

SALS microsatellite population structure manuscript

- Map figure documentation

The below information contains a complete record of the data sources, input file generation and structure, analysis and Rcode, output data and/or figures generated, and session information I used to complete the above title-referenced analysis for publication.

The following can also be compiled in RStudio from the file [SALSusat-Map.Rnw](#)

Motivation: Presented below is a map of the 24 breeding sites sampled in this paper. The output of the below figure should be essentially identical to the old figure given to me by Adrienne (see page 2 of file [SALS PopGen Summary 2-2018.pdf](#), from February 4, 2018 or slide/page 13 of file [SHARP - metapop gen PI retreat.pdf](#), from August 17, 2016). I have only rerun it here in order to make the final figure a vector format and cleaned up to be more suitable for publication; as well as to more fully document the figure generation for posterity.

Add required packages:

```
> library(ggplot2)
> library(ggmap)
> library(ggspatial)
> library(sf)
> library(rnaturalearth)
> library(maps)
> library(cowplot)
> #library(tools)
> library(dplyr)
```

Read in data*:

```
> sites <- read.table("~/Documents/ModernSparrowGenomics/Copies of Associated ELN
  files/SALS usat ms/Data and Analyses/SiteCoords.txt", header = TRUE)
```

*Documentation of data sources and structure:

1.) File [SiteCoords.txt](#).

Source: These data are the decimal degree coordinates for each marsh found on the 'Coords' tab in the [SHARP Marsh Coords.xlsx](#) file. I'm not sure who originally generated this file (dated July 29, 2016) or supplied the marsh coordinates, presumably Jen Walsh, Adrienne Kovach or maybe Bri Benvenuti. This text file is a copy of the Coords tab, formatted for R (Genalex-specific formatting removed, spaces in site names replaced with underscores).

Data structure: Rectangular matrix of latitude ("Y") and longitude ("X") coordinates for each breeding marsh studied here ("Site")

Data units: Decimal degrees

Data snippet:

```
> head(sites)

      X      Y      Site
1 -70.26303 43.58182   Spurwink
2 -70.34428 43.56702  Scarborough
3 -70.36634 43.53317      Jones
4 -70.54228 43.34247 Little_River
5 -70.57074 43.31229   Eldridge
6 -70.58672 43.27407    Furbish
```

Convert coordinates to spatial vector object:

```
> (sites <- st_as_sf(sites, coords = c("X", "Y"), remove = FALSE,
+   crs = 4326, agr = "constant"))
```

Notes on above code:

remove = a logical (true/false) argument for whether or not to remove the coordinate columns from data frame.

crs = coordinate reference system; the number 4326 refers to EPSG:4326, which is the European Petroleum Survey Group code for the World Geodetic System 1984 (WGS84) datum.

From the 'sf' R package documentation vignette: "agr, attribute-geometry-relationship, specifies for each non-geometry attribute column how it re-lates to the geometry, and can have one of following values: "constant", "aggregate", "identity"."constant" is used for attributes that are constant throughout the geometry (e.g. land use)" - or in this case, the site names.

Setup site names formatting and designation of hybrid zone sites:

```
> labels <- c("Spurwink"="Spurwink", "Scarborough"="Scarborough", "Jones"="Jones",
  "Little_River"="Little River", "Eldridge"="Eldridge", "Furbish"="Furbish", "
  Lubberland"="Lubberland Creek", "Chapmans"="Chapman's Landing", "Fairhill"="
  Fairhill", "Hampton"="Hampton", "Parker_River"="Parker River", "Monomoy"="
  Monomoy", "Sachuest"="Sachuest", "Chafee"="Chafee", "Barn_Island"="Barn Island
  ", "East_River"="East River", "Hammonasset"="Hammonasset", "Shirley"="Wertheim
  ", "Idlwild"="Idlewild", "MNC"="Marine Nature Center", "Sawmill"="Sawmill", "
  Four_Sparrow"="Four Sparrow", "ATT"="ATT", "OC-MU"="OC-Mullica")
> HybridSites <- subset(sites, Site %in% c("Spurwink", "Scarborough", "Jones", "
  Little_River", "Eldridge", "Furbish", "Lubberland", "Chapmans", "Fairhill", "
  Hampton", "Parker_River"))
```

