkansas and eastern Ontario, which encompassed pre- and post-WNS occurrence (pre-WNS sample size n = 9 (Arkansas) and n = 10 (eastern Ontario); post-WNS sample size n = 31 (Arkansas) and n = 21 (eastern Ontario). Data were visualized using STRUCTURE and DAPC analyses in the same manner as the microsatellite data. In order to distinguish between signatures of selection relative to patterns of genetic drift in identified outlier SNPs from within the pre- and post-WNS occurrence datasets, we generated a 95% confidence interval of the Wright-Fisher expectation (https://gitlab.com/WiDGeT_TrentU/; File S4.1). Intervals were based on 1,000 iterations of the model, where a)generations are non-overlapping, population size is static, b) randomly sampled generation time was set to 2-10 years, and c) using the pre-WNS sampling size and allele counts to determine N_e that was assumed constant over the simulation (File S4.1). While some of these model parameters are likely not maintained over our sampled time periods, we used raw frequency data data to avoid inflating population sizes and presumed N_e (while maintain the same allele frequencies) to produce more conservative estimates of influences associated with genetic drift. Using these Wright-Fisher simulations, each SNP was assigned as a potential outlier, where the post-WNS allele frequency fell outside of the generated 95% confidence interval, or as drift where post-WNS allele frequency

fell within the interval. The selection coefficient (s) was estimated following Thompson et al. (2019) and Haworth et al., (2021), for any SNP from these two datasets that was identified as an F_{ST} outlier, had a frequency change unlikely to be due to drift, and resulted in a missense mutation. Seleciton coefficients were estimated using dominant, codominant (heterozygous), and recessive models under the assumption of HWE, providing insights on the relative fitness of the three possible genotypes based on the investigated biallelic markers. Seleciton coefficients =1 demonstrate purging of a genotype, where s = 0, is representative of no selection on the genotype.

4.3 Results

4.3.1 Microsatellite marker analyses

We generated 147 complete *M. leibii* profiles at all ten loci genotyped (Table S4.4). Across all sampled regions, the number of observed alleles at each locus ranged from 12-31 (Table 4.2). Based on the observed clustering from Discriminant Analyses of Principal Components (DAPC) and STRUCTURE analyses, we infered three likely clusters, largely corresponding to geography: 1) Arkansas, 2) North Carolina and West Virginia, and 3) New York/New Hampshire/Vermont and Ontario (Fig 4.2, Fig S4.8), where hierarchical analyses did not distinguish further clustering within the second and third clusters (Fig S4.1, S4.11). These three identified clusters were therefore used to inform our grouping of samples for subsequent analyses. Observed mean heterozygosity ranged from 0.80 – 0.87 across the ten microsatellite loci for each genetic cluster (Table 4.2) with an average intra-population diversity

(Hs) ranging from 0.82 – 0.87. Inter-population diversity (Dst) was 0.03 (Table 4.2). F_{ST} estimates, following Nei (1973), were 0.05 between genetic clusters one & two; 0.07 between clusters one & three; and 0.03 between clusters two & three. We also estimated F_{ST} following Weir & Cockerham (1984) with estimates of 0.05, 0.07, and 0.03 among the same cluster comparisons, respectively. Under the Monte Carlo permutations exact test, locus b22 and MS3DO2 were not in Hardy-Weinberg equilibrium (HWE) in cluster 2, along with MME24 in cluster 3 (Fig S4.3). Based on a X^2 test, b22 and MME24 remained out of HWE in their respective clusters in addition to MMF19 in cluster 2 and H10E11 in cluster 3 (Fig S4.3). No locus was out of HWE across all populations, as such all loci were retained for further interpretation of these data.

4.3.2 Immunogenetic assay read and SNP filtering

FastQC identified no apparent issues with raw sequence data. After Trimmomatic, and the initial GATK processing pipeline, an average of ~ 222,393 reads (range of 74,522 – 430,600) from each library remained, of which ~ 99.75% mapped to the *M. lucifugus* reference genome. We observed an average depth of coverage of ~ 86X on a per-individual and per-locus basis. Variant discovery resulted in a total of 15,188 SNPs, with 14,212 passing GATK hard filtering, where 13,311 were bi-allelic. After all other filtering steps (MAF, max-missing), we retained 4,373 immunogenetically relevant SNPs for subsequent analyses.

4.3.3 Immunogenetic assay analyses across the M. leibii range

The dataset of 4,373 immunogenetic SNPs was pruned for linkage disequilibrium leaving 105 SNPs to initially assess spatial immunogenetic structure prior to outlier analyses (Fig 4.3, S4.9; Table S4.5). From these 105 SNPs, DAPC identified three likely clusters (Fig 4.3a, b) that aligned with the three clusters identified using neutral microsatellite markers. STRUCTURE analysis found two immunogenetic clusters; Arkansas and a pooled geographic cluster of New York/New Hampshire/Vermont/Ontario (Fig 4.3c, d). Hierarchical analyses identifying K = 3 as most likely K within this pooled genetic cluster, segregated this cluster further between New York/New Hampshire/Vermont, eastern Ontario, and southwestern Ontario (Fig S4.2, S4.12).

Investigating the same dataset of 4,373 immunogenetic SNPs, specifically screening for outlier loci distinguishing the genetic clusters, we identified 114 F_{ST} outlier SNPs. After performing linkage disequilibrium pruning, a final subset of 16 F_{ST} outlier SNPs was retained for genetic structure visualization (Table S4.5). An optimal K = 2 clusters was identified from the 16 F_{ST} outlier SNP subset that produced consistent patterns of genetic structure observed in the initial 105 LDP SNPs assessment without outlier identification (Fig 4.3, S4.9). Within this subset of 16 F_{ST} outlier SNPs, three were missense mutations (base substitution changes encoded amino acid) associated with TLR6, NOS2, and LBP genes.

4.3.4 Immunogenetic assay analyses of pre- and post-WNS occurrence data

Geographic locations encompassing both pre- and post-WNS occurrence (sampled 2004-2018), included 71 samples (n = 40 from Arkansas and n = 31 from eastern Ontario), with 3,711 immunogenetically relevant SNPs passing filtering parameters. We identified a subset of 23 linkage disequilibrium pruned F_{ST} outlier SNPs to visualize patterns of immunogenetic structure before and after the arrival of WNS (Table S4.5). Both DAPC and STRUCTURE analyses indicated an optimal clustering of K = 2, distinguishing between geographic locations (Fig 4.4, S4.10). From hierarchical DAPC analyses within each sampled geographic location, we inferred patterns suggestive of weak genetic structure within locals that remained discernable when visualizing data for both populations at K = 4 (noting the K = 4 model was less supported than K = 2; Fig 4.4, S4.10).

Given the observation that immunogenetic data predominantly demonstrated geographic structure, we performed within-cluster outlier analyses to determine if similar or differing outliers best explained the weak patterns of selection observed within each of the populations. We analyzed 2,999 and 2,520 SNPs remaining after MAF and missing data filters for the Arkansas and eastern Ontario clusters, respectively. From these SNP subsets, we identified 28 and 29 F_{ST} outlier SNPs associated with WNS occurrence in the Arkansas and eastern Ontario clusters, respectively, noting only one outlier overlapped between regions (Tables S4.6). Based on the simulated confidence interval of the expected change in allele frequency from the pre-WNS occurrence samples, these outlier loci are unlikely to

be caused by genetic drift (Table S4.7). Changes in allele frequencies between pre- and post-WNS sampling at each locus ranged from 0.179 to 0.412 in Arkansas and 0.107 to 0.517 in eastern Ontario (Tables S4.6). To better understand if selection was imposed on *M. leibii* at each site, we estimated selection coefficients under a dominant model. Coefficients ranged from 0.17 (± 0.11) to 0.74 (± 0.19) for each of the 28 SNPs identified in the Arkansas dataset and 0.11 (± 0.08) to 0.97 (± 0.02) in the eastern Ontario dataset of 29 SNPs, indicative of a selective pressure exerted on *M. leibii* (Table S4.7). Between both outlier SNP subsets one non-sense mutation associated with the TAP1 gene was observed (position 89,998 on scaffold GL430751). A single outlier SNP consistent between both data sets was associated with the HRG gene (Table S4.6). Of those outliers SNPs within coding regions, 6 were missense SNPs in the eastern Ontario subset associated with the DRB1-Exon2-like gene (n = 1), the HLA-DPB1 gene (n = 1), the TICAM1 gene (n = 1), the TLR1 gene (n = 1), and the TLR6 gene (n = 2). The seven missense SNPs in the Arkansas subset were associated with DLA-DMB (n = 2) and TAP1 (n = 5). Simulated coefficients of selection for outlier missense SNPs were among the highest observed in our dataset and demonstrated rapid shifts in allele frequencies (Figs S4.6-7; Table S4.6). Most SNPs regardless of designated status as an identified outlier or not, were in genomic regions putatively associated with promoter/regulatory elements where only a smaller proportion were identified within coding regions or leading to larger impacts (Figs S4.4-5). In total across both the Arkansas and eastern Ontario datasets, only five SNPs led to high predicted impact consequences, all of which conveyed a stop

codon gain. One of these was associated with the IL13 gene, and another with TAP1, while the remaining three had no gene symbol available (novel transcripts). Allele frequency shifts at identified high impact sites fluctuated ~<10% between years, with the exception of a stop gain at a novel transcript in the eastern Ontario subset where the substitution resulting in the stop gain increased by 30%.

4.4 Discussion

We sought to explore population connectivity and the relative impacts of white-nose syndrome (WNS) on the rare and elusive bat species, *Myotis leibii*. Using both neutral and immunogenetic data from across the species' range we found clear genetic clustering between sampled geographic regions. Immunogenetic data revealed a small number of outlier SNPs that demonstrated weak patterns of genetic structure within two geographic populations but it remains unclear whether these patterns are a result of WNS, relatedness, or a byproduct of geographic and temporal sampling. These results are broadly consistent with the hypothesis that *M. leibii* were inherently resistant or tolerant to WNS on first exposure, and also align with observed, stable demographic trends in this species irrespective of WNS presence (Langwig et al. 2012; Francl et al. 2012; Frick et al. 2017; Anderson 2018; O'Keefe et al. 2019; Sasse et al. 2023; Hooton et al. 2023).

4.4.1 Microsatellite data informs population management

Neutral microsatellite markers identified three genetic clusters of *M. leibii* along a latitudinal gradient (Fig 4.2), consistent with findings of Ammerman, Lee & Pfau

(2016). Hierarchical analyses of these data did not support further sub-structuring among identified genetic clusters (Fig 4.2, Fig S4.1, S4.8, S4.11). These neutral genetic data provide a baseline evaluation of genetic diversity and structure that contextualize patterns of selective pressure (Kawecki & Ebert 2004). The restricted gene flow observed between distinct genetic clusters in this system suggests that population extirpations or drastic declines would not be quickly countered by natural genetic rescue.

4.4.2 Immunogenetic data across the M. leibii range

Broad-level immunogenetic structure across the range of *M. leibii* identified 105 linkage pruned SNPs that mirrored neutral genetic structure and sample provenance. Relative to microsatellite analyses, Discriminant Analysis of Principal Components (DAPC) of immunogenetic data demonstrated similar but stronger clustering patterns. Hierarchical analyses identified further segregation of the cluster encompassing New York/New Hampshire/Vermont/Ontario subdividing Ontario into eastern Ontario, and southwestern Ontario clusters. Analyzing candidate loci most likely to differentiate genetic clusters from the original dataset (16 outlier SNPs after linkage pruning), revealed no difference in the observed genetic clustering of samples. These same outlier analyses yielded several loci with frequency shifts of variants between clusters with the potential to alter underlying immune system functions via altered protein structure (Schaefer & Rost 2012). Of the identified outliers with missense mutations, two (TLR6 and LBP) were associated with genes notable in their response to fungal infections. The TLR6 gene has been

associated with recognition and response to both exogenous pathogenic and commensal fungi (Wang et al., 2014). The LBP gene is commonly associated with immune responses to bacterial infections (Schumann et al., 1990; Schumann 2011), but is also part of the signalling pathway leading to the expression of the NOS2 gene in the presence of bacteria and fungi leading to the activation of natural killer cells (Bogdan, Rollinghoff & Diefenbach 2000; Lowenstein & Padalko 2004). While these gene associations are of potential interest in context of WNS, we postulate that they are most likely reflective of local adaptations between populations with limited gene flow.

4.4.3 Immunogenetic data in context of pre- and post-WNS occurrence

Bats sampled from the Arkansas and eastern Ontario regions included both pre- and post-WNS occurrence individuals. Between these regions, the model of K = 2 genetic clusters was most supported, where geography appeared the predominant factor clustering pre- and post-WNS occurrence immunogenetic data. Within both areas, under the K = 2 model, DAPC analyses of identified outlier loci suggested weak genetic clustering aligning with pre-and post-WNS occurrence. These discrete patterns remained discernable in DAPC analyses of both populations under the K = 4 model, where geography remained the prominent factor clustering samples. Only one SNP, associated with the HRG gene, was identified as an F_{ST} outlier in both Arkansas and eastern Ontario *M. leibii* populations with a known innate immune system antifungal role (Rydengard et al., 2008). The HRG-associated SNP is located in a putative regulatory region that could alter expression of the gene; however,

discerning the exact ramification of this mutation would require further investigation.

Although we identified only a few missense mutations within outlier SNP subsets from the immunogenetic loci, these data have links to fungal pathogen immune responses. Between the two sets of missense SNPs (from Arkansas and eastern Ontario), most associated genes (HLA-DPB1, DRB1e2-like, DLA-DMB, TLR1, TLR6) are related to the immune system's capacity to recognize fungal pathogens and subsequently mount a defence (Ekblom et al., 2007; Savage & Zamudio 2011; Noreen & Arshad 2015; Wang et al., 2014). Beyond the aforementioned SNPs recognizing fungal pathogens, two other genes with missense polymorphisms were identified with potentially relevant functions, TICAM1 and TAP1. The TICAM1 gene encodes for a protein that acts as an adaptor for both TLR3 and TLR4, subsequently facilitating dendritic cell maturation that can influence aspects such as MHC responses and inflammation (Seya et al., 2005). The TAP1 gene is responsible for the movement of antigenic peptides and presentation to MHC molecules, although more commonly associated with defence against viral infections (Praest et al., 2018; Ritz et al., 2003). Notably, we also observed a nonsense mutation (base substitution causes a stop codon gain) associated with the TAP1 gene in the Arkansas subset that could result in a loss of protein function from premature termination during transcription. While codon reassignments (including stop codons) are common across taxa (Belinky et al., 2021; Ivanova et al., 2014; Massey 2017), the TAP1 mutation may have no detrimental effect as suppressor tRNAs can rectify nonsense

SNPs (Ward et al., 2024). However, assessing the exact impact of the TAP1 mutation is beyond the scope of our study. The observed missense SNPs identified from both Arkansas and eastern Ontario datasets displayed coefficients of selection and frequency shifts suggestive of selection rather than genetic drift (Figs S4.6-7; Table S4.7).

Inconsistent patterns of selection between pre-and post-WNS occurrence data from Arkansas and eastern Ontario align with literature from other bat species with similar intra-population patterns of allele frequency and diversity shifts associated with disease (Gignoux-Wolfsohn et al. 2021). These incongruent patterns could result from selective pressures acting on different subsets of standing genetic variation in each population by the same selective pressure (e.g., WNS), different selective pressures other than WNS, stochastic effects (e.g., genetic drift), or as a matter of sampling bias such as from increased relatedness among sampled regions. In the case of *M. leibii*, geography strongly predicts observed genetic structure, where shifts in allele frequencies (unexplained by genetic drift) from immunogenetic loci were observed within regions relative to pre-and post-WNS time points. The sample design comparing pre- and post-WNS sampled bats revealed weak patterns of genetic structure and gene associations with fungal pathogen resistance, thus we can not exclude WNS as the selective pressure driving the observed differences between populations. That said, the sampling scheme also biases towards any differences observed being associated with WNS when other

selective pressures may be relevant between these geographically and temporally separated sample points.

4.4.4 Myotis leibii resistance or tolerance to WNS

Myotis leibii have been presumed to be resistant or tolerant to WNS based on observed population demographic stability/increases relative to other hibernating Nearctic bats (Langwig et al. 2012; Frick et al. 2017; Anderson 2018; Sasse et al. 2023). This presumption may be undermined by behavioural and ecological characteristics of *M. leibii* that can mask true population trends, as the vast majority of hibernation sites (underneath rocks and in cracks and crevices) remain unmonitored for this species. Further evidence of resistance or tolerance to WNS comes from the prevalence and fungal loads of *P. destructans* on *M. leibii* that are similar to those of other, less susceptible, small-bodied bat species (e.g., *E. fuscus*; Bernard et al. 2017). Our genomic data provide further support for the hypothesis of inherent resistance or tolerance to WNS in *M. leibii*, given observations of stable population trends and only weak selection within genetic clusters that cannot be conclusively linked to any particular selective pressure based on available data.

4.5 Conclusions

Assessing genetic variation within neutral and immunogenetically relevant loci provides a snapshot of host/pathogen dynamics and facilitates investigations of underlying mechanisms in responding to disease. Data presented herein imply that *M. leibii* were resistant or tolerant to WNS upon introduction. The lack of evidence

indicative of population declines in this species, while suspected, remained uncertain given the difficulty associated with generating a census of *M. leibii* in the field. Genomic tools have been used to identify cryptic declines in other systems; however, we implement these tools to highlight relatively stable population trends over an otherwise catastrophic decline of congeneric species. WNS does not seem to exert strong selective pressures on *M. leibii* that would present a conservation concern requiring management interventions despite continued *P. destructans* (*Pd*) presence.

The assay implemented herein targets immunogenetic data and identifies weak patterns of selection within geographic populations, although the exact driver of these patterns remains unknown. It is possible, however, that there are other genes, or regulatory elements, in *M. leibii* that influence the immune response that remained unexplored by this assay. Although in this case, given observed conservation of immunogenetic genes across mammals (Lindberg, Propert, & Pignolo 2015), we expect to have obtained reasonable coverage of a wide breadth of genomic regions that may have influenced *M. leibii* response to *Pd* exposure. Similarly, alternative patterns of genetic diversity and structure may exist in genomic regions unassociated with the immune system, but instead associated with other biological functions. For example, relative to other Nearctic bats *M. leibii* undergo frequent and typical arousals from torpor, which are unassociated with the introduction *Pd* and spread of WNS, to forage during winter months (Bernard et al. 2021; Jackson, Wilcox, & Bernard 2022). These typical arousals could have reduced

the impact of WNS on this species as disruptions to hibernation cycles were already common and therefore may not act as an additive pressure exerted on *M. leibii*. As WNS spreads across western North America and naïve populations are exposed, these populations may also mitigate the impacts of WNS (or other pressures) through varied physiological or behavioural adaptations relative to their eastern conspecifics as hypothesized for *M. lucifugus* (Blejwas et al. 2023). While impractical, given the conservation status of this species, *Pd* challenge experiments would offer a direct avenue to test *M. leibii* for inherent resistance or tolerance to WNS as assessed in other bat species (Frick et al. 2022; McGuire, Mayberry, & Willis 2017).

The introduction of *Pd* led to a rapid spread of WNS across hibernating Nearctic bat populations, resulting in disparate population trends among species several years after disease occurrence. In *M. lucifugus*, some populations experienced drastic declines followed by recovery (Maslo et al. 2015; Frick et al. 2017; Dobony and Johnson 2018; Ineson 2020). In contrast, *Myotis septentrionalis* (northern long-eared bat) populations may not recover from exposure to this invasive pathogen (Francl et al. 2012; Reynolds et al. 2016). The apparent high survival of *M. leibii* following the introduction of *Pd* and minimal apparent genetic consequences, contrasts with findings from congeners that experienced severe demographic declines, pronounced signatures of selection, and clear host response changes based on immunogenetic and transcriptomic data (Auteri & Knowles 2020; Davy et al. 2017; Donaldson et al. 2017; Cheng et al. 2021; Francl et

al. 2012; Gignoux-Wolfsohn et al. 2021; Lilley et al, 2019; Lilley et al, 2020).Based on the conserved immunogenetically relevant loci analyzed herein, we provide evidence of no extensive temporal immunogenetic clustering, nor major shifts in genetic diversity in *M. leibii* that would be expected from disease-mediated population declines.

Sample Location	Samples used in sequence capture	Pre- or Post-WNS designation	Spatial Designation
Arkansas	9	pre	AR
New York, New Hampshire, and Vermont	21	pre	NY/NH/VT/ON
North Carolina	21	pre	NC/VA/WV
Virginia and West Virginia	25	pre	NC/VA/WV
southwestern Ontario	0	pre	NY/NH/VT/ON
eastern Ontario	10	pre	NY/NH/VT/ON
Arkansas	31	post	AR
New York, New Hampshire, and Vermont	0	post	NY/NH/VT/ON
North Carolina	0	post	NC/VA/WV
Virginia and West Virginia	0	post	NC/VA/WV
southwestern Ontario	22	post	NY/NH/VT/ON
eastern Ontario	21	post	NY/NH/VT/ON
Total	160		

Table 4.1. *Myotis leibii* sampling scheme for targeted immunogenetic assay identifying spatial genetic patterns pre- and post-WNS occurrence. Included are approximate spatial sampling locations, sample numbers, temporal designation to the occurrence of WNS and spatial designations determined by microsatellite analyses and used for subsequent analyses on these data.

Genetic Cluster	Range of alleles per locus	H_o	H_e	H_s	D_{ST}	F_{IS}	F_{IS} 95%CI	Number of Individuals
AR	4-12	0.87	0.79	0.82	--	-0.07	-0.14 to -0.014	17
NC/WV	12-26	0.85	0.87	0.87	--	0.02	-0.01 to 0.05	93
NY/NH/VT/ON	7-16	0.80	0.81	0.82	--	0.04	-0.01 to 0.07	37
Overall	12-31	0.84	0.87	--	0.03	0	-0.05 to 0.04	147

Table 4.2. Summary statistics for three *Myotis leibii* genetic clusters identified using microsatellite loci for estimating genetic connectivity among sampled regions. Summary statistics for each identified genetic cluster, and across all populations. Identified are the number of alleles, observed and expected heterozygosity, intra- and interpopulation diversity (H_S and D_{ST} respectively), and estimates of F_{IS}. Values are averaged across the ten microsatellite loci used in these analyses.

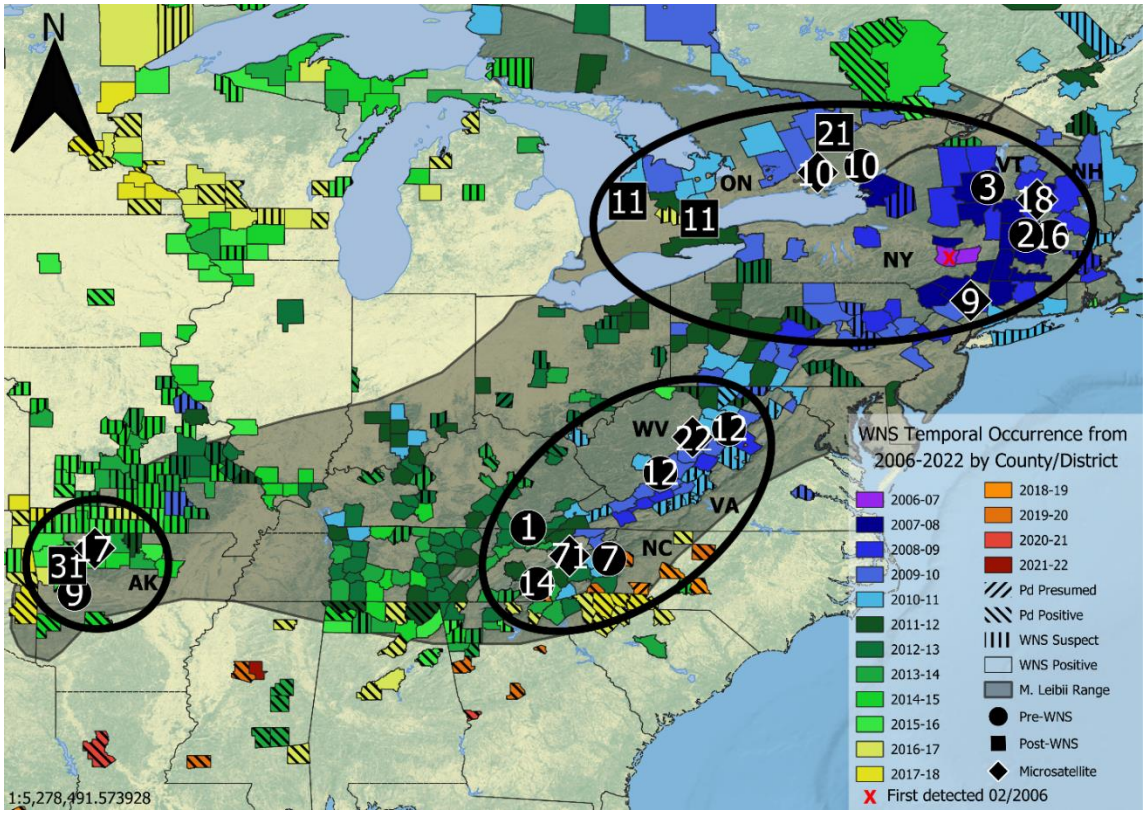

Fig 4.1. Temporal and spatial occurrence of white-nose syndrome overlaid with distribution and sampling scheme of *Myotis leibii*. Figure depicts changing patterns of spread and occurrence of *P. destructans*, and associated disease, WNS, since its introduction in 2006. Sampling included *M. leibii* bats across their range for microsatellite analyses (black-diamonds) and sequencing of immunogenetic coding-regions of the genome pre- (black-circles) and post-WNS (black-squares) occurrence. Numbers within shapes indicate close-proximity sample sizes at each location. Ellipses indicate identified genetic clusters using both microsatellite and immunogenetic datasets where results were essentially mirrored. Map created using Natural Earth and contains information licensed under the Open Government License – Ontario, The United States/U.S. Department of Commerce, The United States/U.S. Fish & Wildlife Service, and the Creative Commons Attribution 4.0 International License – Quebec; created July 2022.

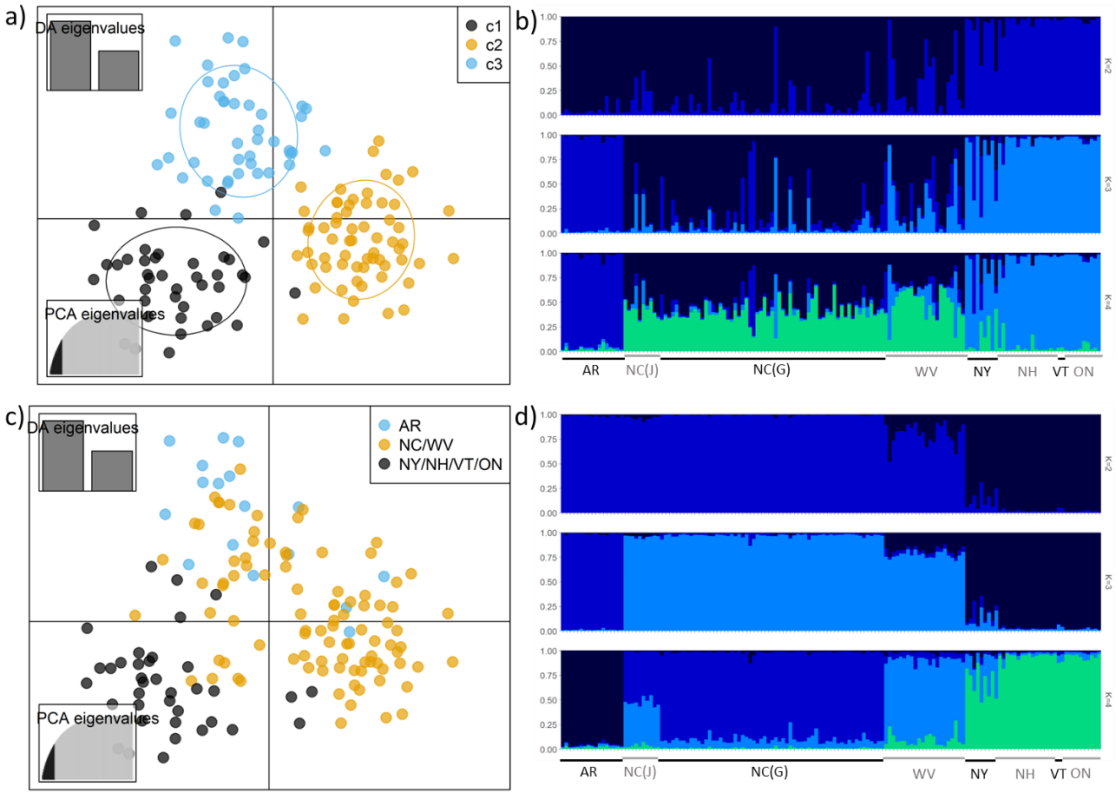

Fig 4.2. Genetic structure analyses of 10 microsatellite loci for 147 *Myotis leibii* indicates three neutral genetic clusters across their range. Inserts a) and b) illustrate discriminant analyses of principal components, where a) displays the inferred 3 clusters of the samples and b) displays the same clustering with samples colour-coded based on geographic origins. Inserts c) and d) provide bar plots of estimated population assignments using STRUCTURE for K = 2 through K = 4 where each partitioned vertical bar represents an individual's proportional membership to the inferred populations. Insert c) provides visualization of STRUCTURE results with LOCPRIOR = 0 and d) with LOCPRIOR = 1; noting K = 3 was the most supported by the Evanno method (Evanno, Regnaut, & Goudet, 2005). Further delineation of samples from the K = 3 clusters were unsupported by hierarchical analyses. AR = Arkansas; NC(J) = North Carolina (Jackson County); NC(G) = North Carolina (Graham County); WV = West Virginia; NY = New York; NH = New Hampshire; VT = Vermont; ON = Ontario).

