MOLECULAR ECOLOGY

Molecular Ecology (2015) **24**, 3009–3025 doi: 10.1111/mec.13140

Differentially expressed genes match bill morphology and plumage despite largely undifferentiated genomes in a Holarctic songbird

NICHOLAS A. MASON*† and SCOTT A. TAYLOR*†

*Department of Ecology and Evolutionary Biology, Cornell University, 215 Tower Rd., Ithaca, NY 14853, USA, †Fuller Evolutionary Biology Program, Laboratory of Ornithology, Cornell University, 159 Sapsucker Woods Road, Ithaca, NY 14850, USA

Abstract

Understanding the patterns and processes that contribute to phenotypic diversity and speciation is a central goal of evolutionary biology. Recently, high-throughput sequencing has provided unprecedented phylogenetic resolution in many lineages that have experienced rapid diversification. The Holarctic redpoll finches (Genus: *Acanthis*) provide an intriguing example of a recent, phenotypically diverse lineage; traditional sequencing and genotyping methods have failed to detect any genetic differences between currently recognized species, despite marked variation in plumage and morphology within the genus. We examined variation among 20 712 anonymous single nucleotide polymorphisms (SNPs) distributed throughout the redpoll genome in combination with 215 825 SNPs within the redpoll transcriptome, gene expression data and ecological niche modelling to evaluate genetic and ecological differentiation among currently recognized species. Expanding upon previous findings, we present evidence of (i) largely undifferentiated genomes among currently recognized species; (ii) substantial niche overlap across the North American *Acanthis* range; and (iii) a strong relationship between polygenic patterns of gene expression and continuous phenotypic variation within a sample of redpolls from North America. The patterns we report may be caused by high levels of ongoing gene flow between polymorphic populations, incomplete lineage sorting accompanying very recent or ongoing divergence, variation in *cis*-regulatory elements, or phenotypic plasticity, but do not support a scenario of prolonged isolation and subsequent secondary contact. Together, these findings highlight ongoing theoretical and computational challenges presented by recent, rapid bouts of phenotypic diversification and provide new insight into the evolutionary dynamics of an intriguing, understudied non-model system.

Keywords: Fringillidae, gene expression, high-throughput sequencing, phenotypic diversity, species limits

Received 4 December 2014; accepted 27 February 2015

Introduction

Inferring the patterns and processes that accompany the generation of phenotypic diversity and new species is an overarching goal of evolutionary biology. In recent years, evolutionary biologists have embraced the notion that different regions of the genome may convey different information about the speciation process (Key 1968; Bazykin 1969; Barton & Hewitt 1981; Rand & Harrison 1989; Harrison & Rand 1989; Harrison 1990; Wu 2001; Nosil & Feder 2012; Seehausen *et al.* 2014). Moreover, the criteria used to delimit species have changed over time—the onset of high-throughput sequencing has provided unprecedented amounts of genetic data that can be used to infer evolutionary history (Lemmon &

Correspondence: Nicholas A. Mason, Fax: (607) 255-8088; E-mail: nicholas.albert.mason@gmail.com

Lemmon 2013; McCormack *et al.* 2013). These novel technologies provide promising tools to study the evolution of phenotypic diversity and speciation, particularly within groups that have experienced recent and rapid diversification, which typically lack coalescence and exhibit incomplete lineage sorting (Maddison & Knowles 2006). Such 'species flocks' present an ongoing challenge for evolutionary biologists to discriminate true speciation events from hybrid swarms and ongoing gene flow, especially when marked phenotypic variation is present.

Reduced-representation approaches, such as double-digest restriction-associated DNA sequencing (ddRAD-Seq) and genotyping by sequencing (GBS), are outperforming traditional Sanger-sequencing methods in their ability to help us understand the evolution of phenotypic variation in young lineages. Hybridization and incomplete lineage sorting are often common in such lineages, which can obscure evolutionary relationships (e.g. Lake Victoria cichlids, Wagner *et al.* 2012; Nicaraguan crater lake cichlids, Elmer *et al.* 2014; *Heliconius* butterflies, Nadeau *et al.* 2013; *Xiphophorus* fishes, Cui *et al.* 2013; Jones *et al.* 2013; wolf-like canids, vonHoldt *et al.* 2011; American oaks, Hipp *et al.* 2014; *Carex*, Escudero *et al.* 2014). Reduced-representation approaches have also generated evidence of strong genetic differentiation in the absence of obvious differences in plumage and morphology in a widely distributed lowland Neotropical bird species (Harvey & Brumfield 2014), provided the resolution necessary to robustly test hypotheses about the generation of hybrid species (e.g. Nice *et al.* 2012) and proved useful for resolving shallow population structure in species of conservation concern (e.g. Larson *et al.* 2013).

