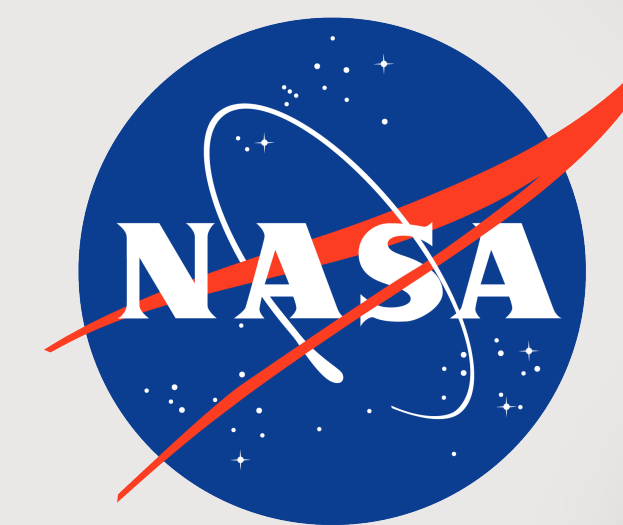# Population Structure and Demography of Gentoo Penguins (*Pygoscelis papua*) in an Area of Range Expansion

Rachael Herman (FI) and Heather Lynch (PI) – Stony Brook University

Introduction

Gentoo penguins are both increasing in population size and expanding their breeding range along the Western Antarctic Peninsula (WAP; Fig. 1) – a region that is undergoing rapid climate warming resulting in declines in overall concentrations of sea ice.

However, concentrations still vary at fine scales from year to year (Fig 2). We hypothesize that years of low sea ice in the area of range expansion during the winter allow for dispersal to novel suitable habitat.