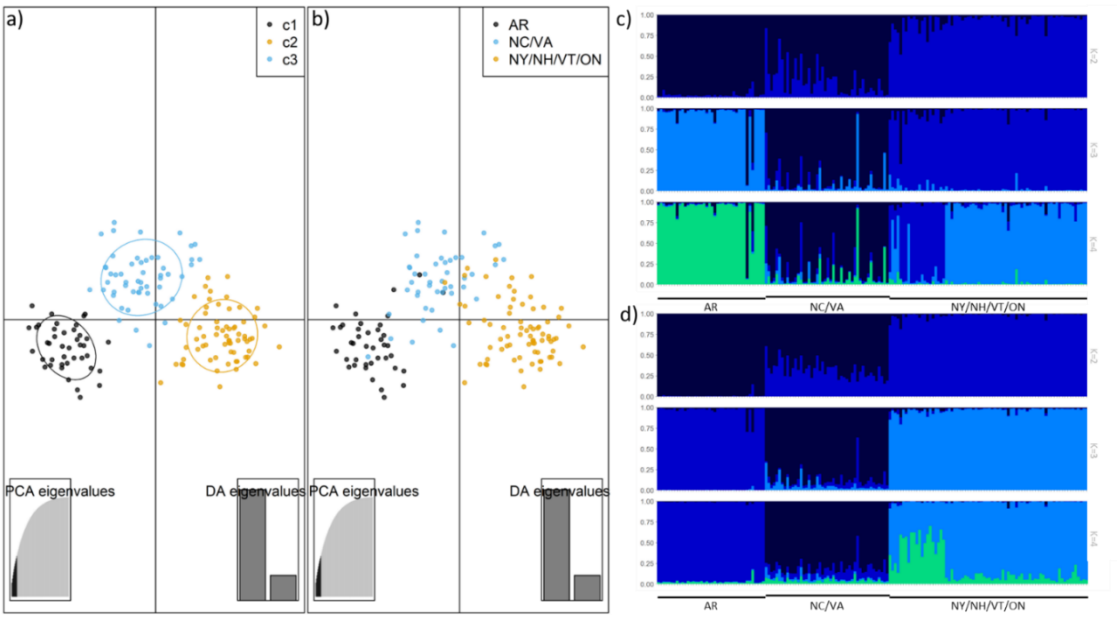

Fig 4.3. Genetic structure of *Myotis leibii* reveals spatial immunogenetic structure consistent with neutral genetic structure, showing three clusters. From a total of 4, 373 identified immunogenetically relevant SNPs, a subset of 105 SNPs remained after linkage-disequilibrium pruning that were used for visualization. Discriminant Analyses of Principal Components were performed and visualized, where a) displays the inferred clustering of the samples and b) displays the same clustering with the samples colour-coded based on geographic origin (AR = Arkansas; NC/VA/WV = North Carolina, Virginia, and West Virginia; NY/NH/VT/ON = New York, New Hampshire, Vermont, and Ontario). Bar plots (c and d) show estimated population assignments using STRUCTURE for K = 2 through K = 4, noting the K = 3 model was most supported by the Evanno method (Evanno, Regnaut, & Goudet, 2005). The K=4 model, with (insert c) and without LOCPRIOR = 1 (insert d), shows some evidence of further genetic partitioning within the NY/NH/VT/ON cluster, which was further supported by a hierarchical analysis of the cluster. Each partitioned vertical bar represents an individual's proportional membership to the inferred populations.

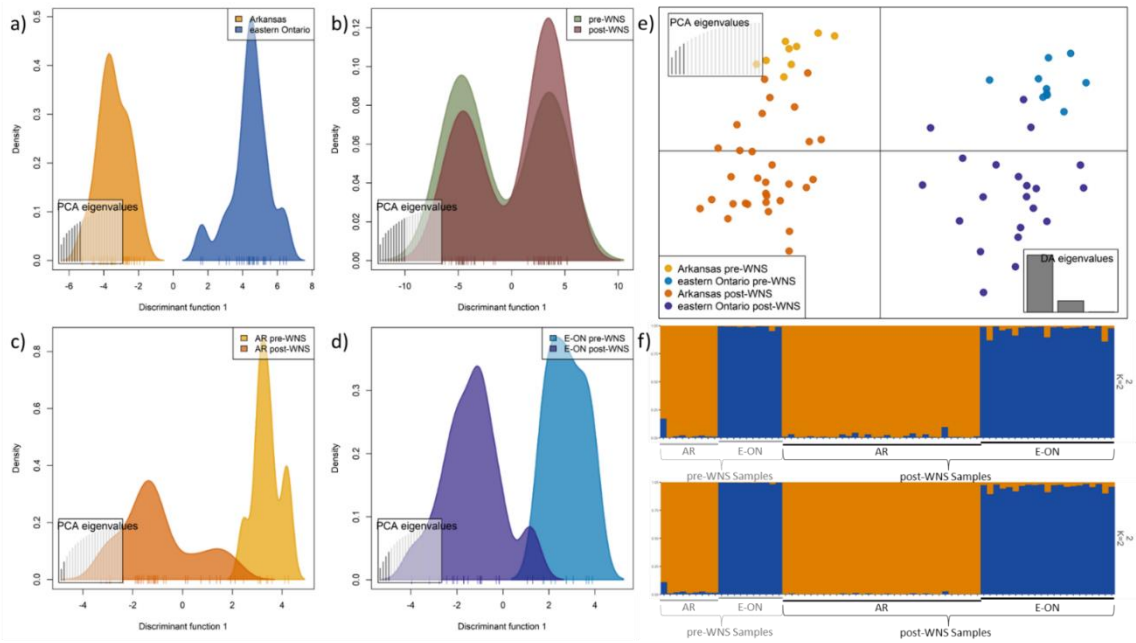

Fig 4.4. Genetic clustering of two *Myotis leibii* populations sampled pre- and post-WNS exposure highlights a prominent geographic influence, and subtle selective patterns associated with the disease. Figure demonstrates consistent patterns of geographic clustering, yet with further subtle intra-population structure correlated with WNS exposure, unattributed to genetic drift, and therefore suggestive of weak selection. Analysis of immunogenetically relevant SNPs from sampled from 2004-2018 in Arkansas and eastern-Ontario, coinciding with before introduction of the disease into North America several years post-introduction produced a dataset of 23 identified outlier SNPs pruned for linkage disequilibrium. Discriminant Analyses of Principal Components identified the optimal number of K clusters = 2, where a) displays the clustering of the 71 samples based on provenance, b) displays the same 71 samples visualized based on WNS occurrence, c) shows the clustering within the 40 samples from Arkansas in context of WNS occurrence, d) displays the WNS occurrence clustering within 31 samples from eastern Ontario, and e) shows k = 4 clustering of the 71 samples where the partition between geographic locations and subtle clustering within populations can be observed. Visualized in insert f) are STRUCTURE plots, where each partitioned vertical bar represents an individual's proportional membership to the inferred populations and the K = 2 model was most supported based on the Evanno method. Results using LOCPRIOR = 0 and LOCPRIOR = 1 are in the top and bottom on the insert, respectively. AR = Arkansas; E-ON = eastern-Ontario; where grey labels indicate samples pre-WNS exposure, and black labels indicate sampled post-WNS exposure.

4.6 References

1. Anderson, B. N. (2018). Immune Function and Metabolism of Hibernating North American Bats with White-Nose Syndrome.
2. Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
3. Ammerman LK, Lee DN, Pfau RS. Patterns of genetic divergence among *Myotis californicus*, *M. ciliolabrum*, and *M. leibii* based on amplified fragment length polymorphism. Acta Chiropterologica. 2016 Dec 30;18(2):336-46.
4. Auteri, G. G., & Knowles, L. L. (2020). Decimated little brown bats show potential for adaptive change. Scientific reports, 10(1), 1-10.
5. Baker, R. E., Mahmud, A. S., Miller, I. F., Rajeev, M., Rasambainarivo, F., Rice, B. L., ... & Metcalf, C. J. E. (2022). Infectious disease in an era of global change. Nature Reviews Microbiology, 20(4), 193-205.
6. Baecklund, T. M., Morrison, J., Donaldson, M. E., Hueffer, K., & Kyle, C. J. (2021). The role of a mechanistic host in maintaining arctic rabies variant distributions: Assessment of functional genetic diversity in Alaskan red fox (*Vulpes vulpes*). Plos one, 16(4), e0249176.
7. Belinky, F., Ganguly, I., Poliakov, E., Yurchenko, V., & Rogozin, I. B. (2021). Analysis of stop codons within prokaryotic protein-coding genes suggests frequent readthrough events. International journal of molecular sciences, 22(4), 1876.
8. Bernard RF, Willcox EV, Jackson RT, Brown VA, McCracken GF. Feasting, not fasting: winter diets of cave hibernating bats in the United States. Frontiers in Zoology. 2021 Dec;18:1-3.
9. Bernard RF, Willcox EV, Parise KL, Foster JT, McCracken GF. White-nose syndrome fungus, Pseudogymnoascus destructans, on bats captured emerging from caves during winter in the southeastern United States. BMC Zoology. 2017 Dec;2(1):1-1.
10. Blehert, D. S., Hicks, A. C., Behr, M., Meteyer, C. U., Berlowski-Zier, B. M., Buckles, E. L., ... & Stone, W. B. (2009). Bat white-nose syndrome: an emerging fungal pathogen?. Science, 323(5911), 227-227.
11. Blejwas K, Beard L, Buchanan J, Lausen CL, Neubaum D, Tobin A, Weller TJ. COULD WHITE-NOSE SYNDROME MANIFEST DIFFERENTLY IN *MYOTIS LUCIFUGUS* IN WESTERN VERSUS EASTERN REGIONS OF NORTH AMERICA? A REVIEW OF FACTORS. Journal of Wildlife Diseases. 2023 Jun 1.
12. Bogdan, C., Röllinghoff, M., & Diefenbach, A. (2000). The role of nitric oxide in innate immunity. Immunological reviews, 173, 17-26.
13. Bohn SJ, Turner JM, Warnecke L, Mayo C, McGuire LP, Misra V, Bollinger TK, Willis CK. Evidence of 'sickness behaviour'in bats with white-nose syndrome. Behaviour. 2016 Jan 1;153(8):981-1003.

14. Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 30(15), 2114-2120.
15. Carlson, C. J., Albery, G. F., Merow, C., Trisos, C. H., Zipfel, C. M., Eskew, E. A., ... & Bansal, S. (2021). Climate change will drive novel cross-species viral transmission. BioRxiv, 2020-01.
16. Cheng TL, Reichard JD, Coleman JT, Weller TJ, Thogmartin WE, Reichert BE, Bennett AB, Broders HG, Campbell J, Etchison K, Feller DJ. The scope and severity of white-nose syndrome on hibernating bats in North America. Conservation Biology. 2021 Oct;35(5):1586-97.
17. Chhatre, V. E., & Emerson, K. J. (2017). StrAuto: automation and parallelization of STRUCTURE analysis. BMC bioinformatics, 18(1), 1-5.
18. Cryan, P. M., Meteyer, C. U., Boyles, J. G., & Blehert, D. S. (2010). Wing pathology of white-nose syndrome in bats suggests life-threatening disruption of physiology. BMC biology, 8(1), 1-8.
19. Cryan, P. M., Meteyer, C. U., Boyles, J. G., & Blehert, D. S. (2013). White-nose syndrome in bats: illuminating the darkness. BMC biology, 11(1), 1-4.
20. Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. Bioinformatics, 27(15), 2156-2158.
21. Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., ... & Li, H. (2021). Twelve years of SAMtools and BCFtools. Gigascience, 10(2), giab008.
22. Davy, C. M., Donaldson, M. E., Bandouchova, H., Breit, A. M., Dorville, N. A., Dzal, Y. A., ... & Kyle, C. J. (2020). Transcriptional host–pathogen responses of Pseudogymnoascus destructans and three species of bats with white-nose syndrome. Virulence, 11(1), 781-794.
23. Davy, C. M., Donaldson, M. E., Rico, Y., Lausen, C. L., Dogantzis, K., Ritchie, K., ... & Kyle, C. J. (2017). Prelude to a panzootic: Gene flow and immunogenetic variation in northern little brown myotis vulnerable to bat white-nose syndrome. Facets, 2(2), 690-714.
24. Davy, C. M., Martinez-Nunez, F., Willis, C. K., & Good, S. V. (2015). Spatial genetic structure among bat hibernacula along the leading edge of a rapidly spreading pathogen. Conservation Genetics, 16(5), 1013-1024.
25. Dzal Y, McGuire LP, Veselka N, Fenton MB. Going, going, gone: the impact of white-nose syndrome on the summer activity of the little brown bat (*Myotis lucifugus*). Biology letters. 2011 Jun 23;7(3):392-4.
26. DeCandia, A. L., Dobson, A. P., & vonHoldt, B. M. (2018). Toward an integrative molecular approach to wildlife disease. Conservation Biology, 32(4), 798-807.
27. DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., ... & Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature genetics, 43(5), 491-498.

28. Dobony CA, Johnson JB. Observed resiliency of little brown myotis to long-term white-nose syndrome exposure. Journal of Fish and Wildlife Management. 2018 Jun 1;9(1):168-79.
29. Donaldson ME, Davy CM, Willis CK, McBurney S, Park A, Kyle CJ. Profiling the immunome of little brown myotis provides a yardstick for measuring the genetic response to white-nose syndrome. Evolutionary applications. 2017 Dec;10(10):1076-90.
30. Ekblom, R., Saether, S. A., Jacobsson, P., Fiske, P., Sahlman, T., Grahn, M., ... & Höglund, J. (2007). Spatial pattern of MHC class II variation in the great snipe (Gallinago media). Molecular ecology, 16(7), 1439-1451.
31. Epstein, B., Jones, M., Hamede, R., Hendricks, S., McCallum, H., Murchison, E. P., ... & Storfer, A. (2016). Rapid evolutionary response to a transmissible cancer in Tasmanian devils. Nature communications, 7(1), 1-7.
32. Erdle SY, Hobson CS. Current status and conservation strategy for the Eastern Smallfooted Myotis (*Myotis leibii*). Heritage Technical Report No. 00-19. Virginia Department of Conservation and Recreation, Division of Natural Heritage, Richmond, VA; 2001.
33. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular ecology. 2005 Jul;14(8):2611-20.
34. Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular ecology resources, 10(3), 564-567.
35. Fenton, M. B. (1972). Distribution and overwintering of *Myotis leibii* and Eptesicus fuscus (Chiroptera: Vespertilionidae) in Ontario.
36. Fischer, N. M., Altewischer, A., Ranpal, S., Dool, S., Kerth, G., & Puechmaille, S. J. (2022). Population genetics as a tool to elucidate pathogen reservoirs: Lessons from Pseudogymnoascus destructans, the causative agent of White-Nose disease in bats. Molecular Ecology, 31(2), 675-690.
37. Fisher MC, Garner TW. Chytrid fungi and global amphibian declines. Nature Reviews Microbiology. 2020 Jun;18(6):332-43.
38. Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics, 180(2), 977-993.
39. Francis, R. M. (2017). pophelper: an R package and web app to analyse and visualize population structure. Molecular ecology resources, 17(1), 27-32.
40. Francl KE, Ford WM, Sparks DW, Brack Jr V. Capture and reproductive trends in summer bat communities in West Virginia: assessing the impact of white-nose syndrome. Journal of Fish and Wildlife Management. 2012 Jun;3(1):33-42.

41. Frick, W. F., Cheng, T. L., Langwig, K. E., Hoyt, J. R., Janicki, A. F., Parise, K. L., ... & Kilpatrick, A. M. (2017). Pathogen dynamics during invasion and establishment of white-nose syndrome explain mechanisms of host persistence.
42. Frick WF, Johnson E, Cheng TL, Lankton JS, Warne R, Dallas J, Parise KL, Foster JT, Boyles JG, McGuire LP. Experimental inoculation trial to determine the effects of temperature and humidity on White-nose Syndrome in hibernating bats. Scientific reports. 2022 Jan 19;12(1):971.
43. Frick, W. F., Puechmaille, S. J., & Willis, C. K. R. (2016). White-nose syndrome in bats. In 'Bats in the Anthropocene: Conservation of Bats in a Changing World'.(Eds CC Voigt and T. Kingston.) pp. 245–262.
44. Gignoux-Wolfsohn, S. A., Pinsky, M. L., Kerwin, K., Herzog, C., Hall, M., Bennett, A. B., ... & Maslo, B. (2021). Genomic signatures of selection in bats surviving white-nose syndrome. Molecular Ecology, 30(22), 5643-5657.
45. Goudet, J. (2005). Hierfstat, a package for R to compute and test hierarchical F-statistics. Molecular ecology notes, 5(1), 184-186.
46. Hayman DT, Pulliam JR, Marshall JC, Cryan PM, Webb CT. Environment, host, and fungal traits predict continental-scale white-nose syndrome in bats. Science advances. 2016 Jan 29;2(1):e1500831.
47. Haworth, S. E., Nituch, L., Northrup, J. M., & Shafer, A. B. (2021). Characterizing the demographic history and prion protein variation to infer susceptibility to chronic wasting disease in a naïve population of white-tailed deer (Odocoileus virginianus). Evolutionary applications, 14(6), 1528-1539.
48. Hohenlohe, P. A., Funk, W. C., & Rajora, O. P. (2021). Population genomics for wildlife conservation and management. Molecular Ecology, 30(1), 62-82.
49. Hooton LA, Adams AA, Cameron A, Fraser EE, Hale L, Kingston S, Fenton MB, McGuire LP, Stukenholtz EE, Davy CM. Effects of bat white-nose syndrome on hibernation and swarming aggregations of bats in Ontario. Canadian Journal of Zoology. 2023 May 29
50. Hoyt, J. R., Sun, K., Parise, K. L., Lu, G., Langwig, K. E., Jiang, T., ... & Feng, J. (2016). Widespread bat white-nose syndrome fungus, northeastern China. Emerging Infectious Diseases, 22(1), 140.
51. Humphrey, C. (2017). Recovery Strategy for the Eastern Small-footed Myotis (*Myotis leibii*) in Ontario. Ontario Recovery Strategy Series. Prepared for the Ontario Ministry of Natural Resources and Forestry, Peterborough, Ontario. vii.
52. Ineson KM. Demography of a recovery: tracking the rebound of little brown bat (*Myotis lucifugus*) populations (Doctoral dissertation, University of New Hampshire).

53. Ivanova, N. N., Schwientek, P., Tripp, H. J., Rinke, C., Pati, A., Huntemann, M., ... & Rubin, E. M. (2014). Stop codon reassignments in the wild. Science, 344(6186), 909-913.
54. Jackson RT, Willcox EV, Bernard RF. Winter torpor expression varies in four bat species with differential susceptibility to white-nose syndrome. Scientific Reports. 2022 Apr 5;12(1):5688.
55. Johnson, J. B., & Gates, J. E. (2008). Spring migration and roost selection of female *Myotis leibii* in Maryland. Northeastern Naturalist, 15(3), 453-460.
56. Johnson, J. S., Kiser, J. D., Watrous, K. S., & Peterson, T. S. (2011). Day-roosts of *Myotis leibii* in the Appalachian Ridge and Valley of West Virginia. Northeastern Naturalist, 18(1), 95-106.
57. Jombart T. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics. 2008 Apr 8;24(11):1403-5.
58. Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. Ecology letters, 7(12), 1225-1241.
59. Kilpatrick, A. M., Briggs, C. J., & Daszak, P. (2010). The ecology and impact of chytridiomycosis: an emerging disease of amphibians. Trends in ecology & evolution, 25(2), 109-118.
60. Kovacova, V., Zukal, J., Bandouchova, H., Botvinkin, A. D., Harazim, M., Martínková, N., ... & Pikula, J. (2018). White-nose syndrome detected in bats over an extensive area of Russia. BMC veterinary research, 14(1), 1-9.
61. Langwig, K. E., Frick, W. F., Bried, J. T., Hicks, A. C., Kunz, T. H., & Marm Kilpatrick, A. (2012). Sociality, density-dependence and microclimates determine the persistence of populations suffering from a novel fungal disease, white-nose syndrome. Ecology letters, 15(9), 1050-1057.
62. Leopardi S, Blake D, Puechmaille SJ. White-nose syndrome fungus introduced from Europe to North America. Current Biology. 2015 Mar 16;25(6):R217-9.
63. Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.3997.
64. Lilley TM, Prokkola JM, Blomberg AS, Paterson S, Johnson JS, Turner GG, Bartonička T, Bachorec E, Reeder DM, Field KA. Resistance is futile: RNA-sequencing reveals differing responses to bat fungal pathogen in Nearctic *Myotis lucifugus* and Palearctic *Myotis myotis*. Oecologia. 2019 Oct;191:295-309.
65. Lilley, T. M., Wilson, I. W., Field, K. A., Reeder, D. M., Vodzak, M. E., Turner, G. G., ... & Paterson, S. (2020). Genome-wide changes in genetic diversity in a population of *Myotis lucifugus* affected by white-nose syndrome. G3: Genes, Genomes, Genetics, 10(6), 2007-2020.
66. Lindborg CM, Propert KJ, Pignolo RJ. Conservation of pro-longevity genes among mammals. Mechanisms of ageing and development. 2015 Mar 1;146:23-7.

67. Longo, A. V., Burrowes, P. A., & Zamudio, K. R. (2014). Genomic studies of disease-outcome in host–pathogen dynamics. American Zoologist, 54(3), 427-438.
68. Lowenstein, C. J., & Padalko, E. (2004). iNOS (NOS2) at a glance. Journal of cell science, 117(14), 2865-2867.
69. Lozier, J. D., & Zayed, A. (2017). Bee conservation in the age of genomics. Conservation Genetics, 18(3), 713-729.
70. Luu, K., Bazin, E., & Blum, M. G. (2017). pcadapt: an R package to perform genome scans for selection based on principal component analysis. Molecular ecology resources, 17(1), 67-77.
71. Maslo B, Valent M, Gumbs JF, Frick WF. Conservation implications of ameliorating survival of little brown bats with white-nose syndrome. Ecological Applications. 2015 Oct;25(7):1832-40.
72. Massey, S. E. (2017). The identities of stop codon reassignments support ancestral tRNA stop codon decoding activity as a facilitator of gene duplication and evolution of novel function. Gene, 619, 37-43.
73. McCallum, H. (2008). Tasmanian devil facial tumour disease: lessons for conservation biology. Trends in ecology & evolution, 23(11), 631-637.
74. McGuire LP, Mayberry HW, Willis CK. White-nose syndrome increases torpid metabolic rate and evaporative water loss in hibernating bats. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology. 2017 Dec 1;313(6):R680-6.
75. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., ... & DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research, 20(9), 1297-1303.
76. Medzhitov R, Schneider DS, Soares MP. Disease tolerance as a defense strategy. Science. 2012 Feb 24;335(6071):936-41.
77. Metheny, J. D., Kalcounis-Rueppell, M. C., Willis, C. K., Kolar, K. A., & Brigham, R. M. (2008). Genetic relationships between roost-mates in a fission–fusion society of tree-roosting big brown bats (Eptesicus fuscus). Behavioral Ecology and Sociobiology, 62(7), 1043-1051.
78. Miller-Butterworth CM, Vonhof MJ, Rosenstern J, Turner GG, Russell AL. Genetic structure of little brown bats (*Myotis lucifugus*) corresponds with spread of white-nose syndrome among hibernacula. Journal of Heredity. 2014 May;105(3):354-64.
79. Moore, M. S., Reichard, J. D., Murtha, T. D., Nabhan, M. L., Pian, R. E., Ferreira, J. S., & Kunz, T. H. (2013). Hibernating little brown myotis (*Myotis lucifugus*) show variable immunological responses to white-nose syndrome. PLoS One, 8(3), e58976.
80. Moore MS, Field KA, Behr MJ, Turner GG, Furze ME, Stern DW, Allegra PR, Bouboulis SA, Musante CD, Vodzak ME, Biron ME. Energy conserving thermoregulatory patterns and lower disease severity in a bat resistant to

the impacts of white-nose syndrome. Journal of Comparative Physiology B. 2018 Jan;188:163-76.
81. Moosman PR, Andersen PR, Frasier MG. Use of rock-crevices in winter by big brown bats and eastern small-footed bats in the Appalachian Ridge and Valley of Virginia. Banisteria. 2017;48:9-13.
82. Moosman PR, Warner DP, Hendren RH, Hosler MJ. Potential for monitoring eastern small-footed bats on talus slopes. Northeastern Naturalist. 2015 Mar;22(1).
83. Neaves, L. E., Zenger, K. R., Prince, R. I., Eldridge, M. D., & Cooper, D. W. (2009). Landscape discontinuities influence gene flow and genetic structure in a large, vagile Australian mammal, Macropus fuliginosus. Molecular Ecology, 18(16), 3363-3378.
84. Noreen, M., & Arshad, M. (2015). Association of TLR1, TLR2, TLR4, TLR6, and TIRAP polymorphisms with disease susceptibility. Immunologic Research, 62(2), 234-252.
85. O'Keefe JM, Pettit JL, Loeb SC, Stiver WH. White-nose syndrome dramatically altered the summer bat assemblage in a temperate Southern Appalachian forest. Mammalian Biology. 2019 Sep;98:146-53.
86. Oswald, P., Rodríguez, A., Bourke, J., Wagner, N., de Buhr, N., Buschmann, H., ... & Pröhl, H. (2020). Locality, time and heterozygosity affect chytrid infection in yellow-bellied toads. Diseases of aquatic organisms, 142, 225-237.
87. Paradis, E., Claude, J., & Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. Bioinformatics, 20(2), 289-290.
88. Peakall, R., Ruibal, M., & Lindenmayer, D. B. (2003). Spatial autocorrelation analysis offers new insights into gene flow in the Australian bush rat, Rattus fuscipes. Evolution, 57(5), 1182-1195.
89. Pina-Martins F, Baptista J, Pappas Jr G, Paulo OS. New insights into adaptation and population structure of cork oak using genotyping by sequencing. Global Change Biology. 2019 Jan;25(1):337-50.
90. Poplin, R., Ruano-Rubio, V., DePristo, M. A., Fennell, T. J., Carneiro, M. O., Van der Auwera, G. A., ... & Banks, E. (2018). Scaling accurate genetic variant discovery to tens of thousands of samples. BioRxiv, 201178.
91. Praest, P., Luteijn, R. D., Brak-Boer, I. G. J., Lanfermeijer, J., Hoelen, H., Ijgosse, L., ... & Wiertz, E. J. H. J. (2018). The influence of TAP1 and TAP2 gene polymorphisms on TAP function and its inhibition by viral immune evasion proteins. Molecular immunology, 101, 55-64.
92. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000 Jun 1;155(2):945-59.
93. Puechmaille, S. J., Wibbelt, G., Korn, V., Fuller, H., Forget, F., Mühldorfer, K., ... & Teeling, E. C. (2011). Pan-European distribution of white-nose syndrome fungus (Geomyces destructans) not associated with mass mortality. PloS one, 6(4), e19167.

