

translocation analysis - format data

steph

2024-09-04

pre-process survival and phenotype data

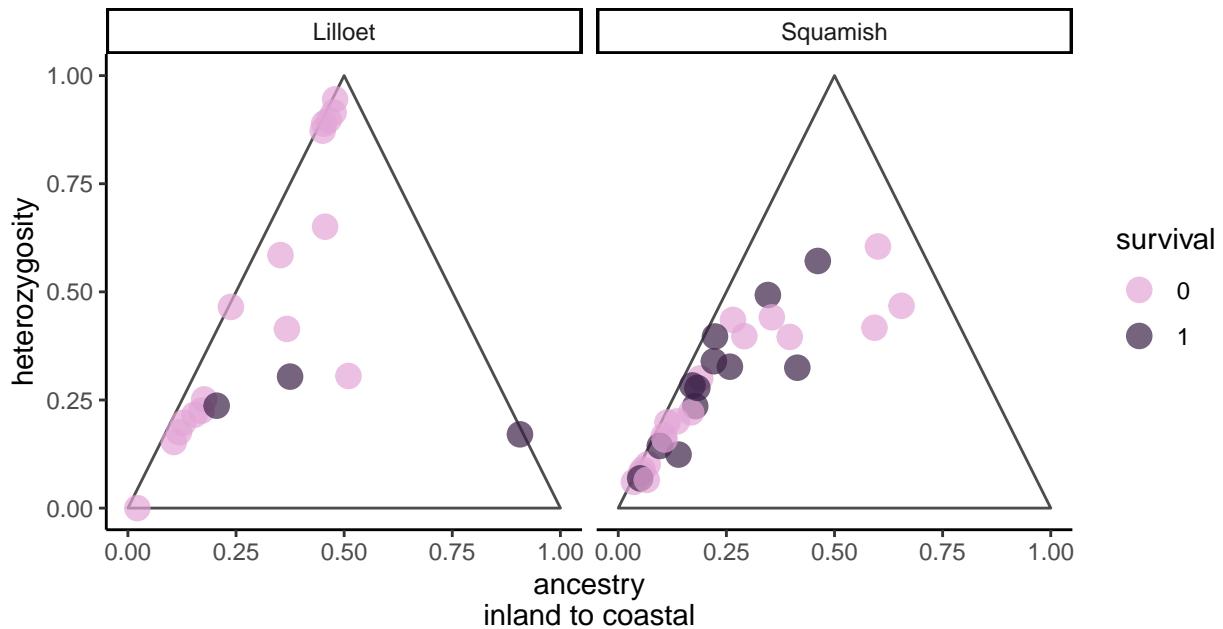
load packages and read in initial dataset

- read in data
- create a merged metadata df

```
lat.df<-read_xlsx("C:/Users/Steph/TAMU_OneDrive/Thrushes/translocation/data_archiving/translocation_data"
birds<-read_xlsx("C:/Users/Steph/TAMU_OneDrive/Thrushes/translocation/data_archiving/translocation_data"
pembertonBirds<-read_xlsx("C:/Users/Steph/TAMU_OneDrive/Thrushes/translocation/data_archiving/translocation_data"
meta_thrush<-rbind(birds,pembertonBirds)

poly.all<-data.frame(X.co=rep(c(0,0.5,1)), #make triangle polygon for plotting
                      Y.co=rep(c(0,1,0)))

ggplot(birds,aes(x=aims_ancestry,y=aims_heterozygosity,colour=as.character(survival)))+
  geom_polygon(data=poly.all,aes(x=X.co,y=Y.co),fill=NA,colour="grey30")+
  geom_point(size=4,shape=19,alpha=0.7)+  
  scale_color_manual(values=natparks.pals("Arches2")[c(4,1)],name='survival')+
  coord_equal()+
  xlim(0,1)+ylim(0,1)+  
  facet_wrap(vars(release_site))+  
  xlab("ancestry\ninland to coastal")+
  ylab("heterozygosity")
```



```
### format capture histories for input to CJS models
```

```
ch.df<-data.frame(matrix(NA,nrow=1,ncol=301))

#for each individual in translocated birds dataset
for(i in c(birds$name_in_vcf,pembertonBirds$name_in_vcf)){
  x1<-rep(0,300)
  x2<-rep(0,300)

  y<-lat.df%>%filter(name_in_vcf==i)%>%
    select(name_in_vcf,tagDays)%>%distinct()

  y1<-y%>%pull(tagDays)
  x1[c(1,y1[y1<301])<-1 #add 1 to detections because bird was caught the day it was tagged
  if(length(y1)>0){if(max(y1)>300){
    x1[300]<-1 }} #fill in last day if bird caught later

  ch.df<-rbind(ch.df,c(i,x1))

}

colnames(ch.df)[1]<-"name_in_vcf"
if(is.na(ch.df[1,1])){ch.df<-ch.df[-1,]}

ch.df<-ch.df%>%unite(ch,-name_in_vcf,sep="")
```

```

ch.df.pemberton<-ch.df%>%filter(name_in_vcf%in%pembertonBirds$name_in_vcf)

ch.df<-ch.df%>%filter(name_in_vcf%in%birds$name_in_vcf)%>%
  left_join(birds)%>%
  mutate(ch=paste("days",ch,sep="_"))%>% #stops capture history from being saved as a very large number
  select(name_in_vcf,motustagid,aims_ancestry,release_site,sex_binary,ch)

#check that number of characters in detection history matches expectation for sets of 10 days
nchar(ch.df$ch[1])==(300+5)