In addition to the novel insights gained from restriction enzyme-based approaches, RNA-Seq experiments have revealed that gene expression differences also play an important role in the speciation process (Wolf *et al.* 2010). Changes in gene expression often underlie phenotypic differences among taxa, and differential gene expression may associate with early stages of the speciation process (Pavey *et al.* 2010; Brawand *et al.* 2012; Harrison *et al.* 2012; Filteau *et al.* 2013). Interactions among genotypes and the environment can also affect gene expression, resulting in phenotypic plasticity and polymorphism within a lineage (West-Eberhard 2005). However, studies that evaluate gene expression in addition to variation among putatively neutral, anonymous loci are scarce. Here, we integrate data from reduced-representation libraries with gene expression data to investigate the role of population differentiation and gene expression in generating phenotypic diversity within a widespread genus of songbirds: the redpoll finches (*Acanthis*).

The redpoll finch complex currently includes three species, *Acanthis flammea*, *Acanthis hornemanni* and *Acanthis cabaret*, which are recognized by most authorities (e.g. Clements *et al.* 2014; Fig. 1B). However, between one and six species have been recognized based on plumage and morphology (Coues 1862; Harris *et al.* 1965; Troy 1985; Herremans 1989; Seutin *et al.* 1992; Marthinsen *et al.* 2008). Collectively, redpolls are distributed throughout the Holarctic and individuals that differ in plumage and morphology (putative species) frequently co-occur within the Holarctic range of the genus. The most widespread taxon, *A. flammea*, has darker plumage overall, heavily streaked sides and undertail coverts, and a longer, wider bill (Clement 2010a). By comparison, *A. hornemanni* is lighter with less streaking and a stubby, narrower, conical bill (Clement 2010b). The most recently recognized species, *A. cabaret*, is the smallest of the redpoll taxa and is browner overall (Knox *et al.* 2001; Sangster *et al.* 2002). *Acanthis flammea* and *A. hornemanni* are both widespread and abundant throughout the Holarctic, although *A. hornemanni* is generally found at higher latitudes. *Acanthis cabaret* was historically restricted to the British Isles, but has recently colonized northern mainland Europe and southern Norway.

Although some studies have suggested geographic structuring or multimodal distributions of phenotypic variation within the redpolls (Molau 1985; Seutin *et al.* 1992), other studies have indicated a high prevalence of intermediate phenotypes and overlap in plumage and morphological characters between currently recognized species (Troy 1985; Herremans 1989). Beyond morphological differences, previous studies have cited differences in vocalizations (Molau 1985; Herremans 1989), phenology (Herremans 1989) and physiology (Brooks 1968) as evidence of multiple species within the complex. Most recently, Lifjeld & Bjerke (1996) suggested that *A. cabaret* and *A. flammea* pair assortatively in southeast Norway. However, mixed pairs have also been documented (Harris *et al.* 1965), and the presence of hybrid offspring has been debated (Molau 1985). Despite phenotypic variation among currently recognized species, molecular studies of redpoll populations have consistently failed to document genetic differentiation (restriction fragment length polymorphisms (RFLPs), Marten & Johnson 1986; RFLPs, Seutin *et al.* 1995; mitochondrial control region, Ottvall *et al.* 2002; mitochondrial control region and 10 microsatellites Marthinsen *et al.* 2008). From a biogeographic perspective, this lack of genetic variation is unusual; most Holarctic birds demonstrate some degree of phylogeographic structuring among temperate ecoregions or