94. Team, R. C. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www. R-project. org/.
95. Råberg L, Graham AL, Read AF. Decomposing health: tolerance and resistance to parasites in animals. Philosophical Transactions of the Royal Society B: Biological Sciences. 2009 Jan 12;364(1513):37-49.
96. Reeder, D. M., Frank, C. L., Turner, G. G., Meteyer, C. U., Kurta, A., Britzke, E. R., ... & Blehert, D. S. (2012). Frequent arousal from hibernation linked to severity of infection and mortality in bats with white-nose syndrome. PloS one, 7(6), e38920.
97. Reynolds DS, Shoemaker K, von Oettingen S, Najjar S, Veilleux JP, Moosman PR. Integrating Multiple Survey Techniques to Document a Shifting Bat Community in the Wake of White-Nose Syndrome. Journal of Fish and Wildlife Management. 2021 Dec 1;12(2):395-411.
98. Reynolds RJ, Powers KE, Orndorff W, Ford WM, Hobson CS. Changes in rates of capture and demographics of *Myotis septentrionalis* (northern long-eared bat) in western Virginia before and after onset of white-nose syndrome. Northeastern Naturalist. 2016 Jun;23(2):195-204.
99. Ricciardi, A., Hoopes, M. F., Marchetti, M. P., & Lockwood, J. L. (2013). Progress toward understanding the ecological impacts of nonnative species. Ecological monographs, 83(3), 263-282.
100. Ritz, U., Drexler, I., Sutter, D., Abele, R., Huber, C., & Seliger, B. (2003). Impaired transporter associated with antigen processing (TAP) function attributable to a single amino acid alteration in the peptide TAP subunit TAP1. The Journal of Immunology, 170(2), 941-946.
101. Rydengård, V., Shannon, O., Lundqvist, K., Kacprzyk, L., Chalupka, A., Olsson, A. K., ... & Schmidtchen, A. (2008). Histidine-rich glycoprotein protects from systemic Candida infection. PLoS pathogens, 4(8), e1000116.
102. Santillán DD, Lama TM, Guerrero YT, Brown AM, Donat P, Zhao H, Rossiter SJ, Yohe LR, Potter JH, Teeling EC, Vernes SC. Large-scale genome sampling reveals unique immunity and metabolic adaptations in bats. Molecular ecology. 2021 Jan 4.
103. Sasse DB, Perry RW. Apparent annual survival of female eastern small-footed bats (*Myotis leibii*) roosting in Arkansas bridges. Journal of Mammalogy. Journal of Mammalogy, 104(6):1257–1263, 2023.
104. Savage, A. E., & Zamudio, K. R. (2011). MHC genotypes associate with resistance to a frog-killing fungus. Proceedings of the National Academy of Sciences, 108(40), 16705-16710.
105. Schaefer, C., & Rost, B. (2012, June). Predict impact of single amino acid change upon protein structure. In BMC genomics (Vol. 13, No. 4, pp. 1-10). BioMed Central.

106. Schumann, R. R. (2011). Old and new findings on lipopolysaccharide-binding protein: a soluble pattern-recognition molecule. Biochemical Society Transactions, 39(4), 989-993.
107. Schumann, R. R., Leong, S. R., Flaggs, G. W., Gray, P. W., Wright, S. D., Mathison, J. C., ... & Ulevitch, R. J. (1990). Structure and function of lipopolysaccharide binding protein. Science, 249(4975), 1429-1431.
108. Schweizer, R. M., Robinson, J., Harrigan, R., Silva, P., Galverni, M., Musiani, M., ... & Wayne, R. K. (2016). Targeted capture and resequencing of 1040 genes reveal environmentally driven functional variation in grey wolves. Molecular ecology, 25(1), 357-379.
109. Seya, T., Oshiumi, H., Sasai, M., Akazawa, T., & Matsumoto, M. (2005). TICAM-1 and TICAM-2: toll-like receptor adapters that participate in induction of type 1 interferons. The international journal of biochemistry & cell biology, 37(3), 524-529.
110. Solari, S. 2018. *Myotis leibii*. The IUCN Red List of Threatened Species 2018: e.T14172A22055716. https://dx.doi.org/10.2305/IUCN.UK.2018-2.RLTS.T14172A22055716.en. Accessed on 07 March 2023.
111. Supple, M. A., & Shapiro, B. (2018). Conservation of biodiversity in the genomics era. Genome biology, 19(1), 1-12.
112. Thompson TQ, Bellinger MR, O'Rourke SM, Prince DJ, Stevenson AE, Rodrigues AT, Sloat MR, Speller CF, Yang DY, Butler VL, Banks MA. Anthropogenic habitat alteration leads to rapid loss of adaptive variation and restoration potential in wild salmon populations. Proceedings of the National Academy of Sciences. 2019 Jan 2;116(1):177-86.
113. Turner GG, Reeder D, Coleman JT. A five-year assessment of mortality and geographic spread of white-nose syndrome in North American bats, with a look at the future. Update of white-nose syndrome in bats. Bat research news. 2011;52:13.
114. Van der Auwera, G. A., & O'Connor, B. D. (2020). Genomics in the cloud: using Docker, GATK, and WDL in Terra. O'Reilly Media.
115. Vander Wal, E., Garant, D., Calmé, S., Chapman, C. A., Festa-Bianchet, M., Millien, V., ... & Pelletier, F. (2014). Applying evolutionary concepts to wildlife disease ecology and management. Evolutionary applications, 7(7), 856-868.
116. Verant, M. L., Boyles, J. G., Waldrep Jr, W., Wibbelt, G., & Blehert, D. S. (2012). Temperature-dependent growth of Geomyces destructans, the fungus that causes bat white-nose syndrome.
117. Verant, M. L., Meteyer, C. U., Speakman, J. R., Cryan, P. M., Lorch, J. M., & Blehert, D. S. (2014). White-nose syndrome initiates a cascade of physiologic disturbances in the hibernating bat host. BMC physiology, 14(1), 1-11.
118. Vonhof, M. J., Russell, A. L., & Miller-Butterworth, C. M. (2015). Range-wide genetic analysis of little brown bat (*Myotis lucifugus*) populations:

estimating the risk of spread of white-nose syndrome. PloS one, 10(7), e0128713.
119. Wang, Z. K., Yang, Y. S., Stefka, A. T., Sun, G., & Peng, L. H. (2014). Fungal microbiota and digestive diseases. Alimentary pharmacology & therapeutics, 39(8), 751-766.
120. Ward C, Beharry A, Tennakoon R, Rozik P, Wilhelm SD, Heinemann IU, O'Donoghue P. Mechanisms and Delivery of tRNA Therapeutics. Chemical Reviews. 2024 May 27.
121. Warnecke L, Turner JM, Bollinger TK, Misra V, Cryan PM, Blehert DS, Wibbelt G, Willis CK. Pathophysiology of white-nose syndrome in bats: a mechanistic model linking wing damage to mortality. Biology Letters. 2013 Aug 23;9(4):20130177.
122. Whiting-Fawcett, F., Field, K. A., Puechmaille, S. J., Blomberg, A. S., & Lilley, T. M. (2021). Heterothermy and antifungal responses in bats. Current Opinion in Microbiology, 62, 61-67.
123. Whitlock, M. C., & Lotterhos, K. E. (2015). Reliable detection of loci responsible for local adaptation: Inference of a null model through trimming the distribution of F ST. The American Naturalist, 186(S1), S24-S36.
124. Wilder, A. P., Frick, W. F., Langwig, K. E., & Kunz, T. H. (2011). Risk factors associated with mortality from white-nose syndrome among hibernating bat colonies. Biology Letters, 7(6), 950-953.
125. Wilder, A. P., Kunz, T. H., & Sorenson, M. D. (2015). Population genetic structure of a common host predicts the spread of white-nose syndrome, an emerging infectious disease in bats. Molecular ecology, 24(22), 5495-5506.
126. Yi X, Latch EK. Nonrandom missing data can bias principal component analysis inference of population genetic structure. Molecular Ecology Resources. 2022 Feb;22(2):602-11.
127. Yıldırım Y, Tinnert J, Forsman A. Contrasting patterns of neutral and functional genetic diversity in stable and disturbed environments. Ecology and Evolution. 2018 Dec;8(23):12073-89.

Chapter 5

Variable population impacts from an invasive pathogen relate to both local demographic and adaptive processes: variable recovery of little brown bat populations subsequent to *P. destructans* epizootics

Preface

A version of this chapter is being prepared for peer-review submission.

Authors: Tristan M. Baecklund, Michael E. Donaldson, Maarten J. Vonhof, D. Blake Sasse, Christina M. Davy, and Christopher J. Kyle

Contributions: MED, CMD, and CJK conceptualized and designed the study; TMB did all the lab work; MED extracted and processed raw sequence data pertaining to microsatellite genomic intervals with Dante; TMB did data processing, analyses and visualization; TMB wrote the manuscript; TMB and CJK have reviewed the manuscript.

Abstract

Undermined barriers to pathogen dispersal from human-mediated translocations and globalization exacerbate changes to host-pathogen dynamics that subsequently threaten species persistence given uncertain capacities to adapt to rapidly shifting suites of selective pressures. *Pseudogymnoascus destructans* (*Pd*), the causative agent of white-nose syndrome (WNS), was anthropogenically introduced into North America in 2006 and resulted in drastic population declines (>90%) for several hibernating bat species, including *Myotis lucifugus*. Of the *Pd*-affected populations, some now show evidence of recovery, where outcomes appear partially influenced by genetic factors. Understanding how population dynamics and biodiversity were impacted throughout the introduction and spread of *Pd* remains limited but vital to ongoing efforts promoting Nearctic bat conservation.

Herein, we explore variation associated with neutral and functionally relevant regions of the *M. lucifugus* genome to assess temporal impacts of *Pd* exposure and evidence of evolutionary rescue in recovering populations. Neutral genetic structure assessments showed no temporal patterns associated with exposure to *Pd*. In contrast, intra-population diversity shifts within immunogenetic regions demonstrated signatures of selection to *Pd* exposure, but loci and patterns were inconsistent among affected populations. These data suggest that observed shifts in genetic diversity are unexplained by drift alone, where signatures of WNS-imposed selection were stronger in Atlantic populations relative to eastern Ontario now showing signs of recovery. We speculate that these contrasting patterns are

associated with varied demographic factors, such as relative population sizes and gene flow, but also local adaptations. Here, we provide further insight into the demographic and evolutionary rescue mechanisms in *M. lucifugus* populations following the introduction of *Pd* that exemplifies the importance of connectivity and large population sizes for population persistence in population-specific disease responses as this pathogen continues to spread westward.

5.1 Introduction

Species diversity is maintained, in part, by balances between species dispersal capacities and local adaptation to differing suites of selective pressures encountered across heterogeneous landscapes (Hohenlohe et al. 2021; Lozier & Zayed 2017; Yildirim et al. 2018). Globalization, human-mediated translocations, and climate change alter these balances; undermining natural barriers to species dispersal and facilitating range expansions into regions previously unsuitable or inaccessible (Ancillotto et al. 2016; Baker et al., 2022; Davy et al. 2017; Fischer et al. 2022; Gallana et al. 2013; Longo et al. 2014; Ricciardi et al. 2013; Rohlin et al. 2013; Vander Wal et al. 2014). Such range shifts can impose novel, selective pressures on endemic populations by introducing invasive species which compete for resources (Langmaier et al. 2020), altering trophic interactions (i.e., new predators; Baker et al. 2020; Jolly et al. 2021), and by changing distributions of infectious diseases (Baker et al. 2022), thereby threatening native biodiversity.

Host-pathogen dynamics are changing at unprecedented rates as human-mediated translocations and climatic changes enable pathogens to expand their ranges into previously inhospitable or inaccessible environments (Baker et al. 2022). Resulting disease often threatens the persistence of naïve populations and species as demonstrated by the global spread of the fungal pathogen, *Batrachochytrium dendrobatidis* (*Bd*) on many amphibian species (Fisher & Garner, 2020; Gallana et al. 2013; Oswald et al. 2020; Walker et al. 2008). Infectious diseases, notably in their epizootic stages, elicit rapid changes in the distribution and diversity of genetic variants associated with susceptibility over short time scales through selective sweeps (Berry et al. 1991; Gallana et al. 2013). Understanding how populations respond to change elicited by disease becomes increasingly important to conservation efforts supporting susceptible species or populations.

In some cases, populations can respond to a selective sweep through a genetic rescue, where population fitness increases through reduced inbreeding resulting from migration and increases the frequency of heterozygous genotypes (Hufbauer et al. 2015; Kanarek et al. 2015). Genetic rescues can also occur through bolstered adaptive potential among populations where the frequency of genetic variants associated with beneficial traits increases (Hufbauer et al. 2015; Kanarek et al. 2015). Herein, we adopt the perspective of Hufbauer et al. (2015), distinguishing genetic from evolutionary rescues, where the latter occurs in the absence of migration and specifically denotes an increase in frequency of advantageous genotypes from the standing variation in the population after initial population

declines. In either scenario of genetic rescue, if the selective pressure is too strong, especially where host populations are small or have low fecundity, extirpation or extinction can occur when populations cannot survive/recover from first contact with the disease. Such a pattern is demonstrated with the extinction of almost 100 amphibian species attributed to *Bd* exposure (Fisher & Garner 2020; Oswald et al. 2020; Smith et al. 2022).

Natural or artificial translocations, however, are not always feasible or advisable given that out-breeding depression can result from adaptations to local suites of selective pressures other than disease (Templeton et al. 1986; Frankham et al. 2011; Kanarek et al. 2015). Alternatively, metapopulations with large effective population sizes and high gene flow may persist after the introduction of infectious disease through an influx of individuals from connected populations or when the number of offspring continues to exceed disease-associated mortality leading to stable or increasing population trends (i.e., demographic rescue; Kanarek et al. 2015). Demographic rescues within *Bd systems* can be illustrated by the persistence of growling grass frogs (*Litoria raniformis*) predicated upon refugia and connectivity amongst metapopulations (Heard et al. 2015). Further evidence of the potential for both demographic and genetic rescues exist among several other species (e.g., Knapp et al. 2023; Byrne et al. 2021; Trumbo et al. 2023). Given the variability in observed (or the potential for) demographic and genetic recuses across amphibian species in the context of *Bd* exposure, it is apparent that broad assessments and conclusions are not always transferable among populations or

species. By understanding mechanisms of recovery, or lack thereof, and factors that foster both demographic and genetic recoveries, these data can enhance vulnerable and/or endangered species' resiliency to future selective pressures at the population scale through informed conservation and management efforts.

Incorporating genetic assays that investigate both neutral and functionally relevant genomic regions offers the potential to quantify population connectivity and genetic diversity associated with heterogeneous suites of selective pressures (Kawecki & Ebert 2004; Neaves et al., 2009; Peakall et al., 2003; Willi et al. 2022). These assays not only elucidate genomic signatures of selection, but contextualize population connectivity that influences selective patterns by diluting the frequency of adaptive genes or enhancing the potential for local adaptations (Davy et al. 2017; Kanarek et al. 2015; Supple & Shapiro 2018; Xuereb et al. 2021; Yi & Latch 2022). Genotyping-by-sequencing (GBS) approaches offer the flexibility and simplified means to selectively and simultaneously sequence genomic regions of interest that encompass both neutral and functionally relevant loci (Pina-Martins et al. 2019; Schweizer et al. 2016). For example, Schweizer et al. (2016), used GBS assays to identify shifts in genetic diversity associated with immune responses in gray wolves despite extensive population connectivity; a finding suggestive of strong local selective pressures. Similar approaches have been used to develop assays investigating genetic diversity among Tasmanian devil (*Sarcophilus harrisii*) populations following the emergence of devil facial tumour disease (Wright et al. 2015). These studies demonstrate the potential of GBS tools to directly assess

signatures of selection associated with disease in the context of population connectivity.

In 2006, the fungal pathogen *Pseudogymnoascus destructans* (*Pd*), the causative agent of white-nose syndrome (WNS), was introduced to North America from Eurasia (Leopardi, Blake, & Puechmaille, 2015). This pathogen rapidly spread across eastern and central North America, resulting in significant population declines in several hibernating bat species (Blehert et al. 2009; Frick et al. 2016; Hoyt et al. 2016; Kovacova et al. 2018; Puechmaille et al. 2011; Wilder et al. 2015; Cheng et al. 2021; Dzal et al. 2011; Francl et al. 2012). White-nose syndrome transmission and infection peaks during hibernation when bats congregate at hibernacula, their immune functions are downregulated, and body temperatures decrease during torpor, thus providing an opportunistic environment for this psychrophilic pathogen (Boddy 2016; Fischer et al. 2022; Langwig et al. 2012; Verant et al. 2012; Whiting-Fawcett et al. 2021). The cascade of physiological changes that bats undergo upon infection increases morbidity by affecting the rate of arousal during torpor and depleting energy stores throughout winter leading to mortality (Anderson 2018; Auteri & Knowles 2020; Blehert et al. 2009; Cryan et al. 2010; Moore et al. 2013; Reeder et al. 2012; Verant et al. 2014). Susceptibility to WNS varies among Nearctic bats due to varying environmental, behavioural, ecological, genetic, and individual factors between habitats and species that influence pathogen exposure outcomes (Cryan et al. 2013; Davy et al. 2020; Hayman et al. 2016). Environmental barriers have not been shown to impede the host-mediated

dispersal of *Pd* explaining its rapid continued spread across the majority of North America (Fig 5.1; Davy et al. 2015; Davy et al. 2017; Miller-Butterworth et al. 2014; Vonhof et al. 2015; Wilder et al. 2015). Specific genetic adaptations are known to influence disease outcomes, where large shifts in genetic diversity associated with an immune response or differences in the regulation of relevant genes are observed (Auteri & Knowles 2020; Davy et al. 2017; Donaldson et al. 2017; Lilley et al. 2020). For example, the big brown bat (*Eptesicus fuscus*), exhibits an upregulated local immune response increasing its resistance or tolerance to WNS, whereas more susceptible species (e.g., *Myotis lucifugus*) display systemic responses (Davy et al. 2020).

Of WNS-susceptible bat species, *Myotis lucifugus* (the little brown bat) experienced > 90% population declines during the initial epizootic phases (Cheng et al. 2021; Dzal et al. 2011; Francl et al. 2012). These declines, combined with the relatively long-lived and ubiquitous nature of the species across North America, have made the *M. lucifugus* system a focal point for understanding the implications this disease has on Nearctic bats (Wilder et al. 2011; Wilder et al. 2015). Most knowledge of the impacts of WNS on *M. lucifugus* focused on eastern distributions of the species, where strong selective sweeps first occurred and placed the species' persistence into question (Cheng et al. 2021). Despite initial declines, some populations, over 4 to 13 years after first documented exposure, now appear to be recovering (Maslo et al. 2015; Frick et al. 2017; Dobony and Johnson 2018; Ineson 2020; Hooton et al. 2023). Given the observed panmictic population at the eastern

extent of *M. lucifugus'* range (Davy et al. 2017), it remains unclear why gene flow and associated influxes of individuals have not led to demographic recoveries for all impacted populations. Some researchers have discussed the potential for a differential response of western *M. lucifugus* (west of the Rocky Mountains) relative to their eastern conspecifics based on several physiological and behavioural factors, such as fewer large congregations of bats in hibernacula (Blejwas et al. 2023). Recent findings from the eastern extent of *M. lucifugus'* distribution suggest different local patterns of WNS-associated impacts exist among subpopulations, despite being masked or unapparent at the scale of the larger metapopulation of the species (Gignoux-Wolfsohn et al. 2021). These findings are indicative of the combined influences of several factors (e.g., selection, gene flow, existing local adaptations) that explain differential genetic impacts or consequences between WNS-affected populations.

Previous genetic assessments of eastern Canadian *M. lucifugus* populations from Manitoba to Atlantic Canada, identified patterns of WNS imposed selection (Donaldson et al. 2017). That study, however, was unable to demonstrate differential WNS-associated patterns of selection among subpopulations as a byproduct of variable demographic factors or pre-existing local pressures, nor temporal patterns of shifting genetic diversity from epizootic to enzootic stages of disease. Incongruent impacts on local populations in the eastern United States have been observed and are hypothesized for western conspecifics (Gignoux-Wolfsohn et al. 2021; Blejwas et al. 2023). As such, for eastern Canadian *M. lucifugus*

subpopulations, there remain knowledge gaps in our understanding of the existence or extent of differential WNS-associated impacts and whether demographic, genetic, evolutionary, or a combination of rescue mechanisms explain the disparate recovery of subpopulations.

Herein, we explore genetic variation associated with neutral and functionally relevant regions of the *M. lucifugus* genome to understand the temporal impacts on genetic diversity and potentially discordant patterns of selection and recovery associated with *Pd* exposure between local populations of the species' eastern extent. Specifically, we expand upon the work of Donaldson et al. (2017) by incorporating a panel of eleven legacy microsatellite markers into an amplicon assay (~300 Kbp of the genome), combining these data with the 92 previously sequenced individuals, and analyzing two subpopulations along a temporal gradient of *Pd* exposure. We use these sequence data to investigate the influence of gene flow, genetic drift, and selection imposed by exposure to WNS on differing patterns of population recovery. These data provide further insights into the signatures of selection left on the genomes of affected *M. lucifugus*, the disparity of these signatures between geographically sampled populations, and mechanisms underlying the recovery of these impacted populations. As WNS continues to spread westward to naïve *M. lucifugus* populations, these data have the potential to inform conservation and management interventions to minimize the impact of this disease.

5.2 Methods

5.2.1 Sample collection, DNA extraction, DNA quality and quantity assessments

3mm wing biopsy punches, were collected non-harmfully following established animal care protocols approved by the institutions of respective sample contributors and encompassed three time-periods of WNS exposure (Table S5.1). Bats sampled preceding the introduction of WNS were collected from Yukon, Manitoba, western Ontario, eastern Ontario, and Atlantic Canada, and are denoted as pre-WNS (PRE; Table S5.1). Deceased bats, where mortality was attributed to the initial spread of WNS, were sampled in eastern Ontario and Atlantic Canada and denoted as early mortalities (EMOR; Table S5.1). Finally, bats persisting in the environment 6+ years after presumable Pd exposure were sampled from eastern Ontario and denoted as survivors (SURV; Table S5.1). Wing punches were stored at -80°C until DNA extraction occurred. All tissue samples were dissolved in lysis buffer containing 600 U/mL proteinase K at 56°C for 2 hours. DNA was extracted from lysed samples using the DNeasy Blood and Tissue Kit (Qiagen), or using a magnetic bead extraction (Magnesil, according to the manufacturer's protocol). DNA samples were assessed for quality and quantity using gel electrophoresis (90 V, 45 minutes, 0.8% agarose gel) stained with ethidium bromide, HighRanger 1 kbp DNA ladder as reference (300 – 10,000 bp; Norgen Biotek) and quantified using a Quant-iT PicoGreen dsDNA Assay Kit (ThermoFischer Scientific).

5.2.2 Library preparation, sequence capture, and high-throughput sequencing

DNA libraries (n = 176) were created following the TruSeq Custom Amplicon Low Input Library Prep Kit (Illumina) and barcoded with Illumina TREx Index Adapters (A501-A508 x R701-R712). Our custom amplicon multiplex covered ~328 kbp of the *M. lucifugus* genome (Myoluc2.0), consisting of 1,323 amplicons in total, encompassing 341 neutral SNPs previously identified (Donaldson et al. 2017), 32 microsatellite markers (simple and complex di- and tetra-nucleotides), and 834 immunogenetic regions (Table S5.3). Libraries were amplified with 25 cycles, following the PCR cycle number guideline specified in the TruSeq Custom Amplicon Low Input Library Prep reference guide, in the context of our number of targeted amplicons. Before pooling, individual DNA libraries were assessed for quality and concentration using ethidium bromide-stained gel electrophoresis (2% agarose) and Quant-iT PicoGreen as outlined above. Final assessments for pooled amplicon libraries were conducted with a bioanalyzer before sequencing on an Illumina MiSeq V3 run using 2x300 bp reads (Advanced Analysis Centre Genomics Facility, University of Guelph).

5.2.3 Sequence alignment, variant annotation, and filtering

Sequence data for 176 amplicon libraries were combined with 92 previously sequenced target-capture libraries (Table S5.1; Donaldson et al. 2017). Raw data were first assessed using FastQC (Andrews 2010) to inspect for any initial issues with the data warranting further investigation, such as per base/sequence quality

scores, sequence duplication, and adapter content. Raw data were then processed with Trimmomatic (Bolger, Lohse, & Usadel 2014) to trim reads for quality and length using the following parameters: i) removing bases at the start and end of reads not surpassing a quality of 3, ii) SLIDINGWINDOW of 4 bp and a quality of 15, and iii) a minimum sequence length of 36 bp. Picard (v.2.20.0-3; http://broadinstitute.github.io/picard/) was used to convert from FastQ to unaligned BAM, mark duplicate reads, and mark indices in sequence data. Paired-end reads for each of the 176 samples were then aligned to the *M. lucifugus* reference genome (Myoluc2.0) using the bwa-mem command in Burrows-Wheeler Aligner (v0.7.17; Li 2013). Sequence metrics were assembled using SAMTOOLS v1.9 (Danecek et al. 2021) and GATK v4.1.0.0 was used for duplicate sequence removal, secondary alignment identification and removal (potential paralogs or copy variants), variant discovery, genotyping, and variant recalibration (DePristo et al. 2011; McKenna et al. 2010; Poplin et al. 2017; Van der Auwera & O'Connor 2020). The following hard filters were applied as part of the GATK pipeline: QualByDepth < 2.0, QUAL (Phred-scaled probability) < 30.0, StrandOddsRatio > 3.0, FisherStrand > 60.0, RMSMappingQuality < 40.0, MappingQualityRankSumTest < -12.5, and ReadPosRankSumTest < -8.0. From this combined dataset of 268 samples, variants associated with the presumably neutral markers and immunome were separated into independent files based on the corresponding genomic intervals for each dataset.

Read data corresponding to targeted microsatellite intervals were processed with Dante (Da Amazing NucleoTide Exposer; Budiš et al. 2019). This workflow uses

an alignment-free algorithm to characterize and genotype STR alleles from high-throughput sequencing data containing repetitive regions. Microsatellite data were filtered accordingly: i) obtained >95% confidence identified by Dante, ii) each locus has < 30% missing data, and iii) only retain individuals with <30% missing data.

Sequence variants associated with presumably neutral SNPs and immunogenetically relevant loci were processed through an iterative filtering process to address potential biases resulting from differences in sequencing approach between generated sequence datasets (amplicon vs. target-capture, respectively). Using VCFtools (Danecek et al. 2011), SNPs were first filtered for a quality score < 20, minor allele count < 3, minimum depth per genotype < 35, minimum mean depth < 15, and to retain only biallelic SNPs. We removed loci presumably on the X chromosome based on targeted gene location in other mammalian species, as well as any site that overlapped a primer binding location. Following a similar approach to O'Leary et al. 2018, data were iteratively filtered to minimize missing data per individual and maximize the genotype call rate. An iterative filtering cycle began with minimizing missing individual data, followed by filtering for genotype call rate and repeating this cycle five times. Filtering of missing individual data started at >90%, decreasing to 70%, >50%, >30%, and >25% after each sequential iteration, and genotype call rate filtering started at <60%, increasing to 70%, 80%, 90%, and <95%, after each sequential iteration. To further refine and maintain high confidence and SNP call qualities for analyses of data subsets was achieved by applying a filter for a minor allele count of five. When required during

subsequent analysis of data subsets, linkage disequilibrium pruning was performed using BCFtools v1.9 (Danecek et al. 2021) with max-LD = 0.2 and retaining a single locus within a 500 kbp window.