Fig 2: Annual mean sea ice concentrations during the austral winter period in the area of range expansion on the WAP. Compiled from NSIDC-processed NASA Team monthly passive microwave sea ice concentrations.
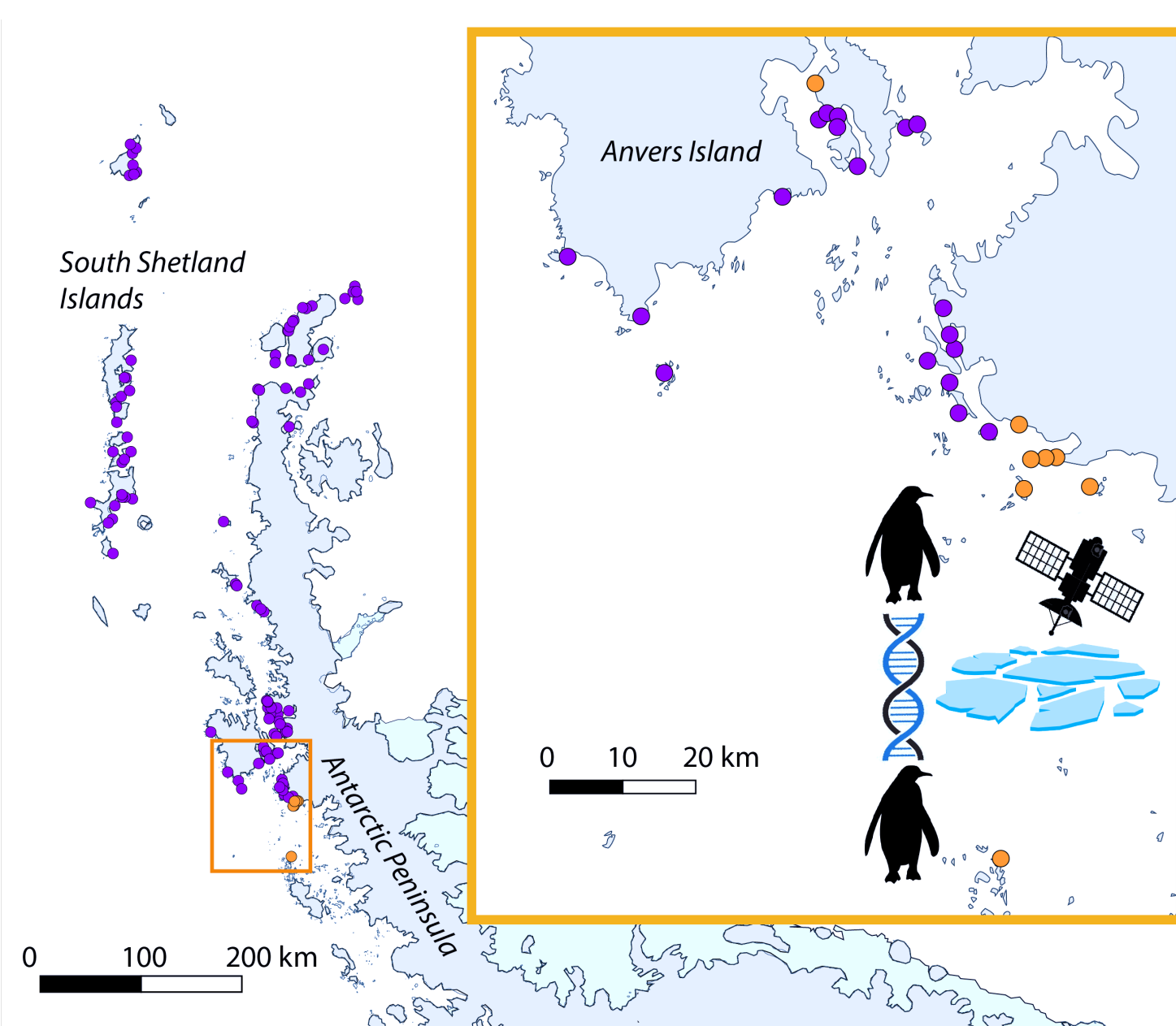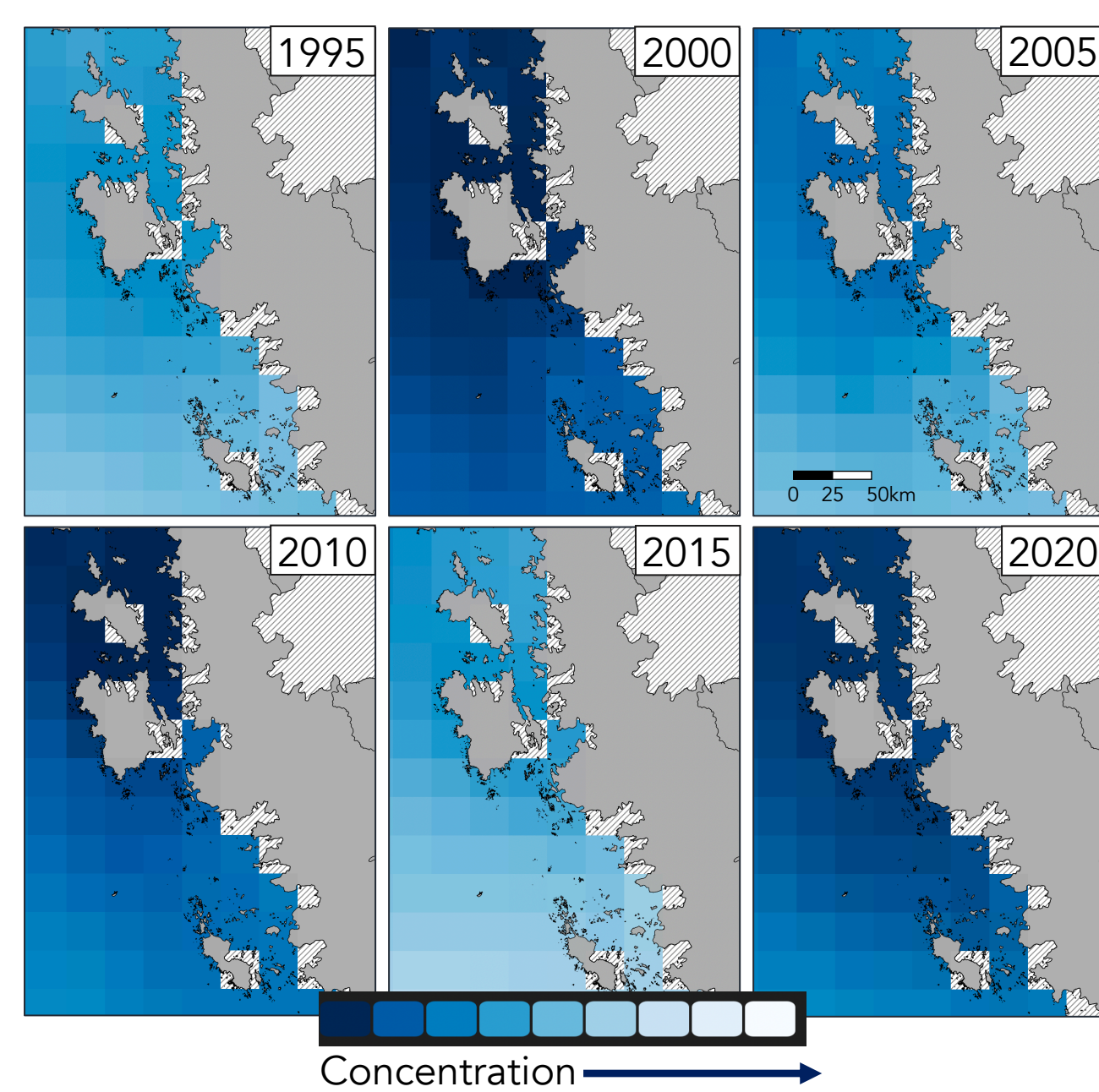