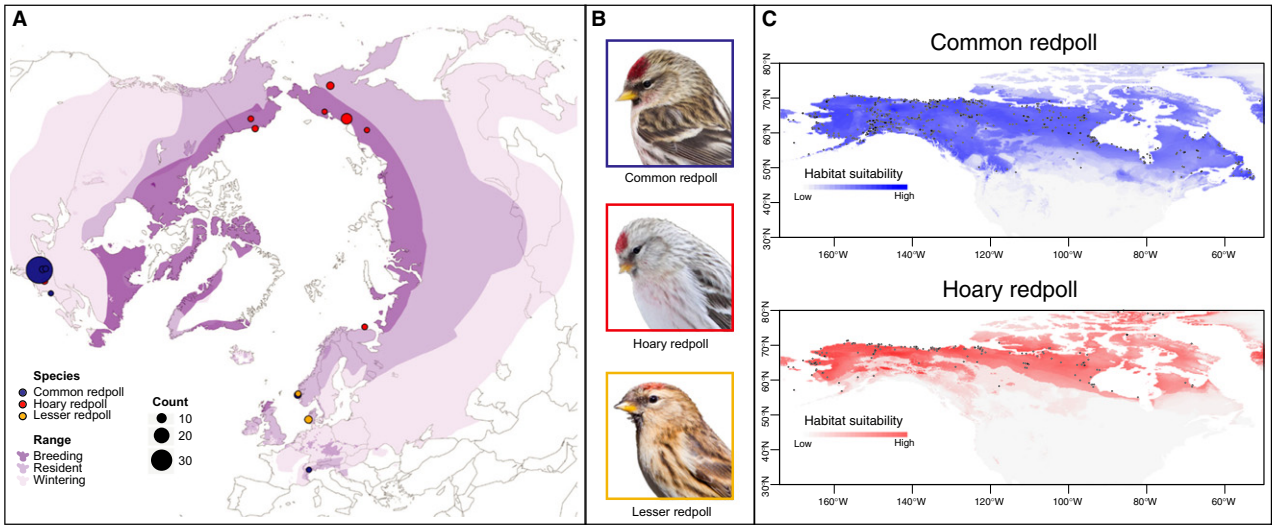

Fig. 1 The redpoll system. (A) Combined Holarctic distribution of redpoll finches. Dark purple indicates breeding grounds, intermediate purple indicates resident status, and light purple indicates wintering distributions. Sampling sites are indicated with coloured circles: the size of the circle corresponds to the number of individuals sampled from each site. (B) Representative phenotypes of common, hoary and lesser redpolls. Note differences in plumage coloration, patterning and bill morphology. (C) Ecological niche models (ENMs) constructed for common and hoary redpolls in North America. Darker colours indicate more suitable habitat. Occurrence data used to create niche models are shown with black dots. ENMs suggest considerable overlap in suitable abiotic conditions between hoary and common redpolls. Nonetheless, hoary redpolls prefer higher latitudes, while common redpolls are more widespread.

between continents (e.g. Questiau *et al.* 1998; Drovetski *et al.* 2004, 2009).

The paucity of genetic differentiation within the redpoll complex, despite marked phenotypic variation across a Holarctic distribution, could be the result of multiple evolutionary scenarios (Marthinsen *et al.* 2008): redpolls may be comprised of (i) a single, undifferentiated gene pool that exhibits phenotypic polymorphism, in which phenotypic differences reflect locally adapted demes or neutral phenotypic variation within a single metapopulation; (ii) multiple gene pools that have recently diverged, in which incomplete lineage sorting has hindered the capacity of previous studies to differentiate populations or species; or (iii) multiple divergent gene pools that are actively exchanging genes through hybridization and introgression via secondary contact.

In this study, we implement high-throughput sequencing to evaluate these hypotheses by examining genome-wide variation in anonymous loci among redpolls sampled from different regions of the Holarctic. We also assess variation among transcriptome sequence data and gene expression in a subset of North American redpolls that span the phenotypic continuum described above. Finally, we use breeding season occurrence records to generate ecological niche models (ENMs) that characterize differences in suitable abiotic conditions between North American *A. flammea* and *A. hornemanni*.

Materials and methods

Sample collections and phenotyping

Molecular analyses were based on 77 individuals, including representatives of all three redpoll species currently recognized by most authorities (e.g. Clements *et al.* 2014), which were from different regions of their current distribution (Fig. 1, Table S1, Supporting information). Based on recently published phylogenies of the family Fringillidae (Zuccon *et al.* 2012), we included two white-winged crossbill (*Loxia leucoptera*) individuals as an out-group in our analyses. Because the main goal of this study was to assess genetic differentiation between redpolls with different plumage and morphology characteristics (i.e. putative species), our geographic sampling was not exhaustive from a phylogeographic perspective and we did not include representatives from all currently recognized redpoll subspecies. For this component of our study, we relied on the classifications of collectors and museum curators to assign individuals to one of the three currently recognized species.

We collected 10 of the 77 redpolls included in this study on the same day at the same wintering locality (Cortland, Cortland County, NY, USA; 42.6°N, 76.2°W; nine males and one female). These individuals were collected because they represented the broadest phenotypic variation possible within the wintering flock. The

flock remained at this location for over 3 weeks before collection. Therefore, individuals experienced similar environmental conditions and foraging opportunities that approximate a common garden setting. This shared experience should have reduced environmentally induced differences in gene expression between flock members; however, we cannot completely rule out differences in microclimate or diet.