5.2.4 Analysis of microsatellite data

STRUCTURE analyses were performed with a burn-in length of 50,000 followed by 200,000 iterations for K = 1 through K = 12, with 20 iterations of each K (Pritchard, Stephens, & Donnelly, 2000). Input files for these analyses were generated using STRAUTO v.1.0 (Chhatre & Emerson 2017). Under R v4.1 (R Core Team 2021), pophelper and pophelpershiny (v2.1.1; Francis 2017) were used to estimate the most likely value of K using the Evanno method (Evanno, Regnaut, & Goudet, 2005) and combine and visualize results. Discriminant Analyses of Principal Components (DAPC) were performed as implemented using adegenet v2.1.7 (Jombart 2008) and ape v5.6-2 (Paradis, Claude, & Strimmer 2004). Here, K was estimated using sussesive K-means but also investigated using K equal to the number of representative geographical sampling locations to evaluate which number of identified clusters seemed most plausible in context of the ecology of the species and hypotheses being tested. The function *xvalDapc* was used to identify the number of principle components achieving lowest mean-squared error and retained for discriminant analysis in an attempt to not overfit these data.

5.2.5 Analyses of immunogenetic SNPs

Immunogenetic variants passing hard and iterative filtering parameters were used as input for F_{ST} outlier identification using PCAdapt (Luu, Bazin, and Blum 2017), OutFLANK (Whitlock & Lotterhos 2015), Arlequin (Excoffier & Lischer 2010), and Bayescan (Foll & Gaggiotti 2008), with an adjusted p-value of ≤ 0.05, followed methods described in Baecklund et al. (2021). F_{ST} outlier identification was performed on all retained individuals and subsets of bats from eastern Ontario and Atlantic Canada. Outliers were identified between pre-WNS (PRE) and early mortalities (EMOR) in Atlantic Canada. For eastern Ontario bats, encompassing all sampled time points of WNS exposure, outliers were identified between pre-WNS (PRE) and early mortalities (EMOR) and between pre-WNS (PRE) and survivors (SURV). Data were visualized using STRUCTURE and DAPC analyses in the same manner as microsatellite data. To distinguish between signatures of selection relative to patterns of genetic drift in identified outlier SNPs from within the pre-and post-WNS occurrence datasets, a 95% confidence interval of the Wright-Fisher expectation was generated (https://gitlab.com/WiDGeT_TrentU/). Intervals were based on 1,000 iterations and a randomly sampled generation time of 2-10 years of the model using pre-WNS allele counts. Each SNP was assigned as a potential outlier, where the post-WNS allele frequency fell outside of the generated 95% confidence interval, or as drift where post-WNS allele frequency fell within the interval. The selection coefficient was estimated following Thompson et al. (2019) and Haworth et al., (2021) for any SNP from these datasets with frequency shifts

unassociated with stochastic effects, identified as an F_{ST} outlier, and resulted in a missense mutation.

5.3 Results

5.3.1 Analysis of microsatellite markers

After processing data with Dante and applying hard filters, 11 microsatellite markers and 235 individuals remained for analysis (Table 5.1). STRUCTURE analyses identified K = 2 model as most supported, largely distinguishing Yukon from all other sampled regions, with no further substructure identified within or between regions when Yukon samples were removed (Fig S5.1, S5.10). Discriminant Analysis of Principle Components (DAPC) identified K = 4 clustering as most supported but demonstrated considerable overlap among geographically sampled populations (Fig S5.1, S5.10). As such, at a coarse geographic scale, we inferred K = 2 as the most supported clustering of samples.

5.3.2 Immunogenetic SNP filtering and analyses

After initial hard filters, data for 30,774 SNPs across 268 individuals remained, covering ~200 kbp of the genome. Following subsequent iterative filtering processes, a final filtered working data set of 3,097 SNPs and 231 individuals remained for data visualization, that identified distinct clustering of Yukon samples (K = 2 model most supported). Modelling of K = 2 remained most supported when Yukon samples were removed from consideration, leaving 2,886 SNPs, and did not show further clustering of remaining samples.

5.3.3 Immunogenetic outlier SNP analyses

From sampled bats in Atlantic Canada, a filtered subset of 1,393 SNPs was assessed to identify the coding consequences of the substitutions, influence of genetic drift, and F_{ST} outliers. Across the entire subset, the majority of gene consequences were located in or associated with promoter/regulatory regions, where only six variants were predicted to have high impact coding consequences (Fig S5.6). Three of the high impact variants resulted in a stop codon gain at the HRG, IL12RB1, and HLA-DPB1 genes. The remaining three high impact variants encoded for splice variants at the 3' end of an intron associated with the CD28 gene, and at the 5' end of an intron related to the both the CD4 gene and a novel bat transcript (unassigned gene symbol). Allele frequency shifts for these high impact substitutions were typically ±<10% between sampled time points, except in the case of the HRG stop gain which increased by ~40%. From the initial 1,393 SNPs, 955 demonstrated simulated frequency shifts expected to be caused by genetic drift (Table S5.5). We identified 29 F_{ST} outliers SNPs associated with eight genes between pre-WNS occurrence and early mortality (attributed to the disease) samples from SNPs with frequency changes unlikely to be caused by drift (Table S5.4). Visualization of these data identified the K = 2 model as optimally supported for both DAPC and STRUCTURE analyses (Fig 5.3, S5.8), distinguishing between PRE and EMOR samples within the region; sampled before introduction of WNS and those bats whose deaths were associated with the disease upon introduction, respectively. Eleven missense SNPs were identified within this F_{ST} outlier subset,

where each missense SNPs demonstrated rapid allele frequency shifts and a high estimated coefficient of selection (Fig S5.3). Missense SNPs were associated with JAGN1, HRG, and IL12RB2 genes (Table S5.4).

For bats sampled from eastern Ontario, a subset of 2,016 immunogenetic SNPs remained for analysis. Visualizing these data, encompassing three sampling periods (PRE, EMOR, SURV), did not indicate any further genetic clustering within this region (Fig S5.2, S5.11). Predictions of coding consequences for all 2,016 SNPs revealed the same proportions high impact SNPs including the exact same variants as in the Atlantic Canada subset apart from the stop gain in HRG, an additional stop gain associated with HLA-DMB, and another 3' intron splice variant at the CD4 gene (Fig S5.7). Over the course of the sampled time points, allele frequency fluctuations for these high impact substitutions were around ±10% , with the exception of the stop gain at HLA-DPB1 which gradually increased in frequency by 20%. Analyses of temporal sample comparisons produced subsets of 1,118 and 1,848 SNPs F_{ST} outliers between comparisons of PRE (sampled before introduction of WNS) versus EMOR (bats whose deaths were associated with the disease upon introduction) and PRE versus SURV (bats persisting in the environment several years after introduction of the disease), respectively. For PRE versus EMOR comparisons, 632 SNPs did not display frequency changes attributable to genetic drift based on simulations, where 16 F_{ST} outliers were identified from this subset, associated with four immunogenetically relevant genes (Tables S5.4 & S5.5). In PRE versus SURV comparisons, however, 873 SNPs were not attributable to drift and 37 F_{ST} outliers

were identified across 23 associated genes (Tables S5.4 & S5.5). Visualization of these outlier loci did not support genetic clustering within either of these comparisons (Fig 5.4, S5.9). A total of eight and seven missense SNPs were identified within each of the F_{ST} outlier subsets for PRE versus EMOR and PRE versus SURV, respectively. In both subsets, these missense SNPs demonstrated rapid allele frequency shifts and a high estimated coefficient of selection (Fig S5.4 & S5.5). Associations with the DRB1e2-like, IL1A, and IL1R1 genes were observed for missense SNPs identified in PRE versus EMOR comparisons, where associations with ICAM1, TAP1, IL1A, and TLR3 genes were noted between PRE and SURV comparisons. Outside of these missense SNPs, 136 SNPs demonstrated frequency shifts unattributed to genetic drift in only one of the temporal comparisons in the context of pre-WNS (Table S5.6). While no loci directly overlapped, identified SNPs associated with 18 different genes between the two datasets, and an additional 15 genes containing unique frequency shifts for each comparison were observed (Table S5.6).

5.4 Discussion

We sought to explore the relative impacts of white-nose syndrome (WNS) on *Myotis lucifugus* in Canada using genomic data across the species' range and encompassing three temporal WNS status periods. Data from neutral and immunogenetic loci identified two genetic clusters: 1) samples originating from Yukon, and 2) a genetic cluster spanning sampled regions in Manitoba through Atlantic Canada. Immunogenetic data from sampled bats in Atlantic Canada

demonstrated clear genetic clustering and frequency shifts between bats persisting before the introduction of the disease, and those that succumbed to WNS soon after *Pseudogymnoascus destructans* (*Pd*), occurrence in the region. In contrast, *M. lucifugus* from eastern Ontario did not demonstrate temporal genetic structure in the context of WNS occurrence. Eastern Ontario samples did display several loci where allele frequencies appear to have undergone shifts in concert with the temporal spread of *Pd* in the environment. These data are consistent with literature identifying prominent gene flow across the eastern extent of the species range (Davy et al. 2017) but demonstrate varied *Pd*-associated impacts among populations of *M. lucifugus* in Canada; a pattern also observed in populations of the species in the northeastern United States (Gignoux-Wolfsohn et al. 2021).

5.4.1 Assessment of neutral genetic structure

The amplicon assay implemented herein included presumably neutral SNPs described and microsatellite makers (Donaldson et al. 2017). Barriers prevented targeted microsatellite data from being obtained previously due to issues with INDELs (insertion-deletions) and absent primer motifs (Donaldson et al. 2017). Implementing the Dante program facilitated the characterization and genotyping of 11 microsatellite markers from the combined sequence capture and amplicon dataset analyzed herein. While we successfully harvested data associated with presumably neutral SNPs, only six were retained for analysis after considering linkage disequilibrium (88 SNPs within six regions). As such, we opted to predicate

observed neutral genetic structure from only microsatellite data and remove the neutral SNP data from further analyses.

Neutral microsatellite markers identified genetic clusters that partitioned Yukon samples from a more expansive cluster encompassing the eastern extent of *M. lucifugus'* distribution in Canada (Manitoba eastward); data consistent with previous analyses (Burns et al. 2014; Davy et al. 2017; Johnson et al. 2015). Hierarchical analyses using STRUCTURE, without the presence of Yukon or within sampled locations with temporal sampling, do not provide any further resolution on genetic structure within this large geographic genetic cluster. There was a complete lack of concordance of genetic clustering associated with geographic sampling in the DAPC relative to identified clustering using STRUCTURE. This disparity between STRUCTURE and DAPC is likely attributable to differing underlying algorithms/parameters to distinguish genetic clusters (notably HWE assumed or not). Given the clustering identified using DAPC does not align with the expected patterns of genetic structure previously observed in *M. lucifugus* (Davy et al. 2017), we defaulted to interpreting genetic connectivity from the STRUCTURE results alone. These data continue to support, at least in part, a longitudinal gradient of retarded gene flow across the range of *M. lucifugus* slowed the westward spread of *Pd* relative to the rampant spread observed in eastern North America upon introduction.

5.4.2 Immunogenetic analyses across the M. lucifugus *range in Canada*

Consistent with both microsatellite data presented herein and previous research, immunogenetic data from across the Canadian range of *M. lucifugus* displayed a large panmictic population across the eastern extent of their distribution (Burns et al. 2014; Davy et al. 2017). DAPC analyses excluding samples from Yukon, suggest samples from Atlantic Canada were divided into two genetic clusters: one more similar to eastern Ontario samples, and the other consistent with samples from Manitoba and western Ontario. These immunogenetic data demonstrate a large panmictic eastern population but also show patterns of further genetic substructure. Given this observation, and what appeared to be a delineation between temporally sampled bats in Atlantic Canada, we further explored patterns of immunogenetic diversity within populations where varying temporal WNS status samples existed.

5.4.3 Immunogenetic analysis of temporally sampled M. lucifugus *in context of WNS occurrence*

5.4.3.1 Assessment of Atlantic Canadian M. lucifugus samples between two time points

Myotis lucifugus sampled within Atlantic Canada included bats present in the region before WNS introduction and early mortalities associated with the disease. Between these available temporal comparisons, we identified loci that had large frequency shifts that were unexplained by genetic drift and therefore taken to suggest evidence of selection acting on Atlantic Canadian *M. lucifugus* across sampled periods.

Evidence of strong selection is prominent in Atlantic Canada samples as made clear by observed genetic structure between PRE and EMOR samples and strong estimated coefficients of selection for several SNPs resulting in missense mutations that could affect protein function (Schaefer & Rost 2012). For example, ten SNPs led to missense mutations in the JAGN1 and HRG genes (one and nine SNPs, respectively), where both genes are implicated as playing prominent roles in antifungal components of the immune system (Erwig & Gow 2016; Khanagale et al 2018; Rydengård et al. 2008). Furthermore, a missense SNP associated with the IL12RB2 gene was also observed. The IL12RB2 gene encodes for an important receptor that is an integral component of JAK2 and STAT4 pathways increasing IFN-gamma production and aiding in immune cell pathogen recognition and destruction (van Rietschoten et al. 2000; Jorgovanovic et al. 2020; Kato-Kogoe et al. 2016; Núñez-Marrero et al. 2020). The association of these missense SNPs to these genes suggests that between the two sampled periods in Atlantic Canada, there was a strong selective pressure acting on the population, that we interpret to be exerted by a fungal pathogen.

5.4.3.2 Assessment of eastern Ontario M. lucifugus samples between three time points

Bat samples from eastern Ontario encompassed three sampled time frames, allowing for comparisons of genetic data from early WNS mortalities and bats surviving several years after WNS occurrence in the context of genetic diversity observed preceding *Pd* introduction. While no distinct patterns of genetic structure were observed between PRE, EMOR, and SURV bats, underlying shifts in genetic

diversity were observed. Similar to data observed in *M. lucifugus* samples from Atlantic Canada, these shifts included missense mutations that demonstrated strong selection coefficients and rapid frequency shifts.

When comparing PRE and EMOR samples, missense mutations were associated with three genes encompassing aspects of the major histocompatibility complex (MHC) and interleukins. Five mutations were associated with MHC, DRB1 exon-2-like genes, an extremely polymorphic region associated with antibody-mediated pathogen defences (Palmer et al. 2016; Wang et al. 2020). Other missense SNPs involve the IL1A gene and the gene encoding for one of its receptors, IL1R1. This gene encodes for trans-membrane glycoprotein allowing IL1A products to enact physiological responses as part of the immune system (Khazim et al. 2018). The cascade of physiological responses enacted by IL1A typically follows tissue damage and includes systemic inflammation and activation of the acquired immune system to eradicate foreign cells (Mistry, Savic, & van der Hilst 2017; Roerink et al. 2017). Comparing pre-WNS bats to those persisting and surviving the occurrence of WNS in eastern Ontario missense SNPs were observed instead in the ICAM1, TAP1, IL1A, and TLR3 gens. These genes cover a wide range of immune functions, including activation of lymphocytes (ICAM1), transportation of MHC-related components (TAP1), inflammation (IL1A), activation of the adaptive immune system (IL1A & TLR3), and in the case of TLR3 also recognizing fungal pathogens (Campos et al. 2019; McDonald et al. 2020; Tsilifis et al. 2021; Singh et al. 2021; Zhang et al. 2013). Interestingly, one of the missense SNPs associated with IL1A has been identified in

both comparisons and previous lesion positive/negative transcriptome work (Field et al. 2015), suggesting that this gene may play an important role in the survival of *M. lucifugus* samples during the sampled period. Within the PRE and SURV sample comparison, several SNPs were associated with the ICAM1 (including a missense SNP), IL6, and TLR8 genes, all of which have been previously implicated in *M. lucifugus* as having impacted expression in the context of WNS (Davy et al. 2020; Field et al. 2015; Field et al. 2018; Lilley et al. 2019). As such, given the lack of genetic structure observed between sampled periods with differing WNS impacts but clear shifts in genetic diversity among previously identified genes, we take these data to suggest weak selection acting upon sampled eastern Ontario *M. lucifugus*.

Direct comparison of resulting allele frequency changes between pre-WNS and the two subsequent time points suggests several candidate genes that may have influenced the survival outcome of eastern Ontario *M. lucifugus*. Frequency changes observed only in EMOR samples are presumably maladaptive substitutions resulting in bats succumbing to WNS, whereas those observed in SURV samples are associated with persistence. While there was no direct overlap in specific loci demonstrating frequency shifts unattributed to genetic drift, there were overlap in genes associated with identified SNPs. In some cases, the directionality of observed frequency shifts within these overlapping genes appeared to be opposite for the two comparisons (e.g., CCL2, IKBKG, IL13, IL23A, LY96, and TICAM1; Table S5.6) suggesting that differences in these genes, or potentially physically linked genes outside the scope of this study, may influence survival.

Further corroborating the potential importance of allele frequency shifts at immunogenetic targets described herein, previous transcriptome research comparing expression of lesion-positive and lesion-negative tissues of *M. lucifugus* has highlighted the significance of several genes consistent with those identified herein: CCL2, IL1B, IL6, IL10, IL23A, TLR2, and TLR9 (Davy et al. 2020; Field et al. 2015; Field et al. 2018; Lilley et al. 2019). Interestingly, IL6 was identified as differentially expressed in all four previously mentioned studies, combined with the frequency shifts observed herein, suggesting this gene may play an integral role in the survival of bats in eastern Ontario. Furthermore, the lack of concordance among the other identified genes and other transcriptome studies may be somewhat unsurprising given the inter-population genetic variation observed herein and elsewhere that could explain this discrepancy (Gignoux-Wolfsohn et al. 2021).

5.4.3.3 Comparing patterns of selection in Atlantic Canada and eastern Ontario M. *lucifugus samples*

Observed patterns of selection within Atlantic Canadian and eastern Ontario *M. lucifugus* appear to be much more prominent in Atlantic Canada. Despite the difference in distinct genetic clustering between sampled time points in Atlantic Canada relative to eastern Ontario, only one visualized F_{ST} outlier SNP identified between PRE and SURV samples in the eastern Ontario dataset revealed gene associations consistent with previous genetic and transcriptome studies (Davy et al. 2020; Donaldson et al. 2017; Field et al. 2015; Field et al. 2018; Lilley et al. 2019). In context of Donaldson et al. (2017), which utilized the same immunogenetic assay

with a subset of the samples analyzed herein, several outlier SNPs were associated with similar genes between the previous work and those identified herein. Of the WNS outlier SNPs identified by Donaldson et al. (2017), our analyses within Atlantic Canada *M. lucifugus* highlighted SNPs associated with IL12RB2 and TLR1 genes. Between eastern Ontario PRE and EMOR samples, loci within the IL1R1 gene were identified as outliers, contrasting the SNPs attributed to the TLR6, DDX8, and TBX21 genes identified between eastern Ontario PRE and SURV comparison. In all three of our comparisons, SNPs associated with MHC DRB1e2-like were identified, which may be unsurprising given correlations between MHC and estimates of population health (Acevedo-Whitehouse et al. 2018; Fu, Eimes, & Waldman 2023; Nayak et al. 2023). Donaldson et al. (2017) provided evidence of evolutionary rescue across the eastern extent of the *M. lucifugus* range, despite a limited sample scheme consisting of only pre-WNS samples from Manitoba and western Ontario and post-WNS samples from Atlantic Canada. Therefore, this sampling scheme did not facilitate assessments of allele frequency shifts within subpopulations that would account for local variation (i.e., local demographic or selective patterns) within the eastern metapopulation. While complete overlap with Donaldson et al. (2017) is not observed, candidate loci identified herein and previously provide further support that these genes may play an integral role in *M. lucifugus* outcomes to *Pd* exposure. It is important to note that these data only encompass immunogenetically relevant regions of the genome and other genomic intervals responsible for additional aspects of bat biology may demonstrate disparate patterns of selection acting upon

sampled populations of *M. lucifugus*. Given the context of the bat samples analyzed herein, we presume this selective pressure to be associated with selective sweeps that WNS has exerted on impacted Nearctic bat populations.

Effective population size (N_E) is a theoretical panmictic population size, lower than census estimates, that experiences the effects of genetic drift on the same scale as the actual population (Wright 1931). Evolutionary potential and population sustainability are therefore directly influenced by N_E (Frankham 2005), offering an alternative explanation for the patterns observed amongst more central (eastern Ontario) and peripheral (Atlantic Canada) populations. More specifically, the combined effects of gene flow and large N_E could have mitigated the impact of WNS on genetic diversity for eastern Ontario bats. This contrasts the smaller N_E expected in Atlantic Canada populations which may have increased the potential for stochastic effects (i.e., genetic drift). Utilizing the same 11 microsatellites herein, Davy et al. (2017) estimated N_E across identified populations and found estimates far lower than expected given the ubiquitous nature of *M. lucifugus* preceding the introduction of *P. destructans*. The identified eastern genetic cluster encompasses a large geographic area, however, there lacks targeted assessments capable of providing insights regarding differences between more central or peripheral subpopulations, relative to the larger meta-population in the context of effective population sizes. As such, we cannot exclude the potential that these intra-population patterns have resulted from an evolutionary rescue, a natural demographic rescue (predicated upon disparate N_E and subsequent emigration

from larger populations), or a combination of the two; hypotheses requiring further testing. Regardless, these data highlight that in the long-lived and incredibly mobile *M. lucifugus*, large population sizes are imperative for mitigating the impacts of novel or changing suites of selective pressures.

5.5 Conclusions

Understanding host-pathogen interactions, and associated impacts on the host population's genetic landscape, is important to better understand the ecology and evolution of a species and provide informed management and conservation efforts. White-nose syndrome has had disparate consequences on impacted Nearctic bats, where some populations experienced drastic declines and subsequent recoveries, but recovery in other populations remains uncertain or unlikely (Maslo et al. 2015; Frick et al. 2017; Dobony and Johnson 2018; Ineson 2020; Francl et al. 2012; Reynolds et al. 2016). As such, investigating shifts in host genetic diversity throughout *Pd* occurrence offers a unique opportunity to understand the relative impacts this pathogen has had on *M. lucifugus*, and elucidate mechanisms of recovery and adaptive capacity moving from an enzootic phase to an endemic one.

Herein, we continue to demonstrate the potential of an evolutionary recovery occurring in Canadian *M. lucifugus*, although variable impacts on host genetic diversity are observed between sampled regions that are consistent with suggestions from assessments of the species in the United States (Gignoux-Wolfsohn et al. 2021). These observed patterns may have arisen due to differences

in population size and associated N_E between central and peripheral subpopulations relative to the larger meta-population spanning eastern Canada. Furthermore, in some sampled regions, given the prominent gene flow, we cannot rule out a natural demographic recovery where distinct patterns of genetic structure in the context of WNS occurrence may have been masked by the influx of variants from other *M. lucifugus* populations, or the congregation of these bats into fewer hibernacula across the landscape. These data highlight the need for intra-specific management considerations, given the apparent local adaptations that may exist despite the prominent gene flow observed in the eastern extent of the Canadian *M. lucifugus* range and has the potential to impact rescue or recovery strategies aimed to increase or maintain diversity among genetically depauperate populations. As such, conservation efforts should continue to bolster or maintain large population sizes, and monitoring approaches implementing full genome strategies would provide increased resolution to explore the extent of observed inter-population disparity at non-immunogenetic loci across the *M. lucifugus* range.

Genetic Cluster	Range of alleles per locus	H_o	H_e	H_s	D_{ST}	F_{IS}	F_{IS} 95%CI	Number of Individuals
YUK Only	3-13	0.66	0.72	0.76	--	0.13	0.01 – 0.26	16
noYUK	6-24	0.67	0.76	0.76	--	0.12	0.05 – 0.18	219
ATL-CAN Only	6-18	0.68	0.68	0.77	--	0.12	0.04 – 0.20	48
E-ON Only	7-24	0.66	0.76	0.76	--	0.12	0.06 – 0.21	143
Overall	6-26	0.67	0.77	--	0.03	0.12	0.08 – 0.19	235

Table 5.1. Summary statistics for *Myotis lucifugus* genetic clusters using microsatellite loci for estimating genetic connectivity among sampled regions. Summary statistics for Yukon samples (YUK Only), all samples excluding those from Yukon (noYUK), and for each of the ATL-CAN and E-ON populations where temporal data was successfully collected. Identified are the number of alleles, observed and expected heterozygosity, intra- and interpopulation diversity (H_S and D_{ST} respectively), and estimates of F_{IS}. Values are averaged across the ten microsatellite loci used in these analyses.

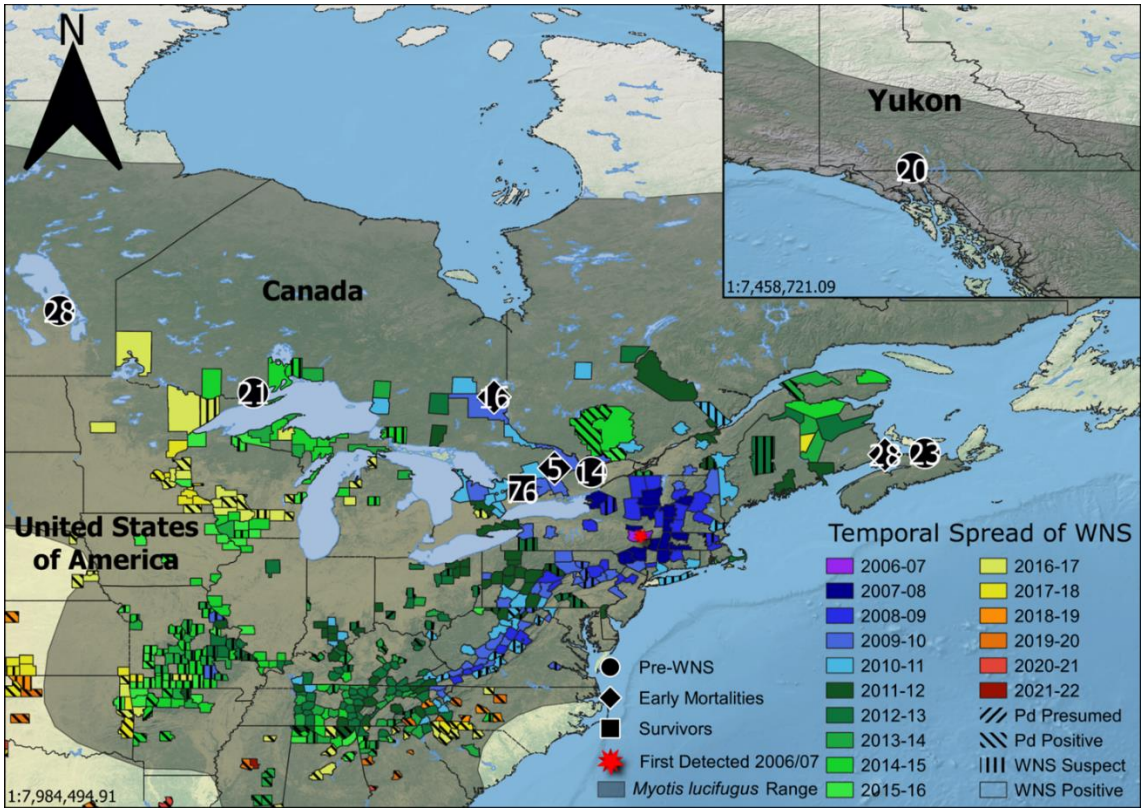

Fig 5.1. Temporal and spatial occurrence of white-nose syndrome (WNS) overlaid with the distribution of *Myotis lucifugus* and sample distributions for this study. The spread and occurrence of *P. destructans,* the causal agent of WNS, has drastically changed since its introduction to North America in 2006. To understand *M. lucifugus* population connectivity and relative impacts on neutral genetic diversity in the northern half of the range, samples were processed at 11 microsatellite markers and at immunogenetically relevant loci to identify patterns of natural selection between *M. lucifugus* populations. Samples were acquired from the regions pre-WNS (PRE; black circles), mortalities during the initial stages of disease spread (EMOR; black diamonds) and >6 years after disease introduction (SURV; black squares). This work is predicated upon the presumption that WNS imposed a strong initial selective pressure on naïve populations, supported by observed demographic declines. Numbers within shapes indicate the number of proximal samples at each location. Map created using Natural Earth and contains information licensed under the Open Government License – Ontario, The United States/U.S. Department of Commerce, The United States/U.S. Fish & Wildlife Service, and the Creative Commons Attribution 4.0 International License – Quebec; created July 2022.