Fig 1: Map and inset of area of range expansion. Purple dots represent all known gentoo penguin colonies and orange denote colonies established since 1994.

OBJECTIVES

Use genomic methods to:

1. Determine how colonies within the area of range expansion are related

2. Identify source populations of newly established colonies

3. Identify potential networks of dispersal that overlap with areas of low sea ice and high lead density.

Materials and Methods

Collected 72 blood samples from breeding gentoo penguin individuals across:

5 potential source colonies:
- Damoy Point
- Danco Island
- Neko Harbor
- Port Charcot
- Petermann Island

2 newly established colonies:
- Moot Point
- Tuxen Rocks

We also collected 1 non-breeding vagrant individual far south at Armstrong Reef

FI Herman and PI Lynch collecting a blood sample from a gentoo penguin at Moot Point.

Bioinformatics Pipeline

Whole genome sequencing (WGS) at 20x coverage with 150bp paired-end reads

Cleaned and mapped reads to published gentoo penguin reference genome

Hard filtered for high quality SNPs using *HaplotypeCaller* in GATK

Ran *BQSR* in GATK to recalibrate base quality scores

Variant calling with VCFtools

Filtered out SNPs on sex chromosomes

Genomic Analyses

- smartPCA: employs the combination of Principle Component Analysis and statistical models specific to genomic data to identify population structure and clustering patterns (Fig. 3)

- ADMIXTURE: maximum likelihood estimation of individual ancestries from multilocus, autosomal SNP genotype datasets (Fig. 4)

Preliminary Results

SmartPCA

- Clear genetic clustering of regions within southern area of the gentoo penguin breeding range

- Clustering of the most southern four colonies suggests founders of Moot Point and Tuxen Rocks likely came from Petermann Island and/or Port Charcot

- Vagrant individual found at Armstrong Reef likely originated from the more northern colonies (such as Neko Harbor or Danco Island)

ADMIXTURE

- K of 3 reflects the same clustering and population structure as the PCA results

- Two distinct ancestral populations, blue (Moot Point, Port Charcot, Petermann Island, and Tuxen Rocks) and gold (Damoy Point), suggesting no gene flow between them

- Danco Island and Neko Harbor assigned to third population (purple), but results suggest individuals from Damoy Point have dispersed to these colonies

What does this mean?

Gentoo penguins that established new colonies in the area of southern range expansion likely originated from colonies close by (stepping-stone pattern)

Yet some penguins may have the propensity to disperse much farther, as suggested by the individual found at Armstrong Reef

Next Steps...