Rather than binning these individuals into putative species based on plumage characteristics and bill shape, which are known to vary continuously (Troy 1985), we measured multiple morphological characters for each individual (Table S2, Supporting information). We quantified the amount of streaking on the undertail coverts and rump of each individual by taking digital photographs that were subsequently measured with IMAGEJ 1.48v (Abràmoff *et al.* 2004). We took four measurements of beak shape (width, depth, culmen length, mandible length) for each individual using digital callipers. Bill and plumage measurements were then incorporated into a principal components analysis (PCA) to obtain multivariate dimensions of phenotypic variation (see Fig. S1, Supporting information for loadings, and Fig. S2, Supporting information for PCA scores). PCA scores were then used to assess statistical associations between phenotypic variation and multiple indices of genetic variation.

In addition to collecting genomic DNA from these 10 individuals, we also preserved separate samples of whole brain, liver and muscle in RNAlater within 25 min post-mortem for RNA-Seq data generation and gene expression analyses. Specimens were processed in the order in which they were caught, meaning that some individuals were held captive longer than others before collecting tissues. Genomic DNA and RNA samples were subsequently stored at −80 °C until library preparation.

ddRAD-Seq library preparation and sequencing

We extracted genomic DNA from each sample using Qiagen® DNeasy kits (tissue protocol; Qiagen, Valencia, CA, USA), eluted the DNA in water, concentrated each sample using a vacuum centrifuge and determined the final concentration of each extraction using Qubit Fluorometric Calibration (QFC; Invitrogen, Carlsbad, CA, USA). DNA extractions are archived at the Cornell Lab of Ornithology (Ithaca, NY, USA).

We prepared ddRAD-Seq libraries using a modified version of the protocol outlined in Peterson *et al.* (2012). Following a standardizing dilution (all genomic DNA ~30 ng/μL), we plated the samples and digested each with the restriction enzymes *Sbf*I and *Msp*I while ligating P1 (barcode) and P2 adaptor primers using 19 unique barcodes for each of four subsequent index groups (a total of 76 unique identifiers—the DNA from one sample was excluded due to low quality). Each digestion reaction contained 300 ng genomic DNA, 3 μL 10× CutSmart buffer, 1 μL of 250 nM P1, 1 μL of 25 μM P2, 3 μL 10 mM ATP, 0.75 μL (15 U) each of 20 U/μL *Sbf*I-*HF* and *Msp*I, 0.75 μL of 400 U/μL T4 DNA ligase, and water to a total of 30 μL. Next, samples were incubated at 37 °C for 30 min followed by 1 h at 20 °C, pooled in groups of 19 and cleaned with 1.5× volumes of AMPure beads (Beckman Coulter Inc) and two washes of 70% ethanol. The pooled samples were then eluted into 30 μL of Qiagen EB buffer and quantified using QFC. For each of the four index groups (index primers 6, 12, 1, 2), we set up six replicate PCRs containing 20 ng DNA, 12.5 μL 2× Phusion MM, 1.25 μL of 5 μM P1, 1.25 μL of 5 μM index primer, and water to a total reaction volume of 25 μL. The PCR temperature profile included a 30-s incubation at 98 °C followed by 16 cycles of 98 °C for 5 s, 60 °C for 25 s and 72 °C for 10 s with a final extension step of 5 min at 72 °C. The six replicate PCRs were pooled within each index group and visualized on a 1% agarose gel. We size-selected (200–1000 bp) each index group using a two-step AMPure cleanup with 6.5% PEG/1.2 M NaCl, 25% PEG/1.2 M NaCl and 11.5% PEG/1.2 M NaCl. Final elutions of each index group were analysed using QFC and an Agilent Bioanalyzer 2100 (Agilent, Santa Clara, CA, USA). Finally, we diluted each index group to 2 nM, combined all four in equal proportions and sequenced the total redpoll ddRAD-Seq library on two lanes of an Illumina HiSeq 2000 [100 base pair (bp), paired end] at the Cornell University Life Sciences Core Laboratories Center (Ithaca, NY, USA). The crossbill ddRAD-Seq library, consisting of two individuals, was prepared using identical protocols and was sequenced on 5.2% of one shared lane of an Illumina HiSeq 2000 (100 bp, single end) at the Cornell University Life Sciences Core Laboratories Center.