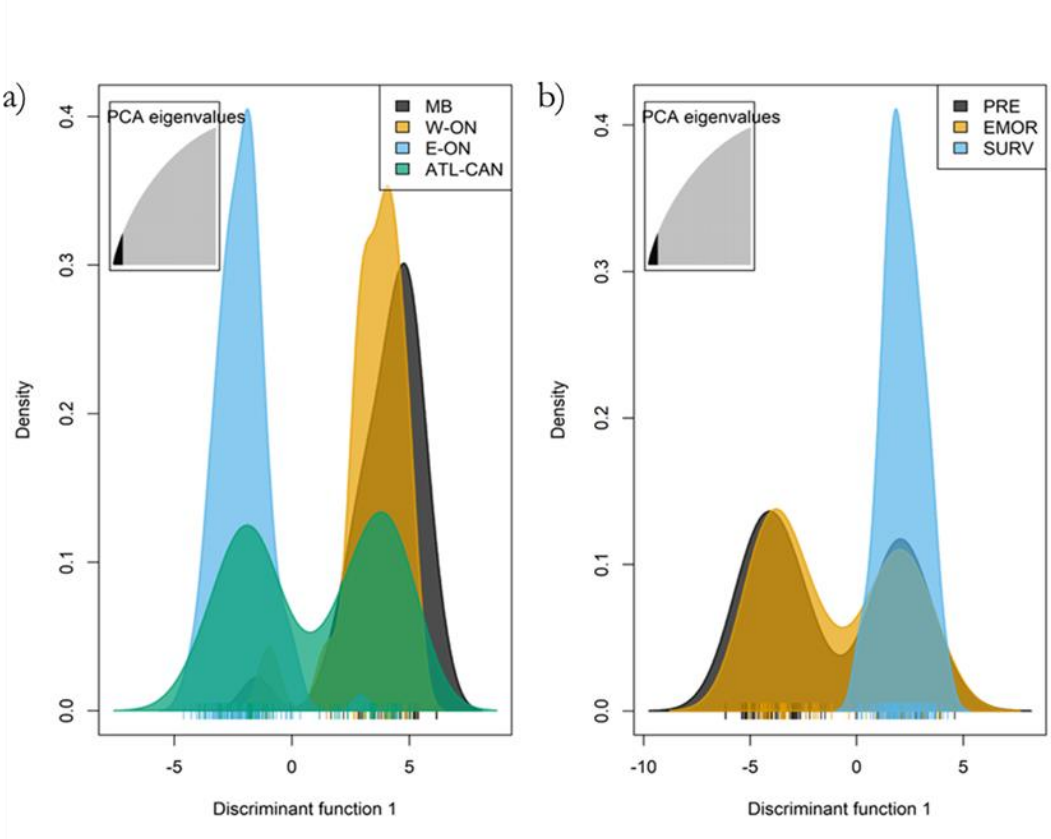

Fig 5.2. Analyses of immunogenetic loci without Yukon samples illustrate genetic substructure within the eastern neutral genetic cluster of *Myotis lucifugus*. A total of 2,886 identified immunogenetically relevant SNPs were used for visualization (after samples from Yukon were removed), demonstrating a partition within Atlantic Canada samples. Discriminant analyses of principal components are visualized in a) and b), where a) displays clustering of samples colour-coded based on geographic origin, and b) based on temporal sampling before WNS (PRE), deaths immediately following WNS-associated demographic declines (EMOR), and surviving bats >6 years post-WNS occurrence (SURV). MB = Manitoba; W-ON = western Ontario; E-ON = eastern Ontario; ATL-CAN = Atlantic Canada.

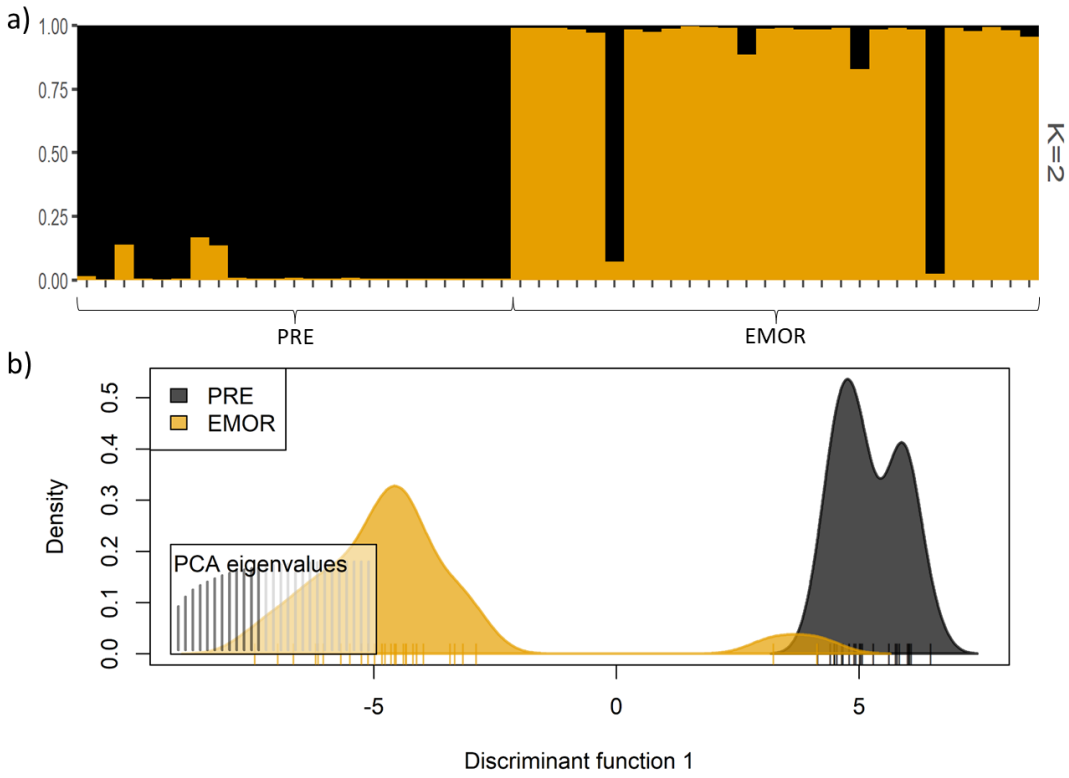

Fig 5.3. Analyses of immunogenetically relevant F_{ST} outlier SNPs from *Myotis lucifugus sampled* from Atlantic Canada show genetic structure distinguishing samples pre-WNS occurrence and early mortalities associated with the disease. Analysis of 29 immunogenetically relevant F_{ST} outlier SNPs from bats sampled before the introduction of WNS to North America (PRE), and individuals associated with early mortalities (EMOR) demonstrate clear genetic structure between the two temporally sampled categories. Inserts a) shows bar plots of estimated population assignments using STRUCTURE for K = 2, where K=2 is most supported by the Evanno method (Evanno, Regnaut, & Goudet, 2005; Fig S5.8). Each partitioned vertical bar represents an individual's proportional membership to the inferred cluster and is visualized using LOCPRIOR = 0. Insert b) demonstrates a discriminant analysis of principal components where samples are organized by temporal sampling associated with WNS.

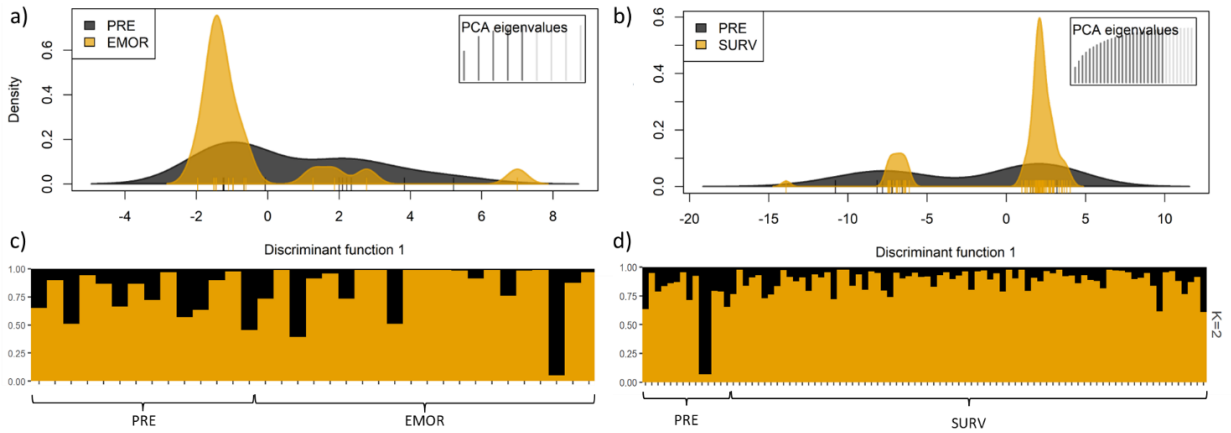

Fig 5.4. Immunogenetic outlier SNPs denote a lack of temporal genetic structure in eastern Ontario *Myotis lucifugus* when comparing pre-WNS samples and two post-WNS occurrence time points. Inserts a and c identify no immunogenetic clustering between bats sampled pre-WNS and early mortalities associated with WNS between 16 outlier SNP loci. Similar observations were made based on 37 outlier SNPs identified between bats sampled pre-WNS occurrence and those persisting several years post-WNS introduction (inserts b and d). Discriminant analyses of principal components are visualized in inserts a & b where the clustering of samples is colour-coded based on temporal sampling: pre-WNS (PRE), deaths immediately following (EMOR), and surviving bats post-WNS occurrence (SURV). Bar plots c) and d) show estimated population assignments using STRUCTURE for K = 2 using LOCPRIOR = 0, where K = 2 was most supported by Evanno method (Evanno, Regnaut, & Goudet, 2005; Fig S5.9). Each partitioned vertical bar represents an individual's proportional membership to the inferred populations.

5.6 References

1. Ancillotto L, Santini L, Ranc N, Maiorano L, Russo D. Extraordinary range expansion in a common bat: the potential roles of climate change and urbanisation. The Science of Nature. 2016 Apr;103:1-8.
2. Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
3. Auteri GG, Knowles LL. Decimated little brown bats show potential for adaptive change. Scientific Reports. 2020 Feb 20;10(1):3023.
4. Baker CM, Plein M, Shaikh R, Bode M. Simultaneous invasive alien predator eradication delivers the best outcomes for protected island species. Biological Invasions. 2020 Mar;22(3):1085-95.
5. Baker RE, Mahmud AS, Miller IF, Rajeev M, Rasambainarivo F, Rice BL, Takahashi S, Tatem AJ, Wagner CE, Wang LF, Wesolowski A. Infectious disease in an era of global change. Nature Reviews Microbiology. 2022 Apr;20(4):193-205.
6. Berry AJ, Ajioka JW, Kreitman M. Lack of polymorphism on the Drosophila fourth chromosome resulting from selection. Genetics. 1991 Dec 1;129(4):1111-7.
7. Blehert DS, Hicks AC, Behr M, Meteyer CU, Berlowski-Zier BM, Buckles EL, Coleman JT, Darling SR, Gargas A, Niver R, Okoniewski JC. Bat white-nose syndrome: an emerging fungal pathogen?. Science. 2009 Jan 9;323(5911):227-.
8. Blejwas K, Beard L, Buchanan J, Lausen CL, Neubaum D, Tobin A, Weller TJ. COULD WHITE-NOSE SYNDROME MANIFEST DIFFERENTLY IN *MYOTIS LUCIFUGUS* IN WESTERN VERSUS EASTERN REGIONS OF NORTH AMERICA? A REVIEW OF FACTORS. Journal of Wildlife Diseases. 2023 Jun 1.
9. Boddy L. Interactions with humans and other animals. InThe fungi 2016 Jan 1 (pp. 293-336). Academic Press.
10. Budiš J, Kucharík M, Ďuriš F, Gazdarica J, Zrubcová M, Ficek A, Szemes T, Brejová B, Radvanszky J. Dante: genotyping of known complex and expanded short tandem repeats. Bioinformatics. 2019 Apr 15;35(8):1310-7.
11. Burns LE, Frasier TR, Broders HG. Genetic connectivity among swarming sites in the wide ranging and recently declining little brown bat (*Myotis lucifugus*). Ecology and Evolution. 2014 Nov;4(21):4130-49.
12. Byrne AQ, Richards-Zawacki CL, Voyles J, Bi K, Ibáñez R, Rosenblum EB. Whole exome sequencing identifies the potential for genetic rescue in iconic and critically endangered Panamanian harlequin frogs. Global Change Biology. 2021 Jan;27(1):50-70.

13. Campos CF, van de Veerdonk FL, Gonçalves SM, Cunha C, Netea MG, Carvalho A. Host genetic signatures of susceptibility to fungal disease. Fungal Physiology and Immunopathogenesis. 2019:237-63.
14. Cheng TL, Reichard JD, Coleman JT, Weller TJ, Thogmartin WE, Reichert BE, Bennett AB, Broders HG, Campbell J, Etchison K, Feller DJ. The scope and severity of white-nose syndrome on hibernating bats in North America. Conservation Biology. 2021 Oct;35(5):1586-97.
15. Cryan PM, Meteyer CU, Boyles JG, Blehert DS. Wing pathology of white-nose syndrome in bats suggests life-threatening disruption of physiology. BMC biology. 2010 Dec;8:1-8.
16. Cryan PM, Meteyer CU, Boyles JG, Blehert DS. White-nose syndrome in bats: illuminating the darkness. BMC biology. 2013 Dec;11:1-4.
17. Davy CM, Martinez-Nunez F, Willis CK, Good SV. Spatial genetic structure among bat hibernacula along the leading edge of a rapidly spreading pathogen. Conservation Genetics. 2015 Oct;16:1013-24.
18. Davy CM, Donaldson ME, Rico Y, Lausen CL, Dogantzis K, Ritchie K, Willis CK, Burles DW, Jung TS, McBurney S, Park A. Prelude to a panzootic: Gene flow and immunogenetic variation in northern little brown myotis vulnerable to bat white-nose syndrome. Facets. 2017 Sep 12;2(2):690-714.
19. Davy CM, Donaldson ME, Bandouchova H, Breit AM, Dorville NA, Dzal YA, Kovacova V, Kunkel EL, Martínková N, Norquay KJ, Paterson JE. Transcriptional host–pathogen responses of Pseudogymnoascus destructans and three species of bats with white-nose syndrome. Virulence. 2020 Dec 31;11(1):781-94.
20. Dzal Y, McGuire LP, Veselka N, Fenton MB. Going, going, gone: the impact of white-nose syndrome on the summer activity of the little brown bat (*Myotis lucifugus*). Biology letters. 2011 Jun 23;7(3):392-4.
21. Dobony CA, Johnson JB. Observed resiliency of little brown myotis to long-term white-nose syndrome exposure. Journal of Fish and Wildlife Management. 2018 Jun 1;9(1):168-79.
22. Donaldson ME, Davy CM, Willis CK, McBurney S, Park A, Kyle CJ. Profiling the immunome of little brown myotis provides a yardstick for measuring the genetic response to white-nose syndrome. Evolutionary applications. 2017 Dec;10(10):1076-90.
23. Erwig LP, Gow NA. Interactions of fungal pathogens with phagocytes. Nature Reviews Microbiology. 2016 Mar;14(3):163-76.
24. Field KA, Johnson JS, Lilley TM, Reeder SM, Rogers EJ, Behr MJ, Reeder DM. The white-nose syndrome transcriptome: activation of anti-fungal host responses in wing tissue of hibernating little brown myotis. PLoS pathogens. 2015 Oct 1;11(10):e1005168.
25. Field KA, Sewall BJ, Prokkola JM, Turner GG, Gagnon MF, Lilley TM, Paul White J, Johnson JS, Hauer CL, Reeder DM. Effect of torpor on host

transcriptomic responses to a fungal pathogen in hibernating bats. Molecular ecology. 2018 Sep;27(18):3727-43.
26. Fischer NM, Altewischer A, Ranpal S, Dool S, Kerth G, Puechmaille SJ. Population genetics as a tool to elucidate pathogen reservoirs: Lessons from Pseudogymnoascus destructans, the causative agent of White-Nose disease in bats. Molecular Ecology. 2022 Jan;31(2):675-90.
27. Fisher MC, Garner TW. Chytrid fungi and global amphibian declines. Nature Reviews Microbiology. 2020 Jun;18(6):332-43.
28. Francl KE, Ford WM, Sparks DW, Brack Jr V. Capture and reproductive trends in summer bat communities in West Virginia: assessing the impact of white-nose syndrome. Journal of Fish and Wildlife Management. 2012 Jun;3(1):33-42.
29. Frankham R. Genetics and extinction. Biological conservation. 2005 Nov 1;126(2):131-40.
30. Frankham R, Ballou JD, Eldridge MD, Lacy RC, Ralls K, Dudash MR, Fenster CB. Predicting the probability of outbreeding depression. Conservation Biology. 2011 Jun;25(3):465-75.
31. Frick WF, Cheng TL, Langwig KE, Hoyt JR, Janicki AF, Parise KL, Foster JT, Kilpatrick AM. Pathogen dynamics during invasion and establishment of white-nose syndrome explain mechanisms of host persistence. 2017; 624-631.
32. Frick WF, Puechmaille SJ, Willis CK. White-nose syndrome in bats. Bats in the Anthropocene: Conservation of bats in a changing world. 2016:245-62.
33. Fu M, Eimes JA, Waldman B. Divergent allele advantage in the MHC and amphibian emerging infectious disease. Infection, Genetics and Evolution. 2023 Jul 1;111:105429.
34. Gallana M, Ryser-Degiorgis MP, Wahli T, Segner H. Climate change and infectious diseases of wildlife: altered interactions between pathogens, vectors and hosts. Current Zoology. 2013 Jun 1;59(3):427-37.
35. Gignoux-Wolfsohn SA, Pinsky ML, Kerwin K, Herzog C, Hall M, Bennett AB, Fefferman NH, Maslo B. Genomic signatures of selection in bats surviving white-nose syndrome. Molecular ecology. 2021 Nov;30(22):5643-57.
36. Haworth, S. E., Nituch, L., Northrup, J. M., & Shafer, A. B. (2021). Characterizing the demographic history and prion protein variation to infer susceptibility to chronic wasting disease in a naïve population of white-tailed deer (Odocoileus virginianus). Evolutionary applications, 14(6), 1528-1539.
37. Hayman DT, Pulliam JR, Marshall JC, Cryan PM, Webb CT. Environment, host, and fungal traits predict continental-scale white-nose syndrome in bats. Science advances. 2016 Jan 29;2(1):e1500831.

38. Heard GW, Thomas CD, Hodgson JA, Scroggie MP, Ramsey DS, Clemann N. Refugia and connectivity sustain amphibian metapopulations afflicted by disease. Ecology Letters. 2015 Aug;18(8):853-63.
39. Hohenlohe PA, Funk WC, Rajora OP. Population genomics for wildlife conservation and management. Molecular Ecology. 2021 Jan;30(1):62-82.
40. Hooton LA, Adams AA, Cameron A, Fraser EE, Hale L, Kingston S, Fenton MB, McGuire LP, Stukenholtz EE, Davy CM. Effects of bat white-nose syndrome on hibernation and swarming aggregations of bats in Ontario. Canadian Journal of Zoology. 2023 May 29;101(10):886-95.
41. Hoyt JR, Sun K, Parise KL, Lu G, Langwig KE, Jiang T, Yang S, Frick WF, Kilpatrick AM, Foster JT, Feng J. Widespread bat white-nose syndrome fungus, northeastern China. Emerging Infectious Diseases. 2016 Jan;22(1):140.
42. Hufbauer RA, Szűcs M, Kasyon E, Youngberg C, Koontz MJ, Richards C, Tuff T, Melbourne BA. Three types of rescue can avert extinction in a changing environment. Proceedings of the National Academy of Sciences. 2015 Aug 18;112(33):10557-62.
43. Ineson KM. Demography of a recovery: tracking the rebound of little brown bat (*Myotis lucifugus*) populations (Doctoral dissertation, University of New Hampshire).
44. Johnson LN, McLeod BA, Burns LE, Arseneault K, Frasier TR, Broders HG. Population genetic structure within and among seasonal site types in the little brown bat (*Myotis lucifugus)* and the northern long-eared bat (*M. septentrionalis*). PLoS One. 2015 May 5;10(5):e0126309.
45. Jolly CJ, Smart AS, Moreen J, Webb JK, Gillespie GR, Phillips BL. Trophic cascade driven by behavioral fine-tuning as naïve prey rapidly adjust to a novel predator. Ecology. 2021 Jul;102(7):e03363.
46. Jorgovanovic D, Song M, Wang L, Zhang Y. Roles of IFN-γ in tumor progression and regression: a review. Biomark Res. 2020; 8: 49.
47. Kanarek AR, Webb CT, Barfield M, Holt RD. Overcoming Allee effects through evolutionary, genetic, and demographic rescue. Journal of biological dynamics. 2015 Jan 1;9(1):15-33.
48. Kato-Kogoe N, Ohyama H, Okano S, Yamanegi K, Yamada N, Hata M, Nishiura H, Abiko Y, Terada N, Nakasho K. Functional analysis of differences in transcriptional activity conferred by genetic variants in the 5' flanking region of the IL12RB2 gene. Immunogenetics. 2016 Jan;68:55-65.
49. Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. Ecology letters, 7(12), 1225-1241.
50. Khandagale A, Lazzaretto B, Carlsson G, Sundin M, Shafeeq S, Römling U, Fadeel B. JAGN1 is required for fungal killing in neutrophil extracellular traps: implications for severe congenital neutropenia. Journal of Leukocyte Biology. 2018 Dec;104(6):1199-213.

51. Khazim K, Azulay EE, Kristal B, Cohen I. Interleukin 1 gene polymorphism and susceptibility to disease. Immunological Reviews. 2018 Jan;281(1):40-56.
52. Knapp RA, Wilber MQ, Byrne AQ, Joseph MB, Smith TC, Rothstein AP, Grasso RL, Rosenblum EB. Reintroduction of resistant frogs facilitates landscape-scale recovery in the presence of a lethal fungal disease. bioRxiv. 2023 May 24:2023-05.
53. Kovacova V, Zukal J, Bandouchova H, Botvinkin AD, Harazim M, Martínková N, Orlov OL, Piacek V, Shumkina AP, Tiunov MP, Pikula J. White-nose syndrome detected in bats over an extensive area of Russia. BMC veterinary research. 2018 Dec;14:1-9.
54. Langmaier M, Lapin K. A systematic review of the impact of invasive alien plants on forest regeneration in European temperate forests. Frontiers in Plant Science. 2020 Sep 3;11:524969.
55. Langwig KE, Frick WF, Reynolds R, Parise KL, Drees KP, Hoyt JR, Cheng TL, Kunz TH, Foster JT, Kilpatrick AM. Host and pathogen ecology drive the seasonal dynamics of a fungal disease, white-nose syndrome. Proceedings of the Royal Society B: Biological Sciences. 2015 Jan 22;282(1799):20142335.
56. Langwig KE, Frick WF, Bried JT, Hicks AC, Kunz TH, Marm Kilpatrick A. Sociality, density-dependence and microclimates determine the persistence of populations suffering from a novel fungal disease, white-nose syndrome. Ecology letters. 2012 Sep;15(9):1050-7.
57. Lilley TM, Prokkola JM, Blomberg AS, Paterson S, Johnson JS, Turner GG, Bartonička T, Bachorec E, Reeder DM, Field KA. Resistance is futile: RNA-sequencing reveals differing responses to bat fungal pathogen in Nearctic *Myotis lucifugus* and Palearctic *Myotis myotis*. Oecologia. 2019 Oct;191:295-309.
58. Lilley TM, Wilson IW, Field KA, Reeder DM, Vodzak ME, Turner GG, Kurta A, Blomberg AS, Hoff S, Herzog CJ, Sewall BJ. Genome-wide changes in genetic diversity in a population of *Myotis lucifugus* affected by white-nose syndrome. G3: Genes, Genomes, Genetics. 2020 Jun 1;10(6):2007-20.
59. Leopardi S, Blake D, Puechmaille SJ. White-nose syndrome fungus introduced from Europe to North America. Current Biology. 2015 Mar 16;25(6):R217-9.
60. Longo, A. V., Burrowes, P. A., & Zamudio, K. R. (2014). Genomic studies of disease-outcome in host–pathogen dynamics. American Zoologist, 54(3), 427-438.
61. Lozier JD, Zayed A. Bee conservation in the age of genomics. Conservation Genetics. 2017 Jun;18:713-29.
62. Maslo B, Valent M, Gumbs JF, Frick WF. Conservation implications of ameliorating survival of little brown bats with white-nose syndrome. Ecological Applications. 2015 Oct;25(7):1832-40.

63. McDonald CA, Longo AV, Lips KR, Zamudio KR. Incapacitating effects of fungal coinfection in a novel pathogen system. Molecular Ecology. 2020 Sep;29(17):3173-86.
64. Miller-Butterworth CM, Vonhof MJ, Rosenstern J, Turner GG, Russell AL. Genetic structure of little brown bats (*Myotis lucifugus*) corresponds with spread of white-nose syndrome among hibernacula. Journal of Heredity. 2014 May;105(3):354-64.
65. Mistry A, Savic S, van der Hilst JCH.Interleukin-1 Blockade: An Update on Emerging Indications. BioDrugs. 2017;31(3):207–221. Epub2017/05/13.https://doi.org/10.1007/s40259-017-0224-7PMID: 28497222.
66. Neaves LE, Zenger KR, Prince RI, Eldridge MD, Cooper DW. Landscape discontinuities influence gene flow and genetic structure in a large, vagile Australian mammal, Macropus fuliginosus. Molecular Ecology. 2009 Aug;18(16):3363-78.
67. Núñez-Marrero A, Arroyo N, Godoy L, Rahman MZ, Matta JL, Dutil J. SNPs in the interleukin-12 signaling pathway are associated with breast cancer risk in Puerto Rican women. Oncotarget. 2020 Sep 9;11(37):3420.
68. Nayak SS, Panigrahi M, Kumar H, Rajawat D, Sharma A, Bhushan B, Dutt T. Evidence for selective sweeps in the MHC gene repertoire of various cattle breeds. Animal Biotechnology. 2023 Dec 15;34(8):4167-73.
69. O'Leary SJ, Puritz JB, Willis SC, Hollenbeck CM, Portnoy DS. These aren't the loci you'e looking for: Principles of effective SNP filtering for molecular ecologists.
70. Palmer JM, Berkman LK, Marquardt PE, Donner DM, Jusino MA, Lindner DL. Preliminary characterization of little brown bats (*Myotis lucifugus*) immune MHC II DRB alleles using next-generation sequencing. PeerJ PrePrints. 2016 Jan 21;4:e1662v1.
71. Peakall R, Ruibal M, Lindenmayer DB. Spatial autocorrelation analysis offers new insights into gene flow in the Australian bush rat, Rattus fuscipes. Evolution. 2003 May;57(5):1182-95.
72. Pina-Martins F, Baptista J, Pappas Jr G, Paulo OS. New insights into adaptation and population structure of cork oak using genotyping by sequencing. Global Change Biology. 2019 Jan;25(1):337-50.
73. Moore MS, Reichard JD, Murtha TD, Nabhan ML, Pian RE, Ferreira JS, Kunz TH. Hibernating little brown myotis (*Myotis lucifugus*) show variable immunological responses to white-nose syndrome. PLoS One. 2013 Mar 20;8(3):e58976.
74. Puechmaille SJ, Wibbelt G, Korn V, Fuller H, Forget F, Mühldorfer K, Kurth A, Bogdanowicz W, Borel C, Bosch T, Cherezy T. Pan-European distribution of white-nose syndrome fungus (Geomyces destructans) not associated with mass mortality. PloS one. 2011 Apr 27;6(4):e19167.
75. Reeder DM, Frank CL, Turner GG, Meteyer CU, Kurta A, Britzke ER, Vodzak ME, Darling SR, Stihler CW, Hicks AC, Jacob R. Frequent arousal from

hibernation linked to severity of infection and mortality in bats with white-nose syndrome. PloS one. 2012 Jun 20;7(6):e38920.
76. Reynolds RJ, Powers KE, Orndorff W, Ford WM, Hobson CS. Changes in rates of capture and demographics of *Myotis septentrionalis* (northern long-eared bat) in western Virginia before and after onset of white-nose syndrome. Northeastern Naturalist. 2016 Jun;23(2):195-204.
77. Ricciardi A, Hoopes MF, Marchetti MP, Lockwood JL. Progress toward understanding the ecological impacts of nonnative species. Ecological monographs. 2013 Aug;83(3):263-82.
78. Rochlin I, Ninivaggi DV, Hutchinson ML, Farajollahi A. Climate change and range expansion of the Asian tiger mosquito (Aedes albopictus) in Northeastern USA: implications for public health practitioners. PloS one. 2013 Apr 2;8(4):e60874.
79. Roerink ME, vanderSchaaf ME, Dinarello CA, Knoop H, vanderMeer JW. Interleukin-1 as a mediator of fatigue in disease: a narrative review. J Neuro inflammation. 2017;14(1):16. Epub2017/01/22. https://doi.org/10.1186/s12974-017-0796-7PMID:28109186;PubMedCentralPMCID: PMCPmc5251329.
80. Rydengård V, Shannon O, Lundqvist K, Kacprzyk L, Chalupka A, Olsson AK, Mörgelin M, Jahnen-Dechent W, Malmsten M, Schmidtchen A. Histidine-rich glycoprotein protects from systemic Candida infection. PLoS pathogens. 2008 Aug 1;4(8):e1000116.
81. Schweizer, R. M., Robinson, J., Harrigan, R., Silva, P., Galverni, M., Musiani, M., ... & Wayne, R. K. (2016). Targeted capture and resequencing of 1040 genes reveal environmentally driven functional variation in grey wolves. Molecular ecology, 25(1), 357-379.
82. Singh M, Thakur M, Mishra M, Yadav M, Vibhuti R, Menon AM, Nagda G, Dwivedi VP, Dakal TC, Yadav V. Gene regulation of intracellular adhesion molecule-1 (ICAM-1): A molecule with multiple functions. Immunology letters. 2021 Dec 1;240:123-36.
83. Smith D, O'Brien D, Hall J, Sergeant C, Brookes LM, Harrison XA, Garner TW, Jehle R. Challenging a host–pathogen paradigm: Susceptibility to chytridiomycosis is decoupled from genetic erosion. Journal of Evolutionary Biology. 2022 Apr 1;35(4):589-98.
84. Supple, M. A., & Shapiro, B. (2018). Conservation of biodiversity in the genomics era. Genome biology, 19(1), 1-12.
85. Templeton AR, Hemmer H, Mace G, Seal US, Shields WM, Woodruff DS. Local adaptation, coadaptation, and population boundaries. Zoo Biology. 1986;5(2):115-25.
86. Thompson TQ, Bellinger MR, O'Rourke SM, Prince DJ, Stevenson AE, Rodrigues AT, Sloat MR, Speller CF, Yang DY, Butler VL, Banks MA. Anthropogenic habitat alteration leads to rapid loss of adaptive variation and restoration potential in wild salmon populations. Proceedings of the National Academy of Sciences. 2019 Jan 2;116(1):177-86.