```

```
## [1] TRUE
```

```

# write.csv(ch.df,
#           paste("C:/Users/Steph/GitHub/thrush_hybrids/translocation/thrush_survival_ch_",today,".csv")
#           )

```

- group detection histories into sets of ten days - 1 if detected in 10 day period

```

ch.df<-ch.df%>%mutate(ch2=gsub("days_","",ch))

ch2<-data.frame(name_in_vcf=ch.df$name_in_vcf)
days<-10
for(i in 1:round((nchar(ch.df$ch2)[1]/days))){
  colID=paste("t",i,sep="")
  ch2[colID]<-rep(0,nrow(ch.df))
  j=i*days-(days-1) #get starting position
  ch2[grep("1",substr(ch.df$ch2,j,j+(days-1))),colID]<-1
}

ch2<-ch2%>%unite(ch,-name_in_vcf,sep="")

ch2<-ch2%>%left_join(birds)%>%
  mutate(ch=paste("days",ch,sep="_"))%>% #stops capture history from being saved as a very large number
  select(name_in_vcf,motustagid,aims_ancestry,release_site,release_year,sex_binary,ch)
# write.csv(ch2,
#           paste("C:/Users/Steph/GitHub_data/translocation/survival.translocatedBirds.ch_10days",today)
#           )

ch.df.pemberton<-ch.df.pemberton%>%mutate(ch2=gsub("days_","",ch))

ch2<-data.frame(name_in_vcf=ch.df.pemberton$name_in_vcf)
days<-10
for(i in 1:round((nchar(ch.df.pemberton$ch2)[1]/days))){
  colID=paste("t",i,sep="")
  ch2[colID]<-rep(0,nrow(ch.df.pemberton))
  j=i*days-(days-1) #get starting position
  ch2[grep("1",substr(ch.df.pemberton$ch2,j,j+(days-1))),colID]<-1
}

ch2<-ch2%>%unite(ch,-name_in_vcf,sep="")

```

```

ch2<-ch2%>%left_join(pembertonBirds)%>%
  mutate(ch=paste("days",ch,sep="_"))%>% #stops capture history from being saved as a very large number
  select(name_in_vcf,motustagid,aims_ancestry,release_site,release_year,sex_binary,ch)

#check that number of characters in detection history matches expectation for sets of 10 days
nchar(ch2$ch[1])==(300/10+5)

## [1] TRUE

# write.csv(ch2,
#           paste("C:/Users/Steph/GitHub_data/translocation/survival.PembertonBirds.2020_2023.ch_10days"
#           )