- Incorporate 57 additional samples into these analyses that were collected from four more potential source populations

- Conduct demographic modelling using haplotype data to detect potential finer-scale population structure and the timing of historical gene flow

- Investigate if colonization events overlap with periods of low sea ice concentrations and high lead density during the non-breeding season using a combination of population census time series, passive microwave data (Fig. 2), and ICEsat-2 data (Fig. 5)

- Build an agent-based model to forecast continued gentoo penguin range expansion based under predicted sea ice changes
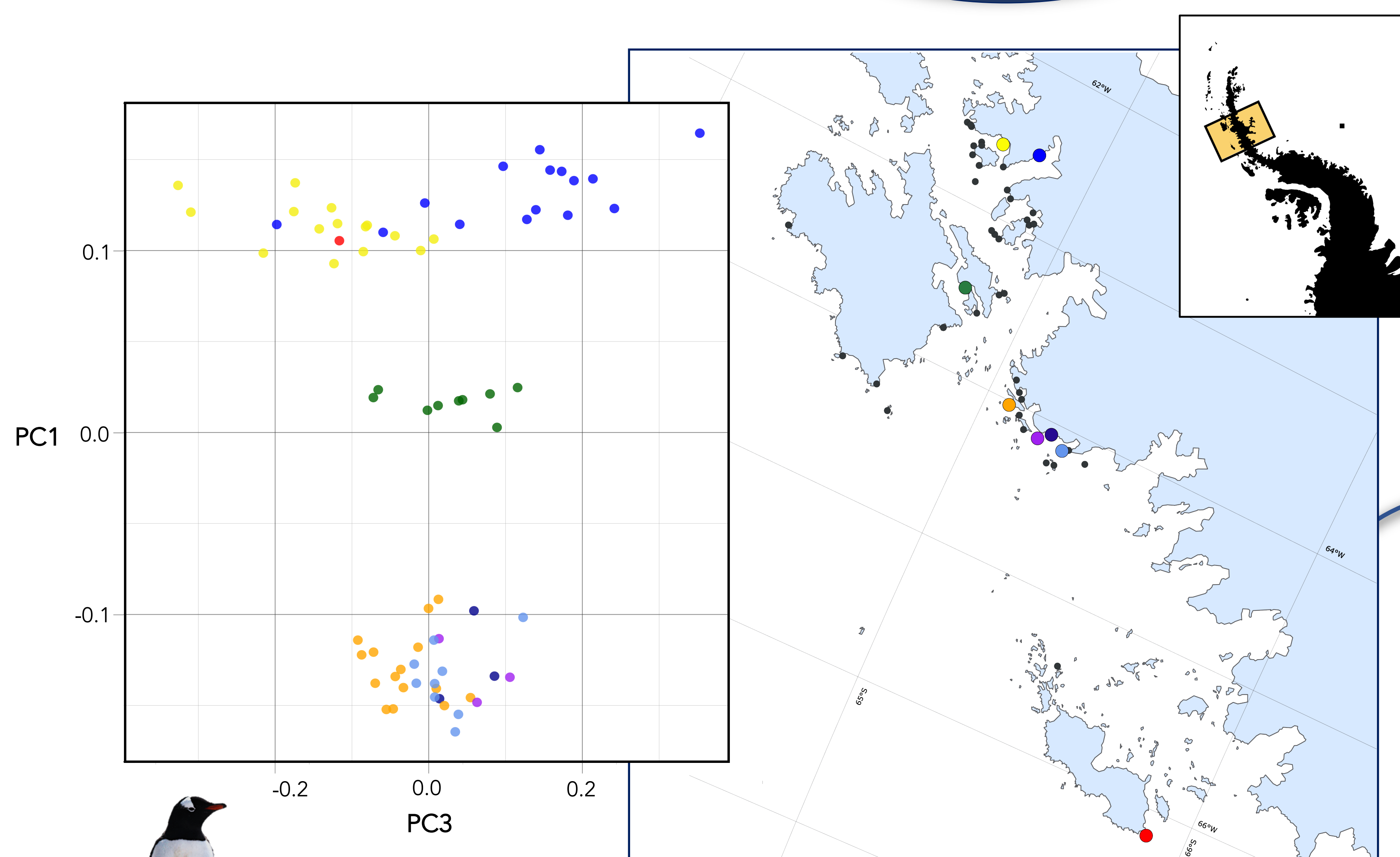
Fig 3: smartPCA results and map of colony locations on the WAP. Smaller circles mark all known gentoo penguin colonies.
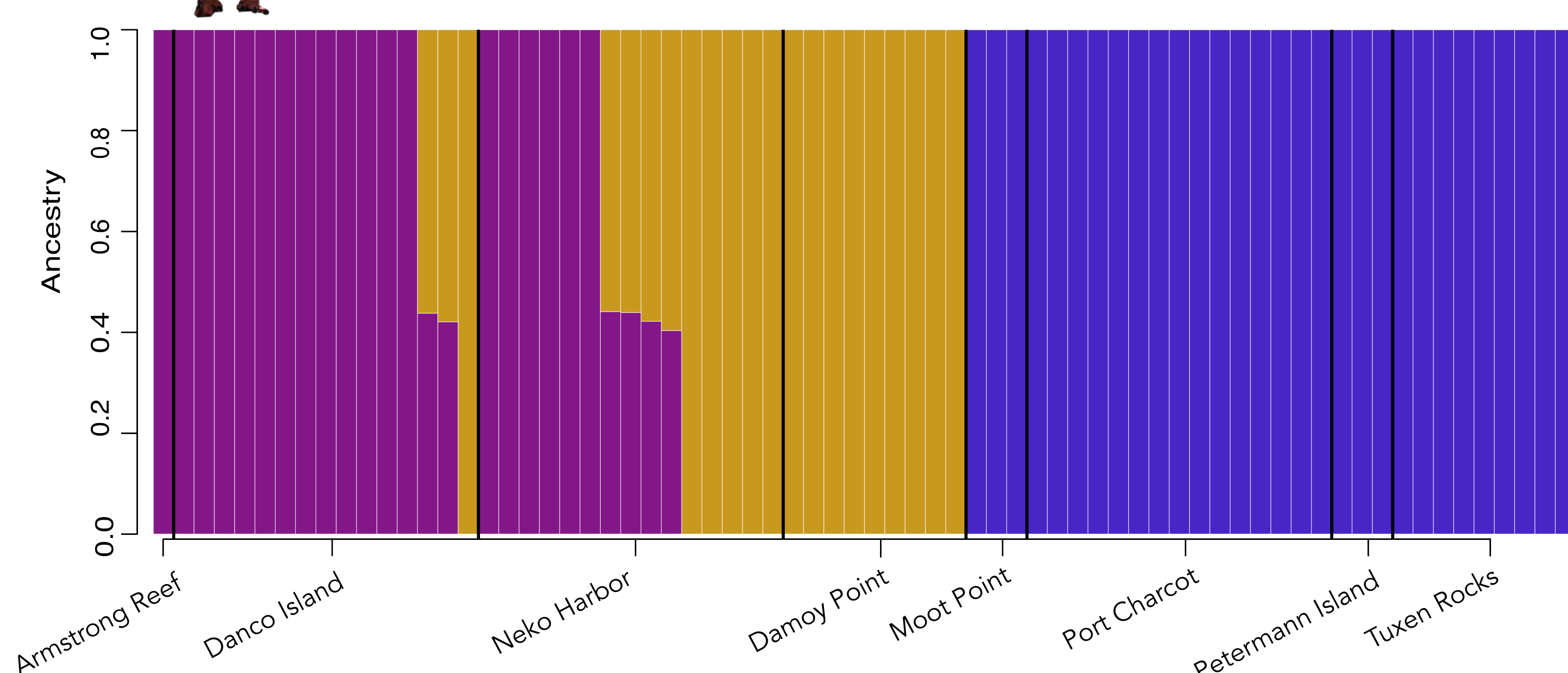