87. Trumbo DR, Hardy BM, Crockett HJ, Muths E, Forester BR, Cheek RG, Zimmerman SJ, Corey-Rivas S, Bailey LL, Funk WC. Conservation genomics of an endangered montane amphibian reveals low population structure, low genomic diversity and selection pressure from disease. Molecular Ecology. 2023 Dec;32(24):6777-95.
88. Tsilifis C, Moreira D, Marques L, Neves E, Slatter MA, Gennery AR. Stem cell transplantation as treatment for major histocompatibility class I deficiency. Clinical Immunology. 2021 Aug 1;229:108801.
89. Vander Wal E, Garant D, Calmé S, Chapman CA, Festa-Bianchet M, Millien V, Rioux-Paquette S, Pelletier F. Applying evolutionary concepts to wildlife disease ecology and management. Evolutionary applications. 2014 Aug;7(7):856-68.
90. van Rietschoten JG, Smits HH, Westland R, Verweij CL, den Hartog MT, Wierenga EA. Genomic organization of the human interleukin-12 receptor β 2-chain gene. Immunogenetics. 2000 Jan;51:30-6.
91. Verant ML, Boyles JG, Waldrep Jr W, Wibbelt G, Blehert DS. Temperature-dependent growth of Geomyces destructans, the fungus that causes bat white-nose syndrome.
92. Verant ML, Meteyer CU, Speakman JR, Cryan PM, Lorch JM, Blehert DS. White-nose syndrome initiates a cascade of physiologic disturbances in the hibernating bat host. BMC physiology. 2014 Dec;14:1-1.
93. Vonhof MJ, Russell AL, Miller-Butterworth CM. Range-wide genetic analysis of little brown bat (*Myotis lucifugus*) populations: estimating the risk of spread of white-nose syndrome. PLoS One. 2015 Jul 8;10(7):e0128713.
94. Walker SF, Bosch J, James TY, Litvintseva AP, Valls JA, Piña S, García G, Rosa GA, Cunningham AA, Hole S, Griffiths R. Invasive pathogens threaten species recovery programs. Current Biology. 2008 Sep 23;18(18):R853-4.
95. Wang K, Liu X, Li Q, Wan K, Gao R, Han G, Li C, Xu M, Jia B, Shen X. MHC-DRB1 exon 2 polymorphism and its association with mycoplasma ovipneumonia resistance or susceptibility genotypes in sheep. Journal of genetics. 2020 Dec;99:1-2.
96. Wilder AP, Frick WF, Langwig KE, Kunz TH. Risk factors associated with mortality from white-nose syndrome among hibernating bat colonies. Biology Letters. 2011 Dec 23;7(6):950-3.
97. Wilder AP, Kunz TH, Sorenson MD. Population genetic structure of a common host predicts the spread of white-nose syndrome, an emerging infectious disease in bats. Molecular ecology. 2015 Nov;24(22):5495-506.
98. Willi Y, Kristensen TN, Sgrò CM, Weeks AR, Ørsted M, Hoffmann AA. Conservation genetics as a management tool: The five best-supported paradigms to assist the management of threatened species. Proceedings of the National Academy of Sciences. 2022 Jan 4;119(1):e2105076119.

99. Whiting-Fawcett F, Field KA, Puechmaille SJ, Blomberg AS, Lilley TM. Heterothermy and antifungal responses in bats. Current Opinion in Microbiology. 2021 Aug 1;62:61-7.
100. Wright B, Morris K, Grueber CE, Willet CE, Gooley R, Hogg CJ, O'Meally D, Hamede R, Jones M, Wade C, Belov K. Development of a SNP-based assay for measuring genetic diversity in the Tasmanian devil insurance population. BMC genomics. 2015 Dec;16:1-1.
101. Wright S. Evolution in Mendelian populations. Genetics. 1931 Mar;16(2):97.
102. Xuereb A, d'Aloia CC, Andrello M, Bernatchez L, Fortin MJ. Incorporating putatively neutral and adaptive genomic data into marine conservation planning. Conservation Biology. 2021 Jun;35(3):909-20.
103. Yıldırım Y, Tinnert J, Forsman A. Contrasting patterns of neutral and functional genetic diversity in stable and disturbed environments. Ecology and Evolution. 2018 Dec;8(23):12073-89.
104. Zhang SY, Herman M, Ciancanelli MJ, de Diego RP, Sancho-Shimizu V, Abel L, Casanova JL. TLR3 immunity to infection in mice and humans. Current opinion in immunology. 2013 Feb 1;25(1):19-33.

CHAPTER 6

6.1 General Discussion

The overarching objective of my thesis was to understand the relative impacts of enzootic/epizootic disease on host immunogenetic diversity and the underlying role this has on the persistence of host populations, thus providing insights required for informed conservation efforts. To achieve this goal, I sequenced immunogenetically relevant regions of several North American mammals associated with a) a long-standing enzootic disease, rabies, relevant to terrestrial meso-carnivores and b) an invasive, epizootic disease, white-nose syndrome that differentially impacts bat species and populations. Specifically, I identified shifts in genetic diversity indicative of selective sweeps associated with viral variant distributions, or pathogen introduction, providing insights into underlying disease dynamics. By investigating signatures of selection imposed by an enzootic disease on an invasive host, and a host with a long-standing pathogen relationship, I was able to evaluate the presence of potential co-evolutionary mechanisms and identify the likely mechanism behind the distribution of disease variants. By exploring the differential impacts of *Pd* exposure among bat populations, I was able to elucidate further evidence of the varied responses of Nearctic bats to this pathogen and the importance of connectivity and large population sizes in population recovery subsequent to strong selective sweeps and associated demographic declines from disease. Overall, these data increase understanding of host-pathogen interactions and factors that influence the impact on host populations, informing management and conservation efforts as persistent anthropogenic and environmental changes continue to influence and alter disease dynamics.

In **Chapters 2** and **3**, I examined host populations of red and arctic foxes (*Vulpes vulpes* and *V. lagopus,* respectively), to explore knowledge gaps in our understanding of how

unique distributions of enzootic arctic rabies viral variants are maintained and provided insights into how continued environmental change may impact this disease system. In **Chapter 2,** I identified weak to moderate genetic structure of 30 outlier loci, associated with genes previously implicated in arctic rabies outcome, demonstrating that red foxes do not exhibit patterns of local adaptation to specific viral variants but rather between regions where the disease is present or absent. This observation contrasts with **Chapter 3**, where 22 outlier loci demonstrate a higher genetic affinity between foxes inhabiting environments, where the same arctic rabies variant circulates, despite the panmictic nature of the species. These data suggest arctic foxes display patterns consistent with expectations of a coevolutionary hot spot, where both pathogen and host impart some level of selection on each other. Conversely, red foxes likely represent a coevolutionary cold spot given the lack of patterns associated with specific viral variants, and the apparent genetic structure of this host between rabies-free and enzootic regions. Given the prolonged generational exposure of arctic foxes to the disease and more recent exposures of red foxes in Alaska due to continued northward range expansions, these observations were largely expected (Berteaux et al. 2015; Mørk and Prestrud 2004). Although unsurprising, observations distinguishing co-evolutionary hot and cold spots may prove informative for management efforts. Identifying where co-evolutionary cold spots are occurring could reveal locations in which host organisms are likely to face strong selective sweeps (i.e., increased likelihood of population impacts) or populations that may be less affected in the event of hot spots. As such, excluding viral variant one, the current dynamics of arctic rabies in North America seem directly tied to the arctic fox, where red fox acts as a maintenance host. However, as environmental change continues to cause range shifts in both fox species, there remains potential that a host transition/shift may occur from arctic to red foxes in Alaska. Such

patterns of host shifting have already been observed in southern Ontario, where arctic rabies viral variant one shifted from arctic foxes as the main reservoir/host to skunk and red fox populations (MacInnes et al. 2001; Mørk and Prestrud 2004; Rosatte et al. 2007). While data from **Chapters 2** and **3** provide a glimpse into current disease dynamics of arctic rabies along northern coasts of North America, they are prone to sampling bias resulting from logistical and financial difficulties associated with sample collection and monitoring in Arctic regions (Mallory et al. 2018; Trivedi et al. 2022). As such, future work should aim to increase sampling for both species of fox across the entirety of the disease's distribution. Including samples known to have survived exposure to rabies, and those that have succumbed would facilitate investigations of patterns of selection in the context of discrete survival and mortality phenotypes and add to the understanding of arctic rabies dynamics. While the arctic rabies viral variants have distinct lineages and distributions, what remains uncertain is if, and how, these lineages differ in their pathogenicity or virulence (Huettmann et al. 2017; Goldsmith et al. 2016; Kuzmin et al. 2008; Nadin-Davis, Sheen, and Wandeler 2012). Further investigations into the differences in arctic rabies viral variants could increase resolution and add to the co-evolutionary dynamics proposed herein.

Approaches similar to those presented in this thesis could be used to investigate other emerging infectious diseases or systems with altered dynamics following anthropogenic or climactic changes. For example, sarcoptic mange, an infestation of the epidermis by the mite *Sarcoptes scabiei*, infects most fox species (Mörner 1992; Niedringhaus et al. 2019; Escobar et al. 2022). Mange is experiencing a global expansion, where previously documented epizootics occur cyclically every 30-45 years in North American canids, and sporadically across European countries (Mörner 1992; Niedringhaus et al. 2019; Escobar et al. 2022). These contrasting disease cycles are important to note, as

some studies have shown previously infected populations demonstrate some level of resistance to the pathogen in subsequent years (Niedringhaus et al. 2019). Varied levels of resistance raise the question of whether there are ongoing co-evolutionary hot and cold spots that may highlight areas requiring additional monitoring or surveillance. While the mite associated with mange is typically unable to survive in colder environments, climate change and range expansions of frequent host species such as red foxes continue; potentially altering the frequency and/or severity of epizootic events. As such, we can better inform future management efforts by profiling host populations to understand current dynamics, impacts on host populations, and disease-susceptible populations or species, for mange and other diseases with the potential to elicit strong selective sweeps.

In **Chapters 4** and **5**, I investigated patterns of selection associated with an invasive fungal pathogen, *Pseudogymnoascus destructans* (*Pd*) responsible for white-nose syndrome (WNS) and drastic population declines among several Nearctic bat species. While much has been learned since the initial introduction and continued spread of *Pd* in 2006/2007, there remained several knowledge gaps in our understanding the consequence of this fungus on some Nearctic bat species and associated underlying mechanisms (life history, demographic or evolutionary) that led to varied species and population impacts.

In **Chapter 4,** immunogenetic data supported the hypothesis that *Myotis leibii* were inherently resistant to WNS upon introduction and did not undergo shifts in genetic diversity as would be expected following population declines associated with a selective sweep. These data are critical to our understanding of WNS susceptibility among Nearctic bats, as the inherent resistance/tolerance to *Pd* observed in *M. leibii* contrasts with documented evolutionary rescues among other populations. While these genomic data demonstrate stable population trends in *M. leibii* since the introduction of *Pd*, transcriptome studies,

similar to those undertaken in susceptible species, would be beneficial to highlight differences in gene expression resulting in the apparent increased resistance or tolerance of this species (Davy et al. 2020; Field et al. 2015; Field et al. 2018; Lilley et al. 2019). Transcriptomic and proteomic studies on *M. leibii* would further add to our understanding of how, and if, genetic variants identified herein affect transcription or underlying protein function, but also facilitate comparisons to other Nearctic species and the type of immune response these bats undergo (i.e., systemic or localized as observed in *M. lucifugus and E. fuscus,* respectively; Davy et al. 2020). In **Chapter 5**, I evaluate temporal shifts in the genetic landscape of *M. lucifugus* associated with the introduction and spread of *Pd* where I found differential shifts in genetic diversity between populations from eastern Ontario and Atlantic Canada. These patterns of varied local impacts are consistent with those observed among other populations of *M. lucifugus* in the United States (Gignoux-Wolfsohn et al. 2021). There is no evidence supporting host-imposed reciprocal selection upon *Pd* since its introduction (Drees et al. 2017). Further, evidence of a potential evolutionary rescue amongst bat populations is widespread within the literature and evident from data presented herein (Auteri & Knowles 2020; Davy et al. 2017; Donaldson et al. 2017; Lilley et al. 2020). The apparent recovery of some populations, despite the continued persistence of *Pd* in the environment, may represent a phase shift from epizootic to enzootic (Frank, Davis, and Herzog 2019). Diversification of *Pd* lineages with time may have future conservation and management implications for bats if another epizootic event occurs as a result of novel strains of the fungal pathogen adapting to circumvent host defence mechanisms.

There are a number of genomic and transcriptomic studies focused on *M. lucifugus* susceptibility to WNS, identifying frequency shifts at specific loci or varied gene expression between infected/non-infected bats (Auteri & Knowles 2020; Blejwas et al. 2023; Davy et al.

2017; Donaldson et al. 2017; Lilley et al. 2020; Davy et al. 2020; Lilley et al 2019). Incongruencies in the loci exhibiting allele frequency shifts, or genes with differing expression levels between these genomic and transcriptomic data highlight the complex nature of the immune response and host-pathogen interactions. While some consistency between the location of genetic variants and gene expression associated with disease outcome in bats exists, it is clear that not all populations or species will be impacted similarly, as hypothesized for western *M. lucifugus* relative to their eastern conspecifics (Auteri & Knowles 2020; Blejwas et al. 2023; Davy et al. 2017; Donaldson et al. 2017; Lilley et al. 2020; Davy et al. 2020; Lilley et al 2019).

Combined with past research, my thesis demonstrates that patterns of demographic/life histories and local suites of selective pressures influence the relative WNS consequences bat populations experience, leading to the observed disparate patterns among species and populations. As *Pd* continues expanding westward, understanding that disparate impacts may arise due to environmental usage and interactions becomes imperative for naïve bat populations, which will likely demonstrate unique signatures of selection given ecological and behavioural differences. Furthermore, my thesis highlights why extrapolating host-pathogen insights from one population/species to another may not always be appropriate and could result in inefficient or ineffective management efforts. As such, population-level assessments are encouraged to provide a complete understanding of the relative impacts imposed on host populations and how host-pathogen relationships may change over time.

6.2 Limitations

My thesis puts heavy reliance on host immunogenetic assays to provide partial insight into disease dynamics by identifying shifts in genetic diversity indicative of selection.

While these assays are extremely cost-effective and informative, reduced representation approaches lack holistic understanding of the genetic landscape of organisms. Given the disease-centered focus employed herein, immunogenetic assays would be sufficient to investigate shifts in genetic diversity associated with the immune response. However, organisms and their responses to disease are much more complex. As such, different signatures of selection may be observed at regions of the genome associated with other relevant biological processes influencing disease outcomes. For example, shifts in genetic diversity among non-immunogenetically relevant genomic regions, such as genes associated with metabolism or hibernation, have been observed in bats impacted by *Pd* (Auteri and Knowles 2020; Lilley et al. 2020; Gignoux-Wolfsohn et al. 2021). Furthermore, variants in regulatory regions that influence gene expression can sometimes be far up/downstream to genes and not captured by these assays, yet have critical functional expression implications (Uslu et al. 2014). Therefore, more frequent adoption of whole genome resequencing should be undertaken in these systems for further investigations to address questions beyond the scope of the methodological approach used in this thesis. Adopting whole genome sequencing methods may prove difficult for some systems, due to lacking or insufficient genomic resources that can facilitate an understanding of the physical linkage of identified loci. As was the case for my research, where the *V. vulpes* genome in **Chapters 2** and **3**, and the *M. lucifugus* genome in **Chapters 4** and **5** were used for aligning sequence data. As such, for both *V. lagopus* and *M. leibii* sequence data, filtering steps accounting for linkage disequilibrium may be incorrect and loci may have been erroneously removed or retained, potentially impacting observed patterns.

Beyond the potential limitations of reduced representation sequencing assays, the methods implemented herein are likely overly conservative as they relate to SNP filtering.

Future work should reevaluate filtering parameters to optimize both the number and potential impact of a wider breadth of loci. Such a re-analysis of these data could have large impacts on the patterns observed given many outlier loci were putatively associated with promoter or regulatory regions. In Chapters 2 & 3, specifically, re-assessing how data were filtered and classified (on-target vs. off-target), as it relates to the identification of paralogs or copy number variants, may reveal further loci of interest.

The identification of outlier loci was a large component of my research. As methods for outlier analyses continue to evolve, the methods I implemented were perhaps too stringent on initial data filtering prior to outlier detection, and inadequately stringent during detection, enhancing the chances of false positive outlier identifications. Recent studies have implemented a combination of outlier detection methods (differention-based methods as implemented herein, and genotype-environment associations) and further analyzed only the intersect of identified outliers between both methods (Shi et al. 2023). Implementing a similar approach, by using several different programs and methods to identify outlier loci would increase the confidence surrounding designating a locus as an outlier and may further clarify or emphasize those loci that are strongly associated with disease variables on the landscape (prescence/absence of disease or disease variants).

Another limitation of my research is sample design given the pragmatic realities in obtaining samples. In **Chapters 2** and **3**, given the difficulties associated with obtaining samples from Arctic regions, it was not feasible to obtain rabies survivor and mortality samples to adequately evaluate the underlying mechanisms of rabies outcomes in these foxes. In contrast, for **Chapters 4** and **5** discrete host phenotypes were obtained, although sample sizes from some geographic regions or phenotypes were lacking. While sampling

numbers were low in some instances, they likely did not impact my ability to draw meaningful conclusions, there remains potential that increased sampling efforts and sizes in all data chapters may improve the resolution of the patterns of selection observed. Re-evaluations of these data would include power analyses to further assess this assumption. Finally, my investigations for disease impacts are host-centred and do not facilitate an assessment of reciprocal selection from the host species onto the respective pathogen's genome. Given that *Pd* can survive on a spectrum of media in caves outside of a bat host for some period of time (Fischer, Dool, and Puechmaille 2020), and has largely accumulated no genetic variation since its introduction to North America (Drees et al. 2017) evidence of reciprocal selection may be limited. However, as it concerns arctic rabies, where it remains unknown if variants have differences in pathogenicity, the existence of reciprocal selection remains plausible and requires further investigation.

Finally, the use of both Wright-Fisher and selection coefficient modeling may require future optimization for the studed systems herein. While I attempted to use these methods to account for the influence of genetic drift on observed frequency shifts, or model the selection imposed on specific genotypes, it is important to note that several assumptions of these test were violated. This includes the potential for overlapping generations and dynamic population sizes in the Wright-Fisher model, or other violoations to HWE as is assumed by the selection coefficient model. Further, because of these violations, results associated with these tests may be prone to error. As such, future work should refine these models in an attempt to account for these variables to better elucidate the true effects of genetic drift, or strength of selection, on genotypes between and within populations.

6.3 Conclusions

The outcomes of the work completed for this thesis further our understanding of disease dynamics in North America and the relative impacts of disease on host populations, highlighting the need to identify and evaluate species on a population level for conservation and management efforts. The interplay and resulting mosaics arising from varied demographic processes (e.g., gene flow and drift), suites of selective pressures populations experience, differing life histories, and types of disease systems (invasive and epizootic vs. endemic and enzootic) are critical for our ability to predict disease outcomes, identify susceptible populations, and determine the capacity of organisms to recover following disease mortalities through genetic, demographic, or evolutionary rescues. Furthermore, the ability to identify co-evolutionary cold spots amongst host-pathogen interactions will provide a means to understand why some populations experience strong selective sweeps from infectious diseases compared to others. These data become increasingly important in the context of invasive pathogens on naïve hosts or inversely, invasive hosts entering disease-established environments where they can disrupt existing dynamics leading to increased morbidity in natural or other spillover hosts. These examples are demonstrated by the impacted role of invasive Australian brushtail possums (*Trichosurus vulpecula*) transmitting *Mycobacterium bovis* between livestock and wildlife (Nugent, Buddie & Knowles 2015), or invasive raccoon dogs (*Nyctereutes procyonoides*) responsible for rabies circulation and subsequent re-emergence's of the pathogen in disease-free countries (Singer et al. 2009). Obtaining an understanding of the extent of change imposed by invasive hosts and/or invasive pathogens on natural populations before it occurs is therefore critical to mitigate epizootic and spillover events.

In order to best inform management and conservation efforts, researchers will need to employ the full spectra of omics technologies (e.g., genomics, transcriptomics, proteomics, and epigenomics) for a holistic understanding of current disease dynamics, and how continued change may impact existing or novel host-pathogen interactions. As environmental change and globalization continue, our understanding of disease dynamics and the factors that may affect them become increasingly important, as they will better equip and inform us for changes within enzootic systems or future epizootic events facilitating endeavours to preserve and maintain genetic and overall biodiversity.

6.4 References

1. Auteri GG, Knowles LL. Decimated little brown bats show potential for adaptive change. Scientific Reports. 2020 Feb 20;10(1):3023.
2. Berteaux D, Gallant D, Sacks BN, Statham MJ. Red foxes (*Vulpes vulpes*) at their expanding front in the Canadian Arctic have indigenous maternal ancestry. Polar Biology. 2015 Jun 1; 38(6):913–7.
3. Blejwas K, Beard L, Buchanan J, Lausen CL, Neubaum D, Tobin A, Weller TJ. COULD WHITE-NOSE SYNDROME MANIFEST DIFFERENTLY IN MYOTIS LUCIFUGUS IN WESTERN VERSUS EASTERN REGIONS OF NORTH AMERICA? A REVIEW OF FACTORS. Journal of Wildlife Diseases. 2023 Jun 1.
4. Davy CM, Donaldson ME, Bandouchova H, Breit AM, Dorville NA, Dzal YA, Kovacova V, Kunkel EL, Martínková N, Norquay KJ, Paterson JE. Transcriptional host–pathogen responses of Pseudogymnoascus destructans and three species of bats with white-nose syndrome. Virulence. 2020 Dec 31;11(1):781-94.
5. Davy CM, Donaldson ME, Rico Y, Lausen CL, Dogantzis K, Ritchie K, Willis CK, Burles DW, Jung TS, McBurney S, Park A. Prelude to a panzootic: Gene flow and immunogenetic variation in northern little brown myotis vulnerable to bat white-nose syndrome. Facets. 2017 Sep 12;2(2):690-714.
6. Donaldson ME, Davy CM, Willis CK, McBurney S, Park A, Kyle CJ. Profiling the immunome of little brown myotis provides a yardstick for measuring the genetic response to white-nose syndrome. Evolutionary applications. 2017 Dec;10(10):1076-90.
7. Donaldson ME, Rico Y, Hueffer K, Rando HM, Kukekova AV, Kyle CJ. Development of a genotype-bysequencing immunogenetic assay as exemplified by screening for variation in red fox with and without endemic rabies exposure. Ecology and evolution. 2018 Jan; 8(1):572–83. https://doi.org/10.1002/ece3.3583 PMID: 29321894
8. Drees KP, Lorch JM, Puechmaille SJ, Parise KL, Wibbelt G, Hoyt JR, Sun K, Jargalsaikhan A, Dalannast M, Palmer JM, Lindner DL. Phylogenetics of a fungal invasion: origins and widespread dispersal of white-nose syndrome. MBio. 2017 Dec 29;8(6):10-128.
9. Escobar LE, Carver S, Cross PC, Rossi L, Almberg ES, Yabsley MJ, Niedringhaus KD, Van Wick P, Dominguez-Villegas E, Gakuya F, Xie Y. Sarcoptic mange: An emerging panzootic in wildlife. Transboundary and Emerging Diseases. 2022 May;69(3):927-42.
10. Field KA, Johnson JS, Lilley TM, Reeder SM, Rogers EJ, Behr MJ, Reeder DM. The white-nose syndrome transcriptome: activation of anti-fungal host responses in wing tissue of hibernating little brown myotis. PLoS pathogens. 2015 Oct 1;11(10):e1005168.

11. Field KA, Sewall BJ, Prokkola JM, Turner GG, Gagnon MF, Lilley TM, Paul White J, Johnson JS, Hauer CL, Reeder DM. Effect of torpor on host transcriptomic responses to a fungal pathogen in hibernating bats. Molecular ecology. 2018 Sep;27(18):3727-43.
12. Fischer NM, Dool SE, Puechmaille SJ. Seasonal patterns of Pseudogymnoascus destructans germination indicate host–pathogen coevolution. Biology Letters. 2020 Jun 24;16(6):20200177.
13. Frank CL, Davis AD, Herzog C. The evolution of a bat population with white-nose syndrome (WNS) reveals a shift from an epizootic to an enzootic phase. Frontiers in Zoology. 2019 Dec;16:1-9.
14. Gignoux-Wolfsohn, S. A., Pinsky, M. L., Kerwin, K., Herzog, C., Hall, M., Bennett, A. B., ... & Maslo, B. (2021). Genomic signatures of selection in bats surviving white-nose syndrome. Molecular Ecology, 30(22), 5643-5657.
15. Goldsmith EW, Renshaw B, Clement CJ, Himschoot EA, Hundertmark KJ, Hueffer K. Population structure of two rabies hosts relative to the known distribution of rabies virus variants in Alaska. Molecular ecology. 2016 Feb; 25(3):675–88. https://doi.org/10.1111/mec.13509 PMID: 26661691
16. Goldsmith EW, Renshaw B, Clement CJ, Himschoot EA, Hundertmark KJ, Hueffer K. Population structure of two rabies hosts relative to the known distribution of rabies virus variants in Alaska. Moleculare cology. 2016 Feb; 25(3):675–88. https://doi.org/10.1111/mec.13509 PMID: 26661691
17. Huettmann F, Magnuson EE, Hueffer K. Ecological niche modeling of rabies in the changing Arctic of Alaska. Acta Veterinaria Scandinavica. 2017 Dec; 59(1):18. https://doi.org/10.1186/s13028-017-0285- 0 PMID: 28320440
18. Kuzmin IV, Hughes GJ, Botvinkin AD, Gribencha SG, Rupprecht CE. Arctic and Arctic-like rabies viruses: distribution, phylogeny and evolutionary history. Epidemiology & Infection. 2008 Apr; 136 (4):509–19.
19. Lilley TM, Prokkola JM, Blomberg AS, Paterson S, Johnson JS, Turner GG, Bartoničkа T, Bachorec E, Reeder DM, Field KA. Resistance is futile: RNA-sequencing reveals differing responses to bat fungal pathogen in Nearctic Myotis lucifugus and Palearctic Myotis myotis. Oecologia. 2019 Oct;191:295-309.
20. Lilley TM, Prokkola JM, Blomberg AS, Paterson S, Johnson JS, Turner GG, Bartoničkа T, Bachorec E, Reeder DM, Field KA. Resistance is futile: RNA-sequencing reveals differing responses to bat fungal pathogen in Nearctic Myotis lucifugus and Palearctic Myotis myotis. Oecologia. 2019 Oct;191:295-309.
21. Lilley TM, Wilson IW, Field KA, Reeder DM, Vodzak ME, Turner GG, Kurta A, Blomberg AS, Hoff S, Herzog CJ, Sewall BJ. Genome-wide changes in genetic diversity in a population of Myotis lucifugus affected by white-nose syndrome. G3: Genes, Genomes, Genetics. 2020 Jun 1;10(6):2007-20.