```

- make a map

```

lat.df.birds<-lat.df%>%filter(name_in_vcf%in%c(birds$name_in_vcf,pembertonBirds$name_in_vcf))
#load map data
world <- ne_countries(scale = "medium", returnclass = "sf")

lat.df.birds$release_site<-factor(lat.df.birds$release_site,levels=c("Squamish","Pemberton","Lilloet"))
lat.df.birds<-lat.df.birds%>%left_join(meta_thrush%>%dplyr::select(name_in_vcf,aims_ancestry))
lat.df.birds$ancestry<-lat.df.birds$aims_ancestry*(-1)+1

map1<-ggplot() +
  geom_sf(data = world,fill="grey90",colour=NA)+
  geom_path(data=lat.df.birds,aes(x=recvDeployLon,y=recvDeployLat,
                                    group=name_in_vcf,colour=ancestry),size=0.3,alpha=0.4)+
  #scale_colour_manual(values=natparks.pals("Banff")[c(4,1)])+
  geom_point(data=lat.df.birds,aes(x=recvDeployLon,y=recvDeployLat,
                                    colour=ancestry,group=name_in_vcf),size=1.2)+
  #scale_shape_manual(values=c(19,15,17))+
  #scale_colour_gradientn(colours=c(natparks.pals("Banff")[4],"grey10",natparks.pals("Banff")[1]),
  #                       name="ancestry",values=c(0,0.5,1))+

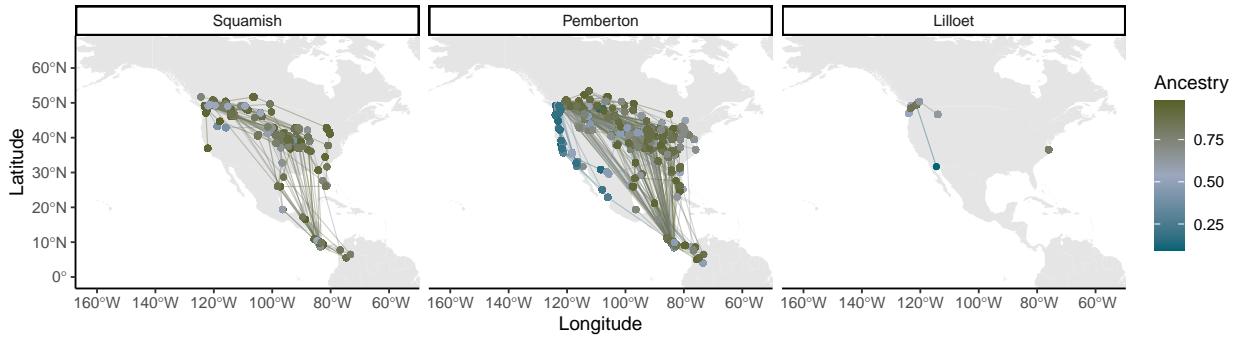
  scale_colour_gradientn(colours=natparks.pals("Banff")[c(1,7,4)],
                         name="Ancestry",values=c(0,0.5,1))+

  xlab("Longitude")+ylab("Latitude")+
  #theme(legend.position="none")+
  ylim(0,66)+xlim(-162,-55)+

  facet_wrap(vars(release_site))

map1

