**Name:** Erik R. Funk
**Address:** 1900 Pleasant St., 334 UCB, Boulder Co., 80309
**Contact:** erik.funk@colorado.edu, 619-916-7354
**Institution:** University of Colorado Boulder
**Co-investigators:** Scott A. Taylor, Nicholas A. Mason

## A large chromosomal inversion explains plumage variation in a globally confusing bird system

Introduction and Goals: Almost every year, redpoll finches make their way into Colorado as they escape the harsher winters of more norther latitudes. The similarity among different redpoll species provides a challenging identification puzzle for birders, and every year Colorado birders are left stumped. While these challenges can be fun field tests for birders, understanding the significance of these differences from an evolutionary perspective, as well as determining their genetic basis, remains challenging. Repolls (*Acanthis* spp.; Figure 1a) in particular are a group that is notorious for presenting identification challenges. These Holarctic birds vary in plumage color from white to brown, and have a long history of taxonomic debate. While many authorities currently recognize three species (Clements et al. 2019), classifications based on plumage and morphology have ranged from one to six species (Coues 1862, Troy 1985, Herremans 1990, Seutin et al. 1992). Despite morphological variation, genetic data from more recent studies (Marthinsen et al. 2008, Mason and Taylor 2015) have failed to find support for any population genetic structure within redpolls, even when examining ~20k single nucleotide polymorphisms (SNPs), suggesting limited to no divergence. As a result, it remains uncertain if morphological differences between redpolls are the result of a balanced polymorphism in an otherwise undifferentiated population, or represent the early stages of speciation.

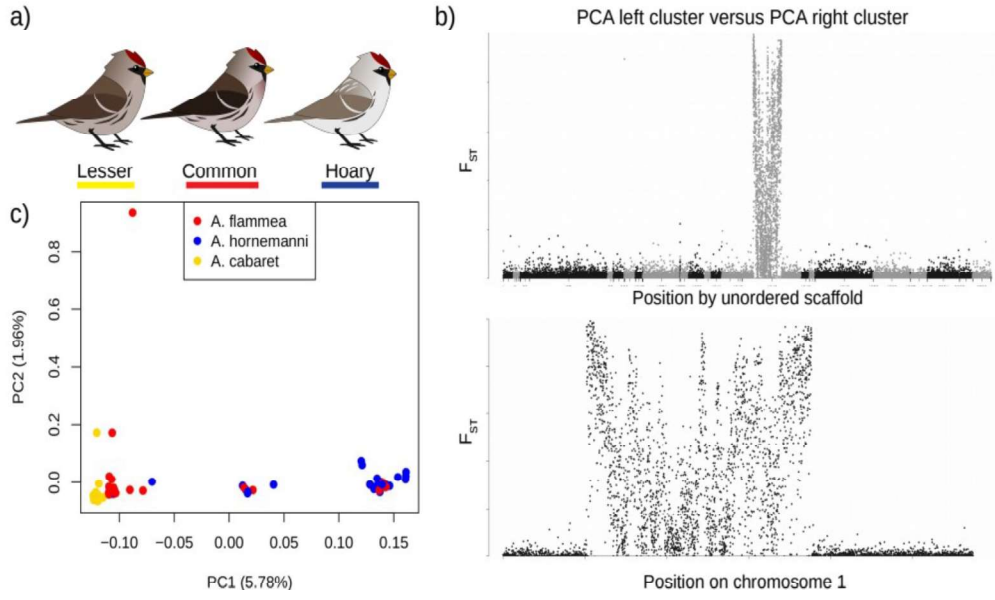


**Figure 1. Redpolls and their genomes. a)** The three currently described redpoll species, including the Lesser Redpoll (*Acanthis cabaret*), the Common Redpoll (*A. flammea*), and the Hoary Redpoll (*A. hornemanni*). **b)** Pairwise FST between divergent redpoll groups, highlighting a single, highly divergent region on chromosome 1. **c)** Principal component analysis of whole genome sequence data, indicating divergent karyotype groups.

Speciation relies on the formation of barriers to reproduction, and traits that differ among closely related species (e.g. coloration) may play an

important role in the formation and maintenance of these barriers (Price 2008). Identifying the genetic basis of such traits is an integral part of understanding the evolutionary mechanisms that either form new species, or maintain variation within species. Structural variants, such as large inversions, may play an important role in generating trait variation, and can have important genetic consequences for a population. For example, chromosomal inversions suppress recombination between inversion types (heterokaryotype individuals), and allow for the co-adaptation of the genes contained within the inversion (Kirkpatrick and Barton 2006). Inversions have the potential to act as isolating barriers if two types of an inversion are incompatible, where a particular combination results in a decrease in fitness, or is lethal (Coyne and Orr 2004). Alternatively, if an inversion does not confer differences in fitness, it may be maintained in a population as a balanced polymorphism (Kim et al. 2017).

The persistence of plumage polymorphisms in redpolls, despite a lack of population genetic structure, suggests redpolls may function as a single population, without fitness differences associated with morphological variation. Preliminary investigations I have conducted using whole genome shotgun data have revealed a region of divergence approximately 55Mb in length along chromosome 1 (Figure 1b). Patterns of population differentiation, measured as $F_{ST}$, between divergent groups identified from principal component analyses (Figure 1c), match patterns of previously reported inversions in other taxa (Kim et al. 2017, Pearse et al. 2019), and may drive balanced plumage polymorphisms despite a lack of population genetic structure. However, in order to characterize this inversion, and better understand the role of this genomic region in generating redpoll diversity, additional long read sequence data is required. Our goal is to sequence long read data from redpolls to better understand the role of this structural variant in generating observed patterns of plumage polymorphism and population structure.