22. MacInnes CD, Smith SM, Tinline RR, Ayers NR, Bachmann P, Ball DG, et al. Elimination of rabies from red foxes in eastern Ontario. Journal of wildlife diseases. 2001 Jan; 37(1):119–32. https://doi.org/ 10.7589/0090-3558-37.1.119 PMID: 11272485 55.
23. Mallory ML, Gilchrist HG, Janssen M, Major HL, Merkel F, Provencher JF, Strøm H. Financial costs of conducting science in the Arctic: examples from seabird research. Arctic Science. 2018 Jun 4;4(4):624-33.
24. Mørk T, Prestrud P. Arctic rabies–a review. Acta Veterinaria Scandinavica. 2004 Mar; 45(1):1. 56.
25. Mörner T. Sarcoptic mange in Swedish wildlife. Revue scientifique et technique (International Office of Epizootics). 1992 Dec 1;11(4):1115-21.
26. Nadin-Davis SA, Sheen M, Wandeler AI. Recent emergence of the Arctic rabies virus lineage. Virus research. 2012 Jan 1; 163(1):352–62. https://doi.org/10.1016/j.virusres.2011.10.026 PMID: 22100340
27. Niedringhaus KD, Brown JD, Sweeley KM, Yabsley MJ. A review of sarcoptic mange in North American wildlife. International Journal for Parasitology: Parasites and Wildlife. 2019 Aug 1;9:285-97.
28. NugentG,BuddleBM,KnowlesG.EpidemiologyandcontrolofMycobacteriumbovisinfection inbrushtail possums (Trichosurus vulpecula), the primary wildlife host of bovine tuberculosis in New Zealand. N ZVetJ.2015;63:28–41. https://doi.org/10.1080/00480169.2014.963791 PMID: 25290902
29. Rosatte RC, Power MJ, Donovan D, Davies JC, Allan M, Bachmann P, et al. Elimination of arctic variant rabies in red foxes, metropolitan Toronto. Emerging Infectious Diseases. 2007 Jan; 13(1):25. https://doi.org/10.3201/eid1301.060622 PMID: 17370512
30. Shi Y, Bouska KL, McKinney GJ, Dokai W, Bartels A, McPhee MV, Larson WA. Gene flow influences the genomic architecture of local adaptation in six riverine fish species. Molecular Ecology. 2023 Apr;32(7):1549-66.
31. SingerA,KauhalaK,HolmalaK,SmithGC.RabiesinNortheasternEurope—thethreatfrominvasive raccoon dogs. J Wildl Dis. 2009; 45:1121–1137. https://doi.org/10.7589/0090-3558-45.4.1121 PMID: 19901385
32. Trivedi CB, Keuschnig C, Larose C, Rissi DV, Mourot R, Bradley JA, Winkel M, Benning LG. DNA/RNA preservation in glacial snow and ice samples. Frontiers in Microbiology. 2022 May 23;13:894893.
33. Uslu VV, Petretich M, Ruf S, Langenfeld K, Fonseca NA, Marioni JC, Spitz F. Long-range enhancers regulating Myc expression are required for normal facial morphogenesis. Nature genetics. 2014 Jul;46(7):753-8.

Appendices

Appendix I: Supplemental materials

Chapter 2

Fig S2.1. Weak signature of neutral genetic structure among red fox populations within Alaska (not including Ontario). Analyses of the 123 putatively neutral SNPs after filtering for a minor allele frequency threshold = 2% and pruning for linkage disequilibrium a) Principle component analysis where 1 = Central, 2 = Seward Peninsula, 3 = Southwest b) Power analysis results c) STRUCTURE analyses of K = 2 and K = 3 where individuals are represented by each bar along the x-axis and assignment to clusters is represented by the y-axis and the different colours; CN = Central Alaska; SP = Seward Peninsula; SW = Southwest Alaska d) DAPC of the inferred clustering of the data e) DAPC organized by sampling location (1 = Central Alaska; 2 = Seward Peninsula; 3 = Southwest Alaska).
https://doi.org/10.1371/journal.pone.0249176.s001

Fig S2.2. Schematic of the outlier SNPs before linkage disequilibrium pruning. a) among red fox from Alaska and Ontario; 131 SNP outliers were identified from the sub-dataset that included Ontario red fox. Only PCAdapt and Arlequin identified outliers. b) within Alaska (not including Ontario); 221 SNP outliers were identified from the sub-dataset that did not include Ontario red fox. Only PCAdapt, Arlequin, and OutFLANK identified outliers.
https://doi.org/10.1371/journal.pone.0249176.s002

Fig S2.3. p-value of iHS detects weak signals of selection within three populations of red fox in Alaska. Within population measurements of selective sweeps based on the p-value of the iHS statistic within the Central, Seward Peninsula and Southwest Alaskan red fox populations. Solid grey bars indicate the identified candidate region of iHS outliers for each population.
https://doi.org/10.1371/journal.pone.0249176.s003

Fig S2.4. Assessment of selective sweeps between populations of Alaskan red fox using XP-EHH. Between population measurements of selective sweeps based on the p-value of the calculated XP-EHH statistic between the Central, Seward Peninsula and Southwest Alaskan red fox populations.
https://doi.org/10.1371/journal.pone.0249176.s004

Fig S2.5. Genome wide F_{ST} estimates between 3 populations of red fox in Alaska. Pairwise Weir and Cockerham F_{ST} values between Central, Seward Peninsula, and Southwest Alaskan red fox populations. Identified outliers are highlighted in red and are those represented in the only red fox from Alaska subset in Table S2.5.
https://doi.org/10.1371/journal.pone.0249176.s005

Table S2.1. Red fox sample information. Sample identifiers, location, corresponding arctic rabies variant to the area, and accession numbers. https://doi.org/10.1371/journal.pone.0249176.s006

Table S2.2. 116 genes probe-bait targets enriched for. Describes in reference to the dog genome (per gene): The transcript and gene ID, position in the genome (chromosome, start/stop base pair), number of exons, and the BLASTp hit description. https://doi.org/10.1371/journal.pone.0249176.s007

Table S2.3. GATK filtering results for the 125 red fox samples. Describes (per sample): The number of raw reads, reads passing GATK filters and those reads not passing the GATK filters due to mapping quality, secondary alignments, and duplicate reads. https://doi.org/10.1371/journal.pone.0249176.s008

Table S2.4. Filtered off-target SNP sub-datasets. Describes the position and average coverage for each SNP retained in the filtered sub-dataset among red fox populations within Alaska and Ontario, and the filtered sub-dataset among red fox populations exclusively within Alaska. Filtering parameters were a minor allele frequency threshold = 2% and pruning for linkage disequilibrium. https://doi.org/10.1371/journal.pone.0249176.s009

Table S2.5. Identified outliers before and after disequilibrium pruning among red fox populations across North America. Describes (per SNP): Location, gene association, and predicted gene function in reference to the dog genome. SNPs are further identified to which sub-dataset they belong; red fox from Alaska and Ontario or red fox only from Alaska: i) identified outlier ii) identified outlier retained after filtering iii) missense SNP with potential to alter protein function and finally if the SNP was identified as an outlier by multiple tests. All BLASTp predicted functions are based upon Canis lupus familiaris unless otherwise specified. https://doi.org/10.1371/journal.pone.0249176.s010

Table S2.6. pN/pS ratios for three populations of red fox from Alaska. The ratio of nonsynonymous substitution per nonsynonymous site to synonymous substitutions per synonymous site for 85 genes and three populations of red fox. Those pN/pS ratios >1, suggestive of positive selection, have been bolded. In order for the calculation to be performed, and a pN/pS ratio determined, there must have been at least 1 synonymous and 1 nonsynonymous polymorphism per gene. Per gene, the number of nonsynonymous/synonymous polymorphisms were determined with SnpEff (Cingolani et al. 2012) and the potential nonsynonymous/synonymous sites was estimated from the coding sequence of each gene using DnaSP v6 (Rozas et al. 2017). https://doi.org/10.1371/journal.pone.0249176.s011

Updated Table S2.6 that includes chi-square test results can be viewed at: https://trentu-my.sharepoint.com/:x:/g/personal/tristanbaecklund_trentu_ca/EWeUSxvOCT9Ei92tiDDBUyIB0xbBeMmFXM7zvNsceIomDw?e=IgmbcO

File S2.1. Supplementary methods. Additional details for methods. https://doi.org/10.1371/journal.pone.0249176.s012

Chapter 3

Fig S3.1. Principal component analyses of progressively filtered off-target SNP datasets. Clustering of the 96 arctic fox samples based on on-target SNPs throughout filtering steps. a) PCA of all off-target SNPs after filtering for MAF, missing-data, and biallelic SNPs (n = 6432) and b) PCA of all identified off-target SNPs after analysis with Variant Effect Predictor and prior to linkage pruning (n = 283). Arviat = black square; Chevak = yellow circle; Hooper Bay = blue triangle; Victoria Island = green diamond.
https://doi.org/10.1371/journal.pone.0258975.s001

Fig S3.2. DAPC of final off-target SNP sub-dataset identifies six clusters. Discriminant analysis of principal components on the final on-target SNP sub-dataset (n = 29). a) the clustering of the samples into six inferred clusters, and b) the clustering of the samples into the same six inferred clusters as in a), but individuals are identified based on the geographical region from which the sample originated; cluster 1 (black square) = Arviat; cluster 2 (yellow circle) = Chevak; cluster 3 (blue triangle)–Hooper Bay; cluster 4 (green diamond)–Victoria Island.
https://doi.org/10.1371/journal.pone.0258975.s002

Fig S3.3. Principal component analyses of progressively filtered on-target SNP datasets. Clustering of the 96 arctic fox samples based on on-target SNPs throughout filtering steps. a) PCA of all on-target SNPs after filtering for MAF, missing-data, and biallelic SNPs (n = 2277) and b) PCA of all identified on-target F_{ST} outlier SNPs prior to linkage pruning (n = 107). Arviat = black square; Chevak = yellow circle; Hooper Bay = blue triangle; Victoria Island = green diamond.
https://doi.org/10.1371/journal.pone.0258975.s003

Fig S3.4. DAPC of final on-target SNP sub-dataset identifies two clusters. Discriminant analysis of principal components on the final on-target SNP sub-dataset (n = 22). a) the clustering of the samples into two inferred clusters and b) the clustering of samples into the same two inferred clusters as in a), but individuals are identified based on the geographic region from which the sample originated; cluster 1 = Arviat and Victoria Island samples; cluster 2 = Southwestern Alaska samples. https://doi.org/10.1371/journal.pone.0258975.s004

Fig S3.5. Comparison of identified F_{ST} outlier SNPs using different detection methods. The proportion of F_{ST} outliers identified by each program for a) all identified F_{ST} outlier SNPs from the on-target data and b) the identified F_{ST} outlier SNPs that were retained in the final on-target sub-dataset.
https://doi.org/10.1371/journal.pone.0258975.s005

Fig S3.6. Power analysis results for the final off-target SNP sub-dataset. Assessment of the power for the final off-target SNP sub-dataset (n = 29) with an assumed effective population size of 5000. Chi-squared test results are shown as blue squares, and the Fisher exact test results are shown as grey diamonds. https://doi.org/10.1371/journal.pone.0258975.s006

Fig S3.7. Genome-wide F_{ST} estimates between 3 populations of arctic fox in North America. Pairwise Weir and Cockerham F_{ST} values between Southwestern Alaska (Chevak and Hooper Bay), Arviat, and Victoria Island arctic fox populations. Identified outliers are highlighted in red and are those found in Table S2.5. https://doi.org/10.1371/journal.pone.0258975.s007

Fig S3.8. Principal component analysis of the 4 outlier SNPs after linkage-disequilibrium pruning from the off-target dataset. Based on the lack of outliers identified, and the lack of genetic clustering evident by their visualization, we conclude there is no genetic structure among the arctic fox populations based on these off-target data. It is important to note however, that due to the small size of the dataset, and the nature of biallelic markers that conclusions drawn from these data should be done with caution. Arviat (1—Black squares); Chevak (2—Yellow circle); Hooper Bay (3—Light Blue Triangle); Victoria Island (4—Green diamond). https://doi.org/10.1371/journal.pone.0258975.s008

Fig S3.9. Principal component analysis of the 29 off-target SNPs passing filtering parameters and pruned for linkage disequilibrium. Based on the prominent overlap of all 4 clusters, these data do not suggest genetic structuring. These data were then further investigated with STRUCTURE and DAPC analyses (Fig 2.2). Arviat (AR–Black circles); Chevak (CV–Yellow circles); Hooper Bay (HB–Light blue circles); Victoria Island (VI–Green circles). https://doi.org/10.1371/journal.pone.0258975.s009

Table S3.1. Arctic fox sample information. Sample identifiers, year of collection, area of sampling and corresponding arctic rabies variant circulating area of collection. https://doi.org/10.1371/journal.pone.0258975.s010

Table S3.2. 116 immunogenetic probe-baited targets enriched for. Describes in reference to the dog genome (per gene): the transcript and gene ID, position in the genome (chromosome, start/stop base pair), number of exons, and the BLASTp hit description. https://doi.org/10.1371/journal.pone.0258975.s011

Table S3.3. GATK filtering results for the 96 arctic fox samples. Describes (per sample): the number of raw reads, reads passing GATK filters and those reads not passing the GATK filters due to mapping quality, secondary alignments, and duplicate reads. https://doi.org/10.1371/journal.pone.0258975.s012

Table S3.4. Filtered off-target SNP sub-dataset. Describes the position and average coverage for each SNP retained in the final filtered sub-dataset. Filtering parameters were a minor allele frequency threshold = 2% and pruning for linkage disequilibrium. https://doi.org/10.1371/journal.pone.0258975.s013

Table S3.5. Identified F_{ST} outliers before and after disequilibrium pruning among arctic fox populations across North America. Describes (per SNP): location, gene association, and predicted gene function in reference to the dog genome. The program identifying the SNP as an F_{ST} outlier is also presented along with an indication of those missense SNPs with the potential to alter protein function. All BLASTp predicted functions are based upon Canis lupus familiaris unless otherwise specified. https://doi.org/10.1371/journal.pone.0258975.s014

Table S3.6. Pairwise F_{ST} 97.5% upper and lower confidence intervals. Outlines the pairwise F_{ST} intervals between the three sampled regions. Interval pairs that were significantly different from zero are bolded.
https://doi.org/10.1371/journal.pone.0258975.s015

Table S3.7. Comparison of genes under selection based on pN/pS ratios for each sampling region. For each gene highlights the ratio of non-synonymous substitutions per non-synonymous site (pN) to synonymous substitutions per synonymous site (pS) across the three sampled regions, as well as a test of significance using the Chi-squared p-values. Bolded ratios indicated those pN/pS ratios greater than or equal to 1 suggestive of directional selection and those pN/pS ratios that have statistically significant p values. Polymorphic sites were determined using SnpEff and potential nonsynonymous/synonymous sites were determined using the coding sequence for each gene as input for DnaSP v6. (Rozas et al. 2017 [68]). https://doi.org/10.1371/journal.pone.0258975.s016

File S3.1. Supplementary methods. Additional methods pertaining to linkage disequilibrium pruning and the F_{ST} outlier testing program parameters selected. https://doi.org/10.1371/journal.pone.0258975.s017

Chapter 4

These supplemental materials can be viewed at https://trentu-my.sharepoint.com/:f:/g/personal/tristanbaecklund_trentu_ca/ErF3O__YViRBoHjUgirxjK8BLPCnppd8QDUI5cGphi4F7g?e=6zSIxM.

Fig S4.1. Hierarchical analyses of two *M. leibii* genetic clusters using ten microsatellite loci. Bar plots demonstrate a lack of substructure within two putative populations, where each partitioned vertical bar represents an individual's proportional membership to the inferred genetic cluster. Inserts a) and c) show estimated population assignments within the North Carolina (NC) and West Virgina (WV) cluster. Inserts b) and d) show estimated population assignments within the New York (NY), Vermont (VT), New Hampshire (NH) and Ontario (ON) cluster. a) and b) represent results using LOCPRIOR = 0, where c) and d) use LOCPRIOR = 1 for the respective datasets. See Fig S4.11 for metrics associated with modeled K values using the Evanno method.

Fig S4.2. Hierarchical analysis of immunogenetic data within the identified northern *M. leibii* genetic cluster. Further investigation of 105 immunogenetic LDP SNPs sought to identify potential substructure within the NE-USA and ON cluster. Bar plots demonstrate substructure within the New York (NY), Vermont (VT), New Hampshire (NH) and Ontario (ON) cluster, where a) represent results using LOCPRIOR = 0, and b) uses LOCPRIOR = 1. Each partitioned vertical bar represents an individual's proportional membership to the inferred genetic cluster. K = 3 was identified as the most likely model based on the Evanno method (Fig S4.12).

Fig S4.3. Tests for HWE highlight microsatellite loci not in equilibrium among 8 populations. Utilizing both the Monte Carlo permutations exact test (panel a) and the X^2 approach (panel b) demonstrates loci that fall within HWE (light blue colour) where H_0 was supported, and loci not in HWE (coloured pink) where H_1 was supported. Microsatellite markers are listed on the y-axis and populations across the x-axis. AR = Arkansas; NC/WV = North Carolina & West Virginia; NY/NH/VT/ON = New York, New Hampshire, Vermont, &Ontario). As no locus was out of HWE across all populations, all loci were retained for further analyses.

Fig S4.4. Variant effect predictor results for entire dataset of SNPs identified within the Arkansas *M. leibii* population. Pie charts summarizing the variant annotations for 2,999 SNPs identified within the Arkansas population. These data show the large proportion of mutations within non-coding and potential regulatory regions than those within coding regions. Chart a) shows the proportion of the overall consequences of each substitution, b) shows the coding consequences of each substitution, and c) shows the proportion of the impact of each consequence. Note that there are more consequence impacts relative to the number of processed variants, as some substitutions have multiple potential implications.

Fig S4.5. Variant effect predictor results for entire dataset of SNPs identified within the eastern Ontario *M. leibii* population. Pie charts summarizing the variant annotations for 2,520 SNPs identified within the Arkansas population. These data show the large proportion of mutations within non-coding and potential regulatory regions than those within coding regions. Chart a) shows the proportion of the overall consequences of each substitution, b) shows the coding consequences of each substitution, and c) shows the proportion of the impact of each consequence. Note that there are more consequence impacts relative to the number of processed variants, as some substitutions have multiple potential implications.

Fig S4.6. Simulated allele frequency projections and associated estimated selection coefficient under three models show rapid genetic shifts in bats from Arkansas. For the 28 F_{ST} outlier SNPs where simulations (Table S4.7) suggest allele frequency shifts were not caused by genetic drift, we propagated allele frequencies forward 15 generations using pre-WNS occurrence data and estimated the coefficient of selection under three models in context of the post-WNS occurrence

data. Individual points reflect a simulated allele frequency where lines reflect the average coefficient of selection under each model depicted by the three colours. Each figure demonstrates rapid shifts in allele frequencies at the seven identified missense SNPs within the subset. a) C/T substitution at position 14255 on scaffold GL430751 b) G/A substitution at position 15496 on scaffold GL430751 c) A/C substitution at position 85062 on scaffold GL430751 d) G/T substitution at position 90028 on scaffold GL430751 e) T/C substitution at position 85032 on scaffold GL430751 f) A/C substitution at position 85037 on scaffold GL430751 and g) G/A substitution at position 91360 on scaffold GL430751.

Fig S4.7. Simulated allele frequency projections and associated estimated selection coefficient under three models show rapid genetic shifts in bats from eastern Ontario. For the 29 F_{ST} outlier SNPs where simulations (Table S4.7) suggest allele frequency shifts were not caused by genetic drift, we propagated allele frequencies forward 15 generations using pre-WNS occurrence data and estimated the coefficient of selection under three models in context of the post-WNS occurrence data. Individual points reflect a simulated allele frequency where lines reflect the average coefficient of selection under each model depicted by the three colours. Each figure demonstrates rapid shifts in allele frequencies at the six identified missense SNPs within the subset. a) A/G substitution at position 6,792 on scaffold GL432468 b) T/C substitution at position 83,865 on scaffold GL430787 c) A/G substitution at position 153,470 on scaffold GL430379 d) T/A substitution at position 5,596,948 on scaffold GL429815 e) A/C substitution at position 5,597,065 on scaffold GL429815 and f) G/A substitution at position 5,615,774 on scaffold GL429815.

Fig S4.8. Identified K cluster metrics for Fig 4.2 using the Evanno method. Metrics of assigned K values used to infer clustering from pophelper using the Evanno method. a) plot of the mean likelihood of K clusters and b) delta K plot for 10 microsatellite loci between *Myotis leibii* populations.

Fig S4.9. Identified K cluster metrics for Fig 4.3 using the Evanno method. Metrics of assigned K values used to infer clustering from pophelper using the Evanno method. a) plot of the mean likelihood of K clusters and b) delta K plot for 105 immunogenetic SNPs between *Myotis leibii* populations, where LOCPRIOR=0. c) plot of the mean likelihood of K clusters and d) delta K plot for the same 105 immunogenetic SNPs, but LOCPRIOR=1.

Fig S4.10. Identified K cluster metrics for Fig 4.4 using the Evanno method. Metrics of assigned K values used to infer clustering from pophelper using the Evanno method. a) plot of the mean likelihood of K clusters and b) delta K plot for 23 immunogenetic outlier SNPs between two *Myotis leibii* populations, where LOCPRIOR=0. c) plot of the mean likelihood of K clusters and d) delta K plot for the same 23 immunogenetic outlier SNPs, but LOCPRIOR=1.

Fig S4.11. Identified K cluster metrics for Fig S4.1 using the Evanno method. Metrics of assigned K values used to infer clustering from pophelper using the Evanno method. Plots a,c,e,g) represent the plot of the mean likelihood of K clusters and b,d,f,h) show delta K plot for 10 microsatellite loci between two hierarchical analyses of *Myotis leibii* populations. a-d) substructure of North Carolina and West Virginia and e-h) New York, Vermont, New Hampshire, and Ontario. a,b,e,f) correspond to LOCPRIOR = 0, and c,d,g,h) correspond to LOCPRIOR = 1.

Fig S4.12. Identified K cluster metrics for Fig S4.2 using the Evanno method. Metrics of assigned K values used to infer clustering from pophelper using the Evanno method. a) plot of the mean likelihood of K clusters and b) delta K plot for 105 immunogenetic SNPs within a northeastern United States and Ontario genetic cluster of *Myotis leibii*, where LOCPRIOR=0. c) plot of the mean likelihood of K clusters and d) delta K plot for the same 105 immunogenetic outlier SNPs, but LOCPRIOR=1.

Table S4.1. Sample information for *M. leibii* samples used in microsatellite and targeted immunogenetic assay. Table contains relevant sample identifiers, year of collection, temporal designation of WNS occurrence, and sample collection information.

Table S4.2. Microsatellite loci multiplexes and important considerations for amplification. Table includes primer specifications, number of PCR cycles, and annealing temperatures used to amplify the ten microsatellite markers.

Table S4.3. Genes of interest for targeted immunogenetic-assay. Table highlights identifiers along with descriptions of each targeted gene and their location in the *M. lucifugus* reference genome (Myoluc2.0).

Table S4.4. Microsatellite genotypes at ten loci for 147 *M. leibii* samples. Table denotes complete genotype, approximate spatial population, year of collection, and overlap with the immunogenetic assay for each sample.

Table S4.5. Immunogenetically relevant SNPs identified from three analyses demonstrating genetic clustering consistent with neutral microsatellite markers. Table includes identified immunogenetically relevant SNPs filtered from the initial dataset of 4,373 loci (N = 160 *M. leibii)*, and those identified from a subset of 3,711 SNPs obtained from 71 individuals from Arkansas and eastern Ontario combined (where pre- and post-WNS occurrence samples were obtained). The table identifies SNP locations, base-pair change, gene association, and change in amino acid for of each biallelic SNP retained in three subsets of SNPs. The first two subsets were identified from the initial 4,373 immunogenetically relevant loci and consist of 1) 105 linkage disequilibrium pruned SNPs (visualized in Fig 4.3) and 2) the 16 F_{ST} outlier SNPs differentiating AR vs. NC/VA vs. NY/NH/VT/ON (linkage disequilibrium pruned from a total of 114 identified outliers). The third subset was comprised of 23 identified F_{ST} outliers (differentiating pre- and post-WNS occurrence samples) selected from 3,711 and samples originating from Arkansas and eastern Ontario

samples (visualized in Fig 4.4). Note that blank 'Amino Acid' cells reflect SNPs whose location is upstream of the protein coding gene.

Table S4.6. Identified F_{ST} outlier immunogenetic SNPs presumably under selection associated with pre- and post-WNS occurrence within Arkansas and eastern Ontario. For the 28 and 29 F_{ST} outlier SNPs identified within Arkansas and eastern Ontario respectively, this table indicates the position of each SNP, base pair substitution, absolute shift in allele frequency of the ancestral allele between pre- and post-WNS occurrence samples, associated gene, and highlights where any synonymous, missense, or nonsense changes have occurred (denoted by *). Note that biallelic SNPs upstream/in regulatory regions of these genes would not affect the composition of the resulting protein and as such have been left blank.

Table S4.7. Simulated allele frequency (AF) projections and associated estimated selection coefficient under three models for 3,711 immunogenetic SNPs showing rapid shifts in bats from Arkansas and eastern Ontario. Table denotes simulations for each of the 3,711 immunogenetic SNPs: i) their location, ii) designation as an identified F_{ST} outlier, iii) presumed cause of shift in allele frequency in context of pre- to post-WNS occurrence (drift or outlier - presumable selection), and iv) the estimated coefficient of selection under three models for SNPs with allele shifts unlikely to be caused by drift and identified as F_{ST} outliers.

Chapter 5

These supplemental materials can be viewed at https://trentu-my.sharepoint.com/:f:/g/personal/tristanbaecklund_trentu_ca/EnK0Jq6IMlVJup2230yihUUBi8xXO-3aRJAz7zC2BvPNIQ?e=oopJFS.

Fig S5.1. Lack of population genetic structure at the eastern extent of the *Myotis lucifugus* range from 11 microsatellite loci and 235 samples. Across both STRUCTURE and DAPC analyses, no genetic structuring is apparent across the entire eastern extent of the *M. lucifugus* range, where STRUCTURE identifies a unique genetic cluster consisting of samples from the Yukon. These analyses, simultaneously, do not suggest shifts in genetic diversity associated with the introduction and spread of white-nose syndrome. Insert a) demonstrates a bar plot of estimated population assignments using STRUCTURE for K = 2 where each partitioned vertical bar represents an individual's proportional membership to inferred populations, noting K = 2 clustering was most supported by the Evanno method (Evanno, Regnaut, & Goudet, 2005; Fig S5.10). Insert c) illustrates a discriminant analysis of principal components, where samples are colour-coded based on geographic origins. DAPC analyses indicated K = 4 clustering as optimal, despite significant overlap among clusters. YT = Yukon; MB = Manitoba; W-ON = western Ontario; E-ON = eastern Ontario; ATL-CAN = Atlantic Canada.