```



```
ggsave("C:/Users/Steph/TAMU_OneDrive/Thrushes/translocation/Fig4.png",
       width=10,height=3)
```

```
projection<-"+proj=longlat +datum=WGS84 +no_defs +ellps=WGS84 +towgs84=0,0,0"

lat.df.birds.fall1<-lat.df%>%
  filter(name_in_vcf%in%c(birds$name_in_vcf,pembertonBirds$name_in_vcf))%>%
  mutate(deployyear=as.numeric(substr(tagDeployStart,1,4)))%>%
  filter(detectyear==deployyear)%>%
  filter(detectdoy>212&tagProjID==280)%>%
  mutate(deployDay=substr(tagDeployStart,1,10),tsDay=substr(ts,1,10))%>%
  mutate(tagDays=as.integer(gsub(" days", " ", difftime(tsDay,deployDay,units="days"))))%>%
  select(motustagid,tsDay,tagDeployStart,tagDepLat,tagDepLon,
         recvDeployID,recvDeployLat,recvDeployLon,recvDeployName,
         detectyear,tagsDays,detectdoy,deployDay,
         deployyear)%>%
  mutate(tagDeployStart=substr(tagDeployStart,1,10))%>%
  distinct%>%arrange(tagDays)%>%
  filter(!is.na(tagDays))%>%
  filter(motustagid!=85685)%>% #broken tag
  mutate(tagDepLat=if_else(motustagid%in%c(91383,85681),50.133366,tagDepLat),
        tagDepLon=if_else(motustagid%in%c(91383,85681),-122.95036,tagDepLon))

#estimate the distance between release site and detection site
#estimation needs to be pairwise, otherwise function will generate a matrix of all possible combos
lat.df.birds.fall1$releaseSiteDistance<-NA
for(i in 1:nrow(lat.df.birds.fall1)){
  lat.df.birds.fall1$releaseSiteDistance[i]<-
    distm(lat.df.birds.fall1[i,c("tagDepLon","tagDepLat")],
          lat.df.birds.fall1[i,c("recvDeployLon","recvDeployLat")],
          fun = distHaversine)
}

lat.df.birds.fall1<-lat.df.birds.fall1%>%
  filter(releaseSiteDistance<300000)%>%
  #keep the first record within the tier for each bird
  arrange(tagDays)%>%filter(!duplicated(motustagid))%>%
  mutate(motustagid=as.character(motustagid))%>%
  left_join(meta_thrush)%>%
  mutate(ancestry=aims_ancestry*(-1)+1)%>%
```

```

mutate(release_site=factor(release_site,levels=c("Squamish","Pemberton","Lilloet")))

lat.df.birds.fall1.summary<-lat.df.birds.fall1%>%
  dplyr::select(ancestry,name_in_vcf,release_site,recvDeployLat,tagDepLat)%>%
  pivot_longer(c(tagDepLat,recvDeployLat),names_to='detect_type',values_to='latitude')%>%
  mutate(detect_type=gsub('Lat',' ',detect_type))%>%
  left_join(
    lat.df.birds.fall1%>%
      dplyr::select(name_in_vcf,recvDeployLon,tagDepLon)%>%
      pivot_longer(c(tagDepLon,recvDeployLon),names_to='detect_type',values_to='longitude')%>%
      mutate(detect_type=gsub('Lon',' ',detect_type)))

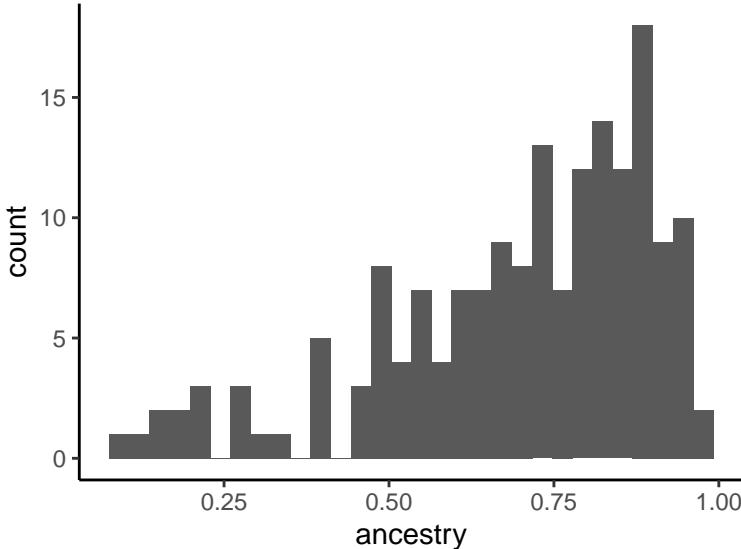
#add bearing measure - check that bearing estimate matches map
lat.df.birds.fall1.summary<-
  lat.df.birds.fall1.summary%>%
  left_join(rbind(birds%>%select(name_in_vcf,fall_releaseSite_bearing_tier1),
                  pembertonBirds%>%select(name_in_vcf,fall_releaseSite_bearing_tier1)))
lat.df.birds.fall1<-lat.df.birds.fall1%>%
  left_join(rbind(birds%>%select(name_in_vcf,fall_releaseSite_bearing_tier1),
                  pembertonBirds%>%select(name_in_vcf,fall_releaseSite_bearing_tier1)))

```

```

ggplot(lat.df.birds.fall1,aes(x=ancestry))+
  geom_histogram()

```



```

ggplot() +
  geom_sf(data = world,fill="grey90",colour=NA)+
  geom_path(data=lat.df.birds.fall1,aes(x=recvDeployLon,y=recvDeployLat,
                                         group=name_in_vcf,colour=ancestry),size=0.4,alpha=0.5) +
  #scale_colour_manual(values=natparks.pals("Banff")[c(4,1)]) +
  geom_point(data=lat.df.birds.fall1,aes(x=recvDeployLon,y=recvDeployLat,
                                         colour=ancestry,group=name_in_vcf),size=1.5) +
  #scale_shape_manual(values=c(19,15,17)) +
  #scale_colour_gradientn(colours=c(natparks.pals("Banff")[4],"grey10",natparks.pals("Banff")[1]),