Create main map:

```
> world <- ne_countries(scale = "medium", returnclass = "sf")
> states <- st_as_sf(map("state", plot = FALSE, fill = TRUE))
> states2 <- states %>% filter(ID %in% c(
+ "maine", "new hampshire", "vermont", "connecticut", "rhode island", "
+ massachusetts", "new york", "new jersey"
+ ))
> sf::sf_use_s2(FALSE)
```

Notes on above code:

sf::sf_use_s2(FALSE) = there was some kind of error when trying to extract the centroid coordinates from the states object for adding state labels; this appears to be a bug in sf version 2, so that line reverts the program back to version 1 (see [this answer on stackoverflow](#))

toupper = converts text in an object field to uppercase. (FYI - can use 'toTitleCase' after installing the 'tools' package to convert to title case.)

nudge_y = allows you to set individual aesthetics, in this case, specify exact postioning of the text labels so they "look" more centered geographically. (Use nudge_x to shift labels left or right.)

geom_sf = draws geometric objects, such as the state polygons and site points.

geom_text_repel = keeps labels from overlapping one another

coord_sf = sets the scale/viewed area of the map based on bounding latitude-longitude coordinates.

annotation_scale = inserts the scale bar in the bottom right corner ("br")

annotation_north_arrow = inserts a north arrow (can be set to true or magnetic north) in the bottom right; the padding adjusts the placement so it doesn't overlap the scale bar.

Create inset map:

```
> usa <- subset(world, admin == "United States of America")
> (mainland <- ggplot(data = usa) + geom_sf(fill = "#456462") + geom_rect(xmin =
  2135380.81, xmax = 2444854.36, ymin = -309447.13, ymax = 359165.16, fill = NA,
  colour = "black", size = 1) + theme_void() + theme(panel.background = element
  _rect(fill = "aliceblue")) + coord_sf(crs = st_crs(2163), xlim = c(-2000000,
  2500000), ylim = c(-2300000, 730000)))
```

Notes on above code:

geom_rect - draws the inset rectangle. I used the same bounding coordinates as were set for the main map (xlim = c(-74.85, -69.05), ylim = c(39.25, 43.87)), however since a different projection is used on the smaller scale map of the US, the decimal degrees had to be converted to the different projection.

*Handy resource for getting transformed coordinates among projections: <https://epsg.io/transform>

EPSG:2163 - US National Atlas Equal Area - "Unspecified datum based upon the Clarke 1866 Authalic Sphere"