Fig S5.2. Lack of temporal genetic substructure within *Myotis lucifugus* from eastern Ontario based on immunogenetic loci. Visualization of 2,016 immunogenetic SNPs from bats sampled throughout the introduction of WNS in eastern Ontario show no apparent genetic structure associated with three sampled time points throughout the introduction and spread of WNS. Insert a) shows bar plots estimating population assignments using STRUCTURE for K = 2-3, noting K = 2 most supported by the Evanno method (Evanno, Regnaut, & Goudet, 2005; Fig S5.11). Each partitioned vertical bar represents an individual's proportional membership to the inferred populations, visualized using LOCPRIOR = 0. A discriminant analysis of principal components is demonstrated in b) where samples are colour-coded based on temporal sampling pre-WNS (PRE), deaths immediately following (EMOR), and surviving bats several years post-WNS occurrence (SURV).

Fig S5.3. Simulated allele frequency projections and associated estimated selection coefficient under three models show rapid genetic shifts in bats from Atlantic Canada. For 29 F_{ST} outlier SNPs, simulations (Table S5.7) suggest allele frequency shifts were not caused by genetic drift. From these outliers, allele frequencies were propagated forward 15 generations using pre-WNS occurrence data to estimate the coefficient of selection under three models in the context of early mortality allele frequency data. Individual points reflect a simulated allele frequency where lines reflect the average coefficient of selection under each model depicted by three colours. Each figure demonstrates rapid shifts in allele frequencies at eleven identified missense SNPs within the outlier subset. a) C/T substitution at position 5676349 on scaffold GL429830 b) G/A substitution at position 15261203 on scaffold GL429775 c) G/T substitution at position 8477578 on scaffold GL429781 d) C/A substitution at position 8477581 on scaffold GL429781 e) A/C substitution at position 8477596 on scaffold GL429781 f) G/A substitution at position 8477622 on scaffold GL429781 g) A/G substitution at position 8483624 on scaffold GL429781 h) T/C substitution at position 8488009 on scaffold GL429781 i) T/C substitution at position 8488012 on scaffold GL429781 j) C/T substitution at position 8490216 on scaffold GL429781 and k) T/C substitution at position 8490234 on scaffold GL429781.

Fig S5.4. Simulated allele frequency projections and associated estimated selection coefficient under three models show rapid genetic shifts in bats from eastern Ontario between bats sampled pre-WNS occurrence and early mortalities. For 16 F_{ST} outlier SNPs, simulations (Table S5.7) suggest allele frequency shifts were not caused by genetic drift. From these outliers, allele frequencies were propagated forward 15 generations using pre-WNS occurrence data to estimate the coefficient of selection under three models in the context of early mortality allele frequency data. Individual points reflect a simulated allele frequency where lines reflect the average coefficient of selection under each model depicted by three colours. Each figure demonstrates rapid shifts in allele frequencies at the eight identified missense SNPs within the outlier subset. a) G/C substitution at position 8228131 on scaffold GL429777 b) A/G substitution at position 8228186 on scaffold GL429777 c) G/C substitution at position 8228187 on

scaffold GL429777 d) T/C substitution at position 8228214 on scaffold GL429777 e) T/G substitution at position 8228216 on scaffold GL429777 f) A/C substitution at position 710342 on scaffold GL430150 g) G/A substitution at position 118693 on scaffold GL430221 and h) T/C substitution at position 118846 on scaffold GL430221.

Fig S5.5. Simulated allele frequency projections and associated estimated selection coefficient under three models show rapid genetic shifts in bats from eastern Ontario between bats sampled pre-WNS occurrence and survivors several years after introduction. For 37 F_{ST} outlier SNPs, simulations (Table S5.7) suggest allele frequency shifts were not caused by genetic drift. From these outliers, allele frequencies were propagated forward 15 generations using pre-WNS occurrence data to estimate the coefficient of selection under three models in the context of data from surviving bats post-WNS occurrence. Individual points reflect a simulated allele frequency where lines reflect the average coefficient of selection under each model depicted by three colours. Each figure demonstrates rapid shifts in allele frequencies at the seven identified missense SNPs within the outlier subset. a) A/G substitution at position 2981860 on scaffold GL429768 b) G/A substitution at position 2984220 on scaffold GL429768 c) C/A substitution at position 2984253 on scaffold GL429768 d) G/A substitution at position 4629 on scaffold GL430078 e) G/A substitution at position 118693 on scaffold GL430221 f) T/C substitution at position 118846 on scaffold GL430221 and g) C/T substitution at position 143020 on scaffold GL430306.

Fig S5.6. Variant Effect Predictor results for entire dataset of SNPs identified within the Atlantic Canada *M. lucifugus* population. Pie charts summarizing the variant annotations for 1,393 SNPs identified within the Atlantic Canadian population. These data show the large proportion of mutations within non-coding and potential regulatory regions than those within coding regions. Chart a) shows the proportion of the overall consequences of each substitution, b) shows the coding consequences of each substitution, and c) shows the proportion of the impact of each consequence. Note that there are more consequence impacts relative to the number of processed variants, as some substitutions have multiple potential implications.

Fig S5.7. Variant Effect Predictor results for entire dataset of SNPs identified within the eastern Ontario *M. lucifugus* population. Pie charts summarizing the variant annotations for 2,016 SNPs identified within the eastern Ontario population. These data show the large proportion of mutations within non-coding and potential regulatory regions than those within coding regions. Chart a) shows the proportion of the overall consequences of each substitution, b) shows the coding consequences of each substitution, and c) shows the proportion of the impact of each consequence. Note that there are more consequence impacts relative to the number of processed variants, as some substitutions have multiple potential implications.

Fig S5.8. Identified K cluster metrics for Fig 5.3 using the Evanno method. Metrics of assigned K values used to infer clustering from pophelper using the

Evanno method. a) plot of the mean likelihood of K clusters and b) delta K plot for 29 immunogenetic outlier SNPs between PRE and EMOR Atlantic Canadian *Myotis lucifugus*.

Fig S5.9. Identified K cluster metrics for Fig 5.4 using the Evanno method. Metrics of assigned K values used to infer clustering from pophelper using the Evanno method. a) plot of the mean likelihood of K clusters and b) delta K plot for 16 immunogenetic outlier SNPs between PRE and EMOR eastern Ontario Myotis lucifugus. c) plot of the mean likelihood of K clusters and d) delta K plot for 37 immunogenetic outlier SNPs between PRE and SURV eastern Ontario *Myotis lucifugus*.

Fig S5.10. Identified K cluster metrics for Fig S5.1 using the Evanno method. Metrics of assigned K values used to infer clustering from pophelper using the Evanno method. a) plot of the mean likelihood of K clusters and b) delta K plot for 11 microsatellite loci across Canadian *Myotis lucifugus*.

Fig S5.11. Identified K cluster metrics for Fig S5.2 using the Evanno method. Metrics of assigned K values used to infer clustering from pophelper using the Evanno method. a) plot of the mean likelihood of K clusters and b) delta K plot for 2, 016 immunogenetic SNPs between eastern Ontario *Myotis lucifugus*.

Table S5.1. Sample information for *Myotis lucifugus* samples used in microsatellite and targeted immunogenetic assay. The table contains relevant sample identifiers, specific analyses samples used for, temporal designation of WNS occurrence, year of collection, and sample collection information. NCBI short-read archive accessions are available for previously sequenced individuals.

Table S5.2. Microsatellite genotypes at 11 loci for 235 *Myotis lucifugus* samples. Table includes sample ID for individuals retained in microsatellite analyses, and corresponding genotype at each of the eleven loci.

Table S5.3. Amplicon intervals for targeted neutral SNPs, microsatellite markers, and immunogenetic regions of interest. The table highlights interval location in *M. lucifugus* reference genome (Myoluc2.0), relative size of generated amplicons, and dataset amplicon was associated with. Note, that some intervals contained data for both neutral SNPs and microsatellite markers.

Table S5.4. Immunogenetically relevant SNP subsets identified from sampled *Myotis lucifugus*. Table denotes location, base substitution, associated gene, and amino acid change (where applicable) for identified F_{ST} outliers from each comparison, and 3,097 SNPs identified from the complete dataset of all *M. lucifugus* samples assessed herein.

Table S5.5. Simulated allele frequency (AF) projections and associated estimated selection coefficient under three models for immunogenetic SNPs showing rapid shifts in bats from Atlantic Canada and eastern Ontario. Allele frequency projection simulation and selection coefficients modelling results for each immunogenetic SNP identified from three temporal comparisons of *M. lucifugus* sampled before and after the initial spread of *Pd* (E-ON PRE vs. EMOR; E-

ON PRE vs. SURV; ATL-CAN PRE vs. EMOR). Table denotes for each SNP in respective comparisons: i) location, ii) designation as an identified F_{ST} outlier, iii) presumed cause of shift in allele frequency in context of pre- to post-WNS occurrence (simulation status; drift or outlier - presumable selection), and iv) estimated coefficient of selection under three models for SNPs with allele shifts unlikely to be caused by drift, identified as F_{ST} outliers, and resulted in a missense mutation.

Table S5.6. SNPs within gene subsets demonstrating frequency shifts, in the context of pre-WNS allele frequencies, between early mortality and surviving *Myotis lucifugus* in eastern Ontario. This table demonstrates data for SNPs where allele frequency shifts were unlikely to have been caused by genetic drift, but resulting shifts were only observed between one comparison. These data reflect potential candidate genes demonstrating frequency shifts that have the potential to elucidate genes influencing survival among eastern Ontario *M. lucifugus*. There were no directly overlapping SNPs that demonstrated a shift between EMOR and SURV, as such, these data are organized into subsets of SNPs based on gene associations, those identified only in one comparison, and genes where SNPs were identified in both; EMOR Only, SURV Only, and EMOR/SURV Overlapping (SNPs were identified in these genes for both comparisons), respectively. For each SNP, the table denotes: i) gene subset, ii) position, iii) allele and frequency for all three sampled periods, iv) genetic drift simulation status, v) absolute frequency shift from observed pre-WNS frequencies, and vi) overall conclusions about the observed frequency shifts.

Appendix II – Copyright Information

Chapter 2

RESEARCH ARTICLE

The role of a mechanistic host in maintaining arctic rabies variant distributions: Assessment of functional genetic diversity in Alaskan red fox (*Vulpes vulpes*)

Tristan M. Baecklund[1☯*], Jaycee Morrison[2☯], Michael E. Donaldson[1‡], Karsten Hueffer[3‡], Christopher J. Kyle[1,4,5‡]

1 Environmental and Life Sciences Graduate Program, Trent University, Peterborough, Ontario, Canada, **2** Forensic Science Undergraduate Program, Trent University, Peterborough, Ontario, Canada, **3** Department of Veterinary Medicine, University of Alaska Fairbanks, Fairbanks, Alaska, United States of America, **4** Forensic Science Department, Trent University, Peterborough, Ontario, Canada, **5** Natural Resources DNA Profiling & Forensic Centre, Trent University, Peterborough, Ontario, Canada

☯ These authors contributed equally to this work.
‡ These authors also contributed equally to this work.
* tristanbaecklund@trentu.ca

OPEN ACCESS

Citation: Baecklund TM, Morrison J, Donaldson ME, Hueffer K, Kyle CJ (2021) The role of a mechanistic host in maintaining arctic rabies variant distributions: Assessment of functional genetic diversity in Alaskan red fox (*Vulpes vulpes*). PLoS ONE 16(4): e0249176. https://doi.org/10.1371/journal.pone.0249176

Editor: Hoh Boon-Peng, UCSI University, MALAYSIA

Received: May 1, 2020

Accepted: March 12, 2021

Published: April 8, 2021

Data Availability Statement: Raw sequence data obtained from high-throughput sequencing are available from the NCBI Sequence Read Archive (accession number PRJNA592970).

Funding: This research was funded by a Natural Sciences and Engineering Research Council Discovery Grant to CJK (PGPIN2016- 05373). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. This research was enabled in part

Abstract

Populations are exposed to different types and strains of pathogens across heterogeneous landscapes, where local interactions between host and pathogen may present reciprocal selective forces leading to correlated patterns of spatial genetic structure. Understanding these coevolutionary patterns provides insight into mechanisms of disease spread and maintenance. Arctic rabies (AR) is a lethal disease with viral variants that occupy distinct geographic distributions across North America and Europe. Red fox (*Vulpes vulpes)* are a highly susceptible AR host, whose range overlaps both geographically distinct AR strains and regions where AR is absent. It is unclear if genetic structure exists among red fox populations relative to the presence/absence of AR or the spatial distribution of AR variants. Acquiring these data may enhance our understanding of the role of red fox in AR maintenance/spread and inform disease control strategies. Using a genotyping-by-sequencing assay targeting 116 genomic regions of immunogenetic relevance, we screened for sequence variation among red fox populations from Alaska and an outgroup from Ontario, including areas with different AR variants, and regions where the disease was absent. Presumed neutral SNP data from the assay found negligible levels of neutral genetic structure among Alaskan populations. The immunogenetically-associated data identified 30 outlier SNPs supporting weak to moderate genetic structure between regions with and without AR in Alaska. The outliers included SNPs with the potential to cause missense mutations within several toll-like receptor genes that have been associated with AR outcome. In contrast, there was a lack of genetic structure between regions with different AR variants. Combined, we interpret these data to suggest red fox populations respond differently to the presence of AR, but not AR variants. This research increases our understanding of AR dynamics in the Arctic, where host/disease patterns are undergoing flux in a rapidly changing Arctic

RESEARCH ARTICLE

Genetic structure of immunologically associated candidate genes suggests arctic rabies variants exert differential selection in arctic fox populations

Tristan M. Baecklund[1]*, Michael E. Donaldson[1☯], Karsten Hueffer[2☯], Christopher J. Kyle[1,3,4☯]

1 Environmental and Life Sciences Graduate Program, Trent University, Peterborough, ON, Canada, **2** Department of Veterinary Medicine, University of Alaska Fairbanks, Fairbanks, AK, United States of America, **3** Forensic Science Department, Trent University, Peterborough, ON, Canada, **4** Natural Resources DNA Profiling & Forensic Centre, Trent University, Peterborough, ON, Canada

☯ These authors contributed equally to this work.
* tristanbaecklund@trentu.ca

OPEN ACCESS

Citation: Baecklund TM, Donaldson ME, Hueffer K, Kyle CJ (2021) Genetic structure of immunologically associated candidate genes suggests arctic rabies variants exert differential selection in arctic fox populations. PLoS ONE 16(10): e0258975. https://doi.org/10.1371/journal.pone.0258975

Editor: Hoh Boon-Peng, UCSI University, MALAYSIA

Received: April 23, 2021

Accepted: October 10, 2021

Published: October 29, 2021

Data Availability Statement: Raw sequence data obtained from high-throughput sequencing are available from the NCBI Sequence Read Archive (http://www.ncbi.nlm.nih.gov/bioproject/535441; accession number PRJNA535441).

Funding: This research was funded by a Natural Sciences and Engineering Research Council Discovery Grant to CJK (PGPIN2016-05373). This research was enabled in part by support provided by Compute Canada (www.computecanada.ca;

Abstract

Patterns of local adaptation can emerge in response to the selective pressures diseases exert on host populations as reflected in increased frequencies of respective, advantageous genotypes. Elucidating patterns of local adaptation enhance our understanding of mechanisms of disease spread and the capacity for species to adapt in context of rapidly changing environments such as the Arctic. Arctic rabies is a lethal disease that largely persists in northern climates and overlaps with the distribution of its natural host, arctic fox. Arctic fox populations display little neutral genetic structure across their North American range, whereas phylogenetically unique arctic rabies variants are restricted in their geographic distributions. It remains unknown if arctic rabies variants impose differential selection upon host populations, nor what role different rabies variants play in the maintenance and spread of this disease. Using a targeted, genotyping-by-sequencing assay, we assessed correlations of arctic fox immunogenetic variation with arctic rabies variants to gain further insight into the epidemiology of this disease. Corroborating past research, we found no neutral genetic structure between sampled regions, but did find moderate immunogenetic structuring between foxes predominated by different arctic rabies variants. F_{ST} outliers associated with host immunogenetic structure included SNPs within interleukin and Toll-like receptor coding regions (IL12B, IL5, TLR3 and NFKB1); genes known to mediate host responses to rabies. While these data do not necessarily reflect causation, nor a direct link to arctic rabies, the contrasting genetic structure of immunologically associated candidate genes with neutral loci is suggestive of differential selection and patterns of local adaptation in this system. These data are somewhat unexpected given the long-lived nature and dispersal capacities of arctic fox; traits expected to undermine local adaptation. Overall, these data contribute to our understanding of the co-evolutionary relationships between arctic rabies and their primary host and provide data relevant to the management of this disease.

Appendix III – Chapter 2 iHS and XP-EHH Analyses

Subsequent to the publication of Chapter 2 in 2021, "*The role of a mechanistic host in maintaining arctic rabies variant distributions: Assessment of functional genetic diversity in Alaskan red fox (Vulpes vulpes)*", the *REHH* package was used to estimate iHS and XP-EHH statistics received updates that made the analysis more amenable to unphased multi-locus genotypes (Gautier & Vitalis, 2012; Gautier, Klassman, & Vitalis, 2017; R Core Team R, 2013; Klassman & Gautier, 2022). Meaning that at the time I failed to comprehend that *REHH* was unable to properly process these data leading to potential bias or error in the evaluation of these metrics and impacting interpretation of selection within/between fox populations. As such, I have re-produced these analyses using a different program, Selscan 2.0, which can calculate iHS and XP-EHH from unphased multi-locus genotypes following their new definition of the statistics for these data types (Sziech 2024). Both iHS and XP-EHH calculations were performed using physical distances rather than genetic distances (--pmap), a minor allele frequency of 1% (--maf 0.01) and allowed the integration of truncated EHH curves at loci near boundaries of the data where haplotype information begins to run out and a stopping point is unlikely to be reached (--trunc-ok). iHS and XP-EHH statistics were normalized using a log transformation of a Gaussian cumulative distribution function with a 'p-value' threshold of 4 denoting loci putatively under selection, as implemented in *REHH*, allowing for a direct comparison of results between programs (Gautier & Vitalis, 2012; Gautier, Klassman, & Vitalis, 2017).

iHS was successfully calculated using Selscan at 208, 210, and 314 loci for the Seward Peninsula, Southwest, and Central red fox populations respectively. Of these loci, 9, 11, and 12 for each of the respective populations had normalized iHS 'p-values' surpassing the threshold of four, indicative of putative selection acting on those loci (Fig A3.1). Although this particular analysis does not facilitate the identification of candidate regions as in the initial analyses with *REHH*, of the loci where outlier iHS values were obtained, two were in common amongst all three populations, two in common between the Central Alaska and Seward Peninsula populations, and two in common between Seward Peninsula and Southwestern Alaska populations (Table A3.1).

In contrast to the analyses in the body of Chapter 2, XP-EHH analyses with Selscan did not reveal any loci between any of the two populations comparisons that were approaching or at fixation in one and not the other (Fig A3.2; Table A3.2). It is important to note however, that this may be due to the filtering required to make the SNP data compatible with Selscan. Selscan is unable to handle missing data, as such loci were pruned to ensure no missing data and to retain the same loci within each population. These filtering steps enabled XP-EHH estimations to be computed for only 1010, 1078, and 1060 loci for the Seward Peninsula & Southwest, Central & Southwest, and Central and Seward Peninsula comparisons respectively (Table A3.2).

Overall, analysis of iHS within each of the populations using Selscan revealed similar signatures of selection as to those previously identified with *REHH*.

Discrepancies regarding the differential signatures of selection between populations using XP-EHH can likely be explained by the filtering steps required to compute the statistics with Selscan, although it should be noted imputing missing data prior to analysis may retain additional loci for the calculations providing a more robust analysis.

Table A3.1. iHS estimates for three populations of red fox from Alaska using Selscan. Table denotes the position of each analyzed SNP, and the analogous p-value used to infer signatures of selection. Loci surpassing a p-value threshold of four were taken as putatively under selection in that population. SNPs that appeared under selection in multiple populations are highlighted in grey.

Table A3.2. XP-EHH estimates between three populations of red fox from Alaska using Selscan. Table denotes the position of each analyzed SNP and the analogous p-value of the XP-EHH statistic used to infer signatures of selection. Loci surpassing a p-value threshold of four were taken as putatively under selection in that population.

Tables available at: https://trentu-my.sharepoint.com/:f:/g/personal/tristanbaecklund_trentu_ca/Erpi0dNjHLFIoVAByK9X5sUBWBgCziTMdkjVwARbfj5wuQ?e=J59mVc

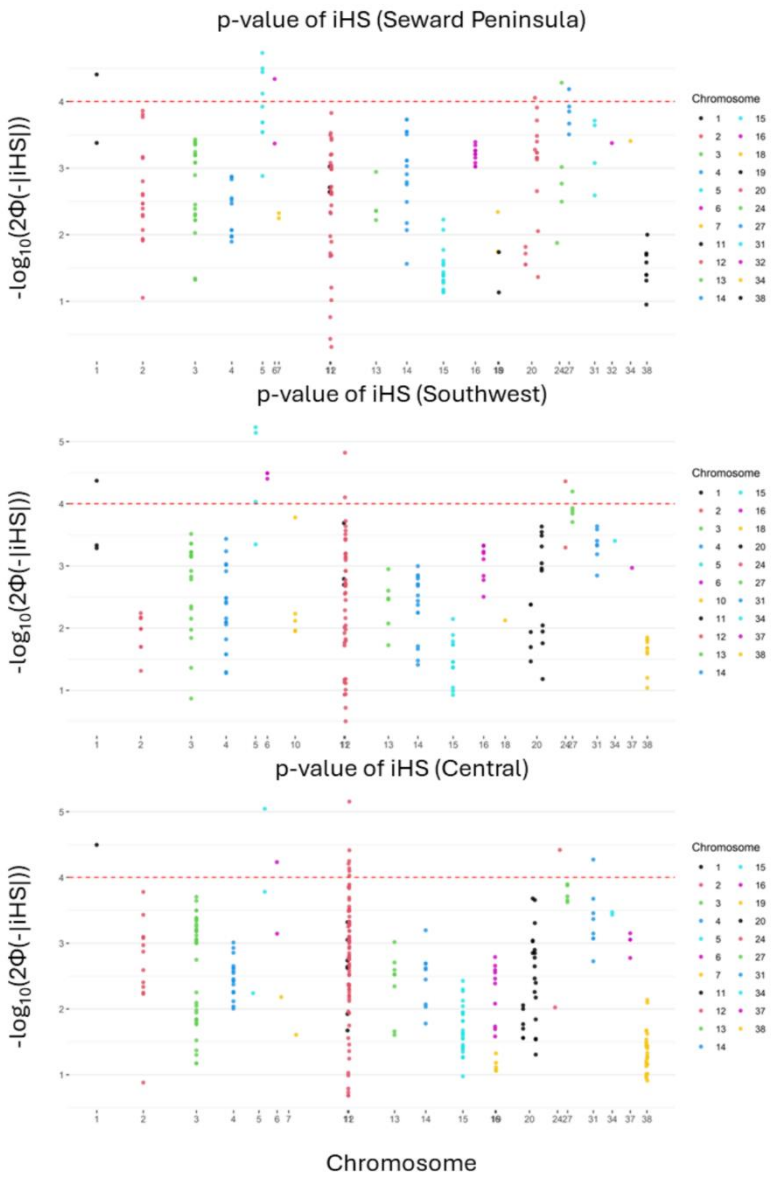

Fig A3.1. Estimates of iHS using Selscan detects weak signals of selection within three populations of red fox in Alaska. Log transformed iHS values were used to normalize the data and provide an analogous p-value, measured against a threshold of 4 (red dotted line), to identify loci putatively under selection within red fox populations from Seward Peninsula, Central, and Southwestern Alaska.

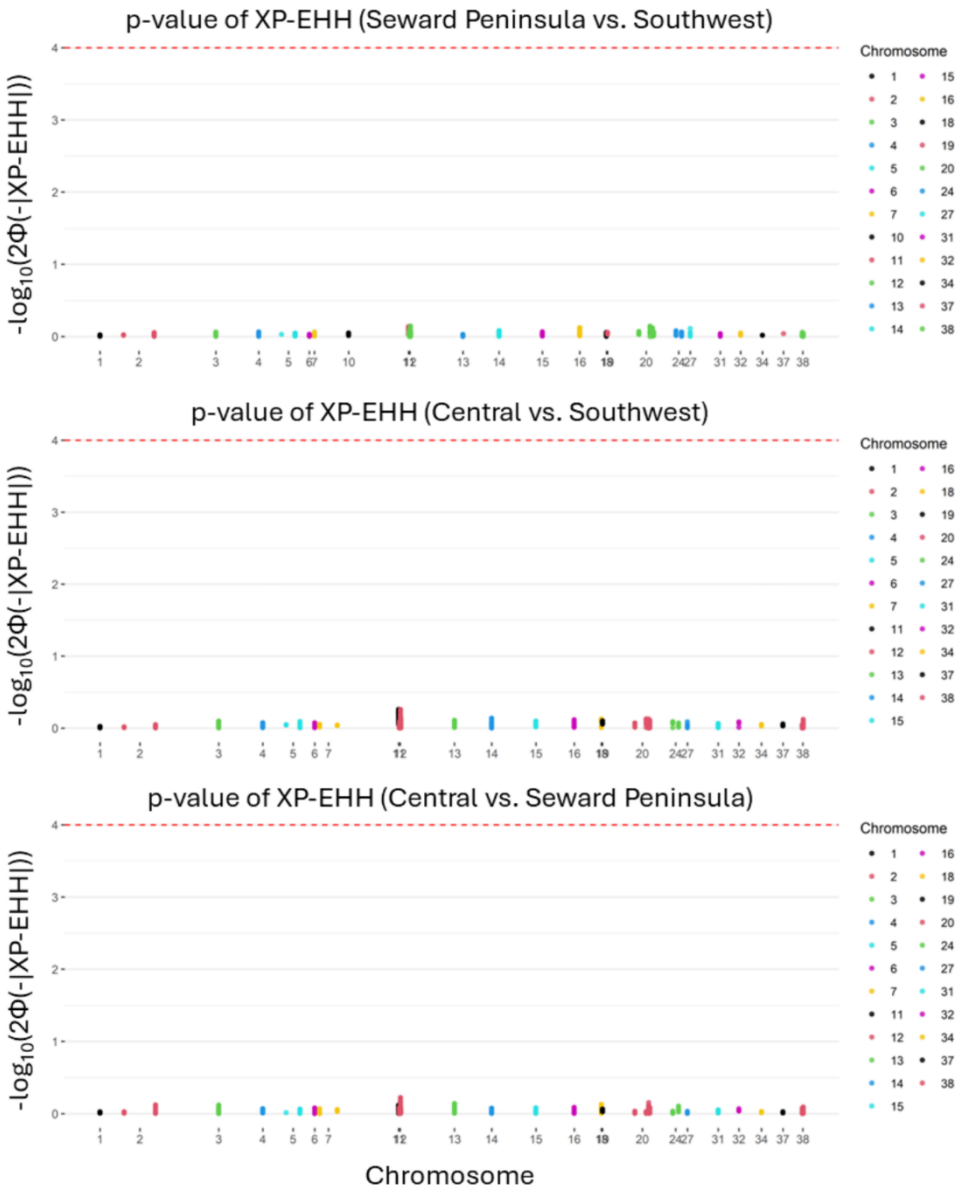

Fig A3.2. Estimates of XP-EHH using Selscan detect no signatures of selection between comparisons of three populations of red fox in Alaska. Log transformed XP-EHH values were used to normalize the data and provide an analogous p-value, measured against a threshold of 4 (red dotted line), to identify loci putatively under selection between red fox populations from Seward Peninsula, Central, and Southwestern Alaska.

Appendix III References

1. Gautier M. and Vitalis R. rehh: An R package to detect footprints of selection in genome-wide SNP data from haplotype structure. Bioinformatics, 2012, 28(8), 1176-1177. https://doi.org/10.1093/bioinformatics/bts115
2. Gautier M., Klassmann A., and Vitalis R. rehh 2.0: a reimplementation of the R package rehh to detect positive selection from haplotype structure. Molecular Ecology Resources 2017(1), 78-90. https://doi.org/10.1111/1755-0998.12634
3. R Core Team R. R: A language and environment for statistical computing. 2013.
4. Klassmann A, Gautier M. Detecting selection using extended haplotype homozygosity (EHH)-based statistics in unphased or unpolarized data. PloS one. 2022 Jan 18;17(1):e0262024.
5. Szpiech ZA. selscan 2.0: scanning for sweeps in unphased data. Bioinformatics. 2024 Jan 1;40(1):btae006.