Location: All samples needed for this project have already been obtained. All work will take place in the Taylor lab at the University of Colorado Boulder.

Project Timeline: DNA will be extracted from all samples in February 2020. Sequencing of these individuals will take place in the month immediately following DNA extracting, and will be completed before summer. Sequence analysis will occur in May and June of 2020, for write up and publication in the Fall of 2020.

Methods: Because short read sequence data does not provide adequate information on large-scale structural variants, we propose to sequence long reads from 6 individuals, including 2 individuals from each karyotype, to characterize the inversion, to examine inversion break points, and to examine gene content; we already have short read whole genome sequences, gene expression data, and morphological data for these individuals. We will re-extract high quality DNA from pectoral muscle, and enrich the samples for large (20-70kb) sequence fragments using a Blue Pippin High Pass Cassette (Sage Scientific, USA). We will prepare sequencing libraries using the DNA ligation sequencing kit from Oxford Nanopore Technologies, and sequence long reads using the Oxford Nanopore Technologies minION (Oxford, UK). All high quality reads will be assembled using the program Flye v.2.4.2 (Kolmogorov et al. 2019), and mapped

to a high-quality chromosome-level Brown-capped Rosy-Finch (*Leucosticte australis*; Dovetail Genomics) reference genome using minimap2 v2.16 (Li 2018). We will then examine all contigs mapping to chromosome 1 to more precisely identify the break points of the structural variant.

To investigate gene content and synteny, we will annotate both inversion karyotypes using the Maker pipeline v2.31.10 (Cantarel et al. 2008). We will use these gene models to characterize the genes and the order that they occur in the inversion. Finally, we will use previously obtained expression data to examine the potential for the inversion to play a role in morphological differences. Specifically, we will identify genes related to morphological differences by extracting genes that are differentially expressed between individuals with the lightest and darkest plumage. We will compare the location of differentially expressed genes with respect to the inversion, and will use both the differentially expressed genes, and genes from within the inversion, as target sets for gene ontology enrichment analyses using GOrilla (Eden et al. 2009).

Long read data are integral for characterizing this inversion and will help us better understand the role of structural variants in maintaining balanced polymorphisms and trait variation in nature, as well as clarify a historically confusing group of birds.

References:


Cantarel, B. L., I. Korf, S. M. C. Robb, G. Parra, E. Ross, B. Moore, C. Holt, A. S. Alvarado, and M. Yandell (2008). MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Research 18:188–196. doi: 10.1101/gr.6743907

Clements, J. F., T. S. Schulenberg, M. J. Iliff, D. Roberson, T. A. Fredericks, B. L. Sullivan, and C. L. Wood (2019). The eBird/Clemenets checklist of birds of the world: v2019. [Online.] Available at http://www.birds.cornell.edu/clementschecklist/download.

Coues, E. (1862). A monograph of the genus Aegiothus with descriptions of new species. Proceedings of the academy of Natural Sciences of Philadelphia 13:373–390.

Coyne, J. A., and H. A. Orr (2004). Speciation. In. Sinauer Associates, Sunderland, MA.

Eden, E., R. Navon, I. Steinfeld, D. Lipson, and Z. Yakhini (2009). GOrilla: A tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics 10:1–7. doi: 10.1186/1471-2105-10-48

Herremans, M. (1990). Taxonomy and evolution in redpolls Carduelis flammea-hornemanni□a multivariate study of their biometry. Ardea 78:441–458.

Kim, K. W., C. Bennison, N. Hemmings, L. Brookes, L. L. Hurley, S. C. Griffith, T. Burke, T. R. Birkhead, and J. Slate (2017). A sex-linked supergene controls sperm morphology and swimming speed in a songbird. Nature Ecology and Evolution 1:1168–1176. doi: 10.1038/s41559-017-0235-2

Kirkpatrick, M., and N. Barton (2006). Chromosome inversions, local adaptation and speciation. Genetics 173:419–434. doi: 10.1534/genetics.105.047985

Kolmogorov, M., J. Yuan, Y. Lin, and P. A. Pevzner (2019). Assembly of long, error-prone reads using repeat graphs. Nature Biotechnology 37:540–546. doi: 10.1038/s41587-019-0072-8

Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. Bioinformatics 34:3094–3100. doi: 10.1093/bioinformatics/bty191

Marthinsen, G., L. Wennerberg, and J. T. Lifjeld (2008). Molecular Phylogenetics and Evolution Low support for separate species within the redpoll complex ( Carduelis flammea – hornemanni – cabaret ) from analyses of mtDNA and microsatellite markers. Molecular Phylogenetics and Evolution 47:1005–1017. doi: 10.1016/j.ympev.2008.03.027

Mason, N. A., and S. A. Taylor (2015). Differentially expressed genes match bill morphology and plumage despite largely undifferentiated genomes in a Holarctic songbird. Molecular Ecology 24:3009–3025. doi: 10.1111/mec.13140

Pearse, D. E., N. J. Barson, T. Nome, G. Gao, M. A. Campbell, A. Abadía-Cardoso, E. C. Anderson, D. E. Rundio, T. H. Williams, K. A. Naish, T. Moen, et al. (2019). Sex-dependent dominance maintains migration supergene in rainbow trout. Nature Ecology & Evolution 3:1–12. doi: 10.1038/s41559-019-1044-6

Price, T. (2008). Speciation in Birds. In. Roberts and Company Publishers, Greenwood Village, Colorado.

Seutin, G., P. T. Boag, and L. M. Ratcliffe (1992). Plumage Variability in Redpolls from Churchill , Manitoba. 109:771–785.

Troy, D. M. (1985). A Phenetic Analysis of the Redpolls Carduelis flammea flammea and C . hornemanni exilipes. The Auk 102:82–96.

See the following pages for budget.