```

```

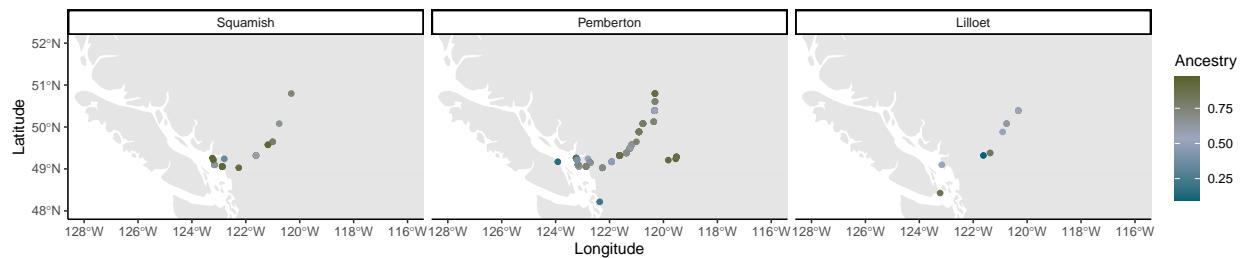
#           name="ancestry",values=c(0,0.5,1))+

scale_colour_gradientn(colours=natparks.pals("Banff")[c(1,7,4)],
                       name="Ancestry",values=c(0,0.5,1))+

xlab("Longitude")+ylab("Latitude")+
#theme(legend.position="none")+
ylim(48,52)+xlim(-128,-116)+

facet_wrap(vars(release_site))

```



```

ggplot() +
  geom_sf(data = world,fill="grey95",colour="grey80")+
  geom_path(data=lat.df.birds.fall1.summary,
            aes(x=longitude,y=latitude,
                 group=name_in_vcf,colour=ancestry),size=0.5,alpha=0.8)+

  #scale_colour_manual(values=natparks.pals("Banff")[c(4,1)])+
  geom_point(data=lat.df.birds.fall1,aes(x=recvDeployLon,y=recvDeployLat,
                                         colour=ancestry,group=name_in_vcf),size=1.5,alpha=0.8)+

  #scale_shape_manual(values=c(19,15,17))+

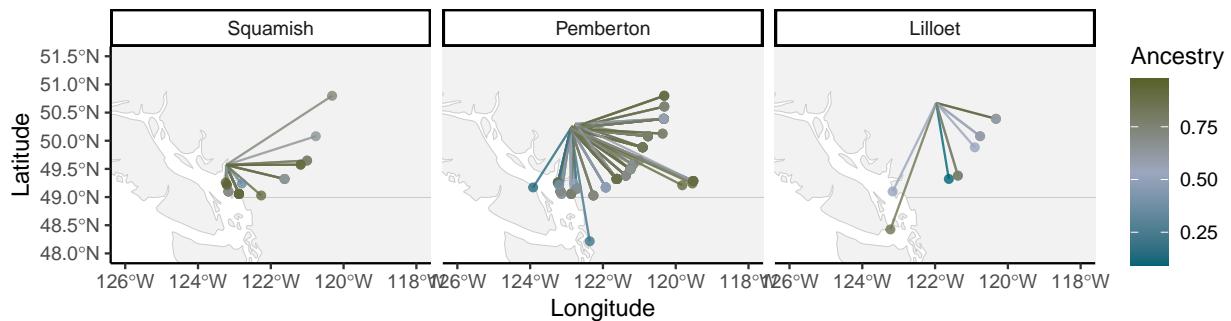
  #scale_colour_gradientn(colours=c(natparks.pals("Banff")[4],"grey10",natparks.pals("Banff")[1]),
  #                       name="ancestry",values=c(0,0.5,1))+

  scale_colour_gradientn(colours=natparks.pals("Banff")[c(1,7,4)],
                         name="Ancestry",values=c(0,0.5,1))+

  xlab("Longitude")+ylab("Latitude")+
#theme(legend.position="none")+
  ylim(48,51.5)+xlim(-126,-118)+

  facet_wrap(vars(release_site))

```



```
ggsave("C:/Users/Steph/TAMU_OneDrive/Thrushes/translocation/Fig3C.png",
       width=8,height=3)
```

verify that estimated bearings match where the birds are going

```
ggplot() +
  geom_sf(data = world,fill="grey95",colour="grey80")+
  geom_path(data=lat.df.birds.fall1.summary,
            aes(x=longitude,y=latitude,
                 group=name_in_vcf, colour=as.numeric(fall_releaseSite_bearing_tier1)),
            size=0.5,alpha=0.8)+
  geom_point(data=lat.df.birds.fall1,aes(x=recvDeployLon,y=recvDeployLat,
                                         colour=as.numeric(fall_releaseSite_bearing_tier1),
                                         group=name_in_vcf),size=1.5,alpha=0.8)+

  scale_colour_viridis(option='plasma',name="orientation")+
  xlab("Longitude")+ylab("Latitude")+
  ylim(48,51.5)+xlim(-126,-118)+

  facet_wrap(vars(release_site))
```