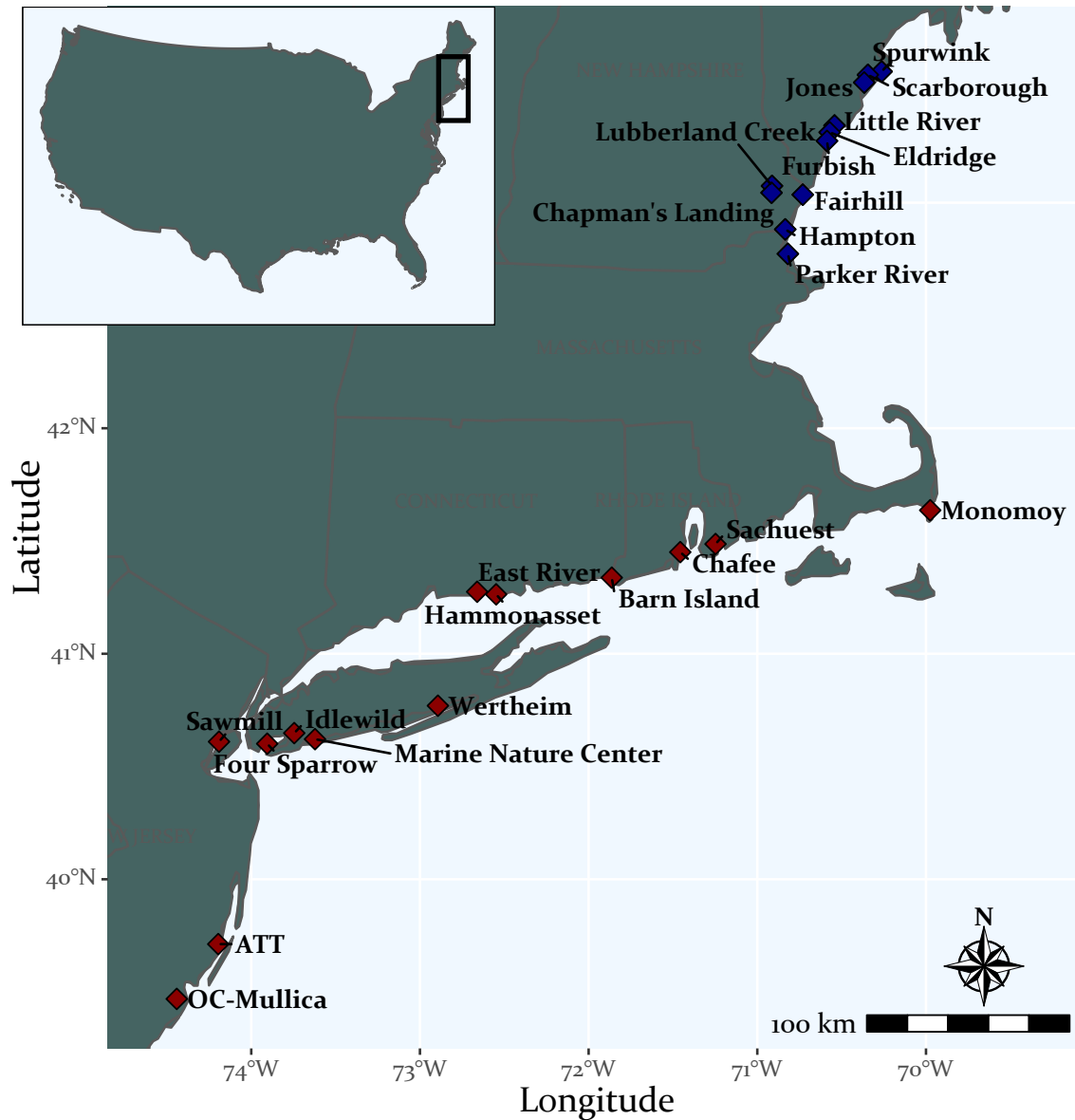
Format final figure for publication:

```
> FinalMap <- ggdraw(map2) +draw_plot(mainland, width = 0.43, height = 0.5, x =  
0.02, y = 0.583)
```

Save final figure:

```
> ggsave(file = "/Users/Lindsey/Documents/ModernSparrowGenomics/Copies of  
Associated ELN files/SALS usat ms/Final Figures/FinalMap.pdf", plot = FinalMap  
, device = cairo_pdf, width = 6.5, units = "in")
```

Copy of resulting **FinalMap.pdf** plot:



Record of Session Info:

```
> sessionInfo()
```

R version 4.0.3 (2020-10-10)

Platform: x86_64-apple-darwin17.0 (64-bit)

Running under: macOS Big Sur 10.16

Matrix products: default

BLAS: /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRblas.dylib

LAPACK: /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRlapack.dylib

locale:

[1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8

attached base packages:

[1] stats graphics grDevices utils datasets methods base

other attached packages:

[1] dplyr_1.0.2 cowplot_1.1.1 maps_3.4.0
[4] rnatuarearth_0.1.0 sf_1.0-3 ggspatial_1.1.5
[7] ggrepel_0.9.1 ggplot2_3.3.5

loaded via a namespace (and not attached):

[1] Rcpp_1.0.8 pillar_1.7.0 compiler_4.0.3
[4] class_7.3-17 tools_4.0.3 digest_0.6.29
[7] lattice_0.20-41 lifecycle_1.0.1 tibble_3.1.6
[10] gtable_0.3.0 pkgconfig_2.0.3 rlang_1.0.1
[13] cli_3.2.0 DBI_1.1.0 rstudioapi_0.13
[16] rnatuarearthdata_0.1.0 e1071_1.7-7 withr_2.4.3
[19] rgeos_0.5-8 generics_0.1.2 vctrs_0.3.8
[22] classInt_0.4-3 grid_4.0.3 tidyselect_1.1.0
[25] glue_1.4.2 R6_2.5.1 fansi_1.0.2
[28] sp_1.4-5 farver_2.1.0 purrr_0.3.4
[31] magrittr_2.0.2 scales_1.1.1 ellipsis_0.3.2
[34] units_0.7-2 colorspace_2.0-3 utf8_1.2.2
[37] KernSmooth_2.23-17 proxy_0.4-26 munsell_0.5.0
[40] crayon_1.5.0