Fig 4: Admixture results for K = 3 (hypothesized number of populations of origin). Each individual is represented by a bar. Individuals are assigned to a population based on the proportion of ancestry from each gene pool present in the data.
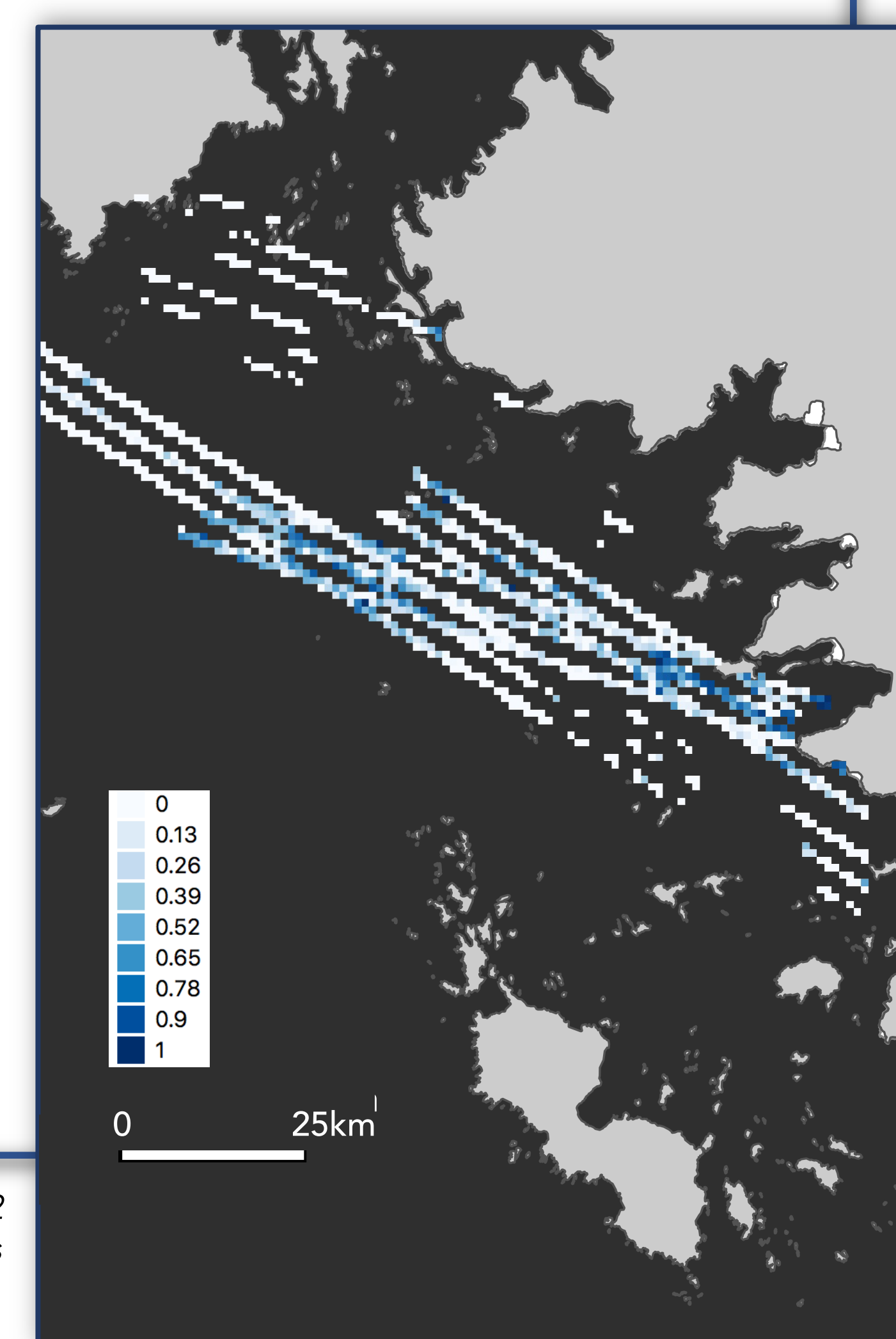

Fig 5: Raster of mean lead density extracted from raw ICEsat-2 tracks. Darker shades of blue represent higher density of leads in sea ice. Data is averaged over the non-breeding season (June, July, and August) of 2020.