RNA-Seq library preparation and sequencing

We combined ~15 mg of homogenized liver, pectoral muscle and brain from each of the 10 individuals that were collected on the same day at the same wintering locality to total 45 mg of tissue per individual library. We recognize that gene expression probably varies across the three tissue types we pooled and that the tissues we chose are not ideal for detecting gene expression differences related to plumage or facial morphology (Abzhanov *et al.* 2004; Ekblom *et al.* 2012; Poelstra *et al.* 2014). Nevertheless, maintenance of plumage patterning and facial morphology (e.g. melanin deposition and bill growth patterns) should be preserved throughout

the life of an individual, and there is the possibility that these gene expression profiles may be detectable in pooled tissue samples (see Results). One possible drawback of pooling tissues is that we are unable to detect organ-specific differential gene expression among redpolls; however, the goal of our RNA-Seq experiment was to detect possible species-level differences and find candidate genes worthy of further exploration under controlled conditions. Future experiments in this system will take a tissue-specific approach to target more relevant tissue types and developmental stages (e.g. Abzhanov *et al.* 2004; Ekblom *et al.* 2012; Poelstra *et al.* 2014).

Following collection, we used a TissueRuptor (Qiagen) to homogenize each individual tissue pool and followed the 'standard' mRNA extraction protocol as detailed in the Dynabeads® mRNA DIRECT™ kit (Invitrogen). To remove as much rRNA as possible from our extraction before constructing complimentary DNA (cDNA) libraries, we performed the mRNA extraction protocol twice. We converted mRNA into cDNA libraries using the NEBNext RNA First Strand Synthesis Module (New England BioLabs, Ipswich, MA, USA). We then performed second-strand cDNA synthesis, end repair, dA-tailing and adaptor ligation for each individual cDNA library. Following adapter ligation and library purification, we performed 12 cycles of the 'denaturation annealing extension' step during the index PCR. Prior to pooling multiplexed individuals, we assessed the quality and quantity of cDNA using QFC and an Agilent Bioanalyzer 2100. Multiplexed individual cDNA libraries were pooled in equimolar ratio and sequenced on a single lane using 100-bp single-end reads on the Illumina HiSeq 2000 at the Cornell Core Laboratories Center. Raw, demultiplexed reads are available through the NCBI Short Read Archive (SRP052607).

Transcriptome assembly and gene expression profiling

Barcoded RNA-Seq reads were demultiplexed, filtered and trimmed prior to assembly. We used TRIMMOMATIC v0.27 (Lohse *et al.* 2012) to remove low-quality reads that dropped below a Phred-scale quality score of 20 or included contamination from Illumina adapters. Filtered reads were then loaded into TRINITY r2013-02-25 (Grabherr *et al.* 2011; Haas *et al.* 2013) to assemble a de novo reference transcriptome using all 10 individual cDNA libraries, including individuals of both *Acanthis flammea* and *Acanthis hornemanni*.

We performed transcript quantification of the de novo assembly with RNA-Seq by Expectation Maximization (RSEM v.1.2.3; Li & Dewey 2011), which estimates transcript abundance by aligning filtered and trimmed reads to the reference transcriptome. Before continuing with downstream analyses, we tested a range of different transcript abundance cut-off values to remove underrepresented contigs from the de novo assembly (Fig. S3, Supporting information). We removed contigs that failed to meet a 1.0 Transcripts Per Million threshold and selected the longest isoform for each component. After selecting the longest isoform to represent each contig, our final Trinity assembly included 30 357 contigs with an N50 of 2715 bp. The transcriptome assembly and raw RNA-Seq reads can be accessed via the NCBI Transcriptome Shotgun Assembly Database (BioProject number PRJNA256306).

To compare gene expression profiles among the 10 individuals that we collected and phenotyped, we aligned individual RNA-Seq libraries back to the de novo assembly with BOWTIE (Langmead *et al.* 2009) and used RSEM to estimate read abundance for each gene. We applied Trimmed Mean of M-values normalization to obtain Fragments Per Kilobase of exon per Million fragments mapped (FPKM) values for each sample and contig, which were then log-transformed (Robinson *et al.* 2009). We applied multidimensional scaling to FPKM counts to generate multivariate dimensions of differential gene expression among individuals using the package LIMMA (Smyth 2005; Ritchie *et al.* 2015) within the R programming environment (R Core Development Team 2014).

To identify differentially expressed genes that are potentially associated with phenotypic variation in redpolls, we used principal component scores of plumage and morphology as the response variable in a generalized linear model (GLM) with normalized, log-transformed FPKM read counts as predictor variables in the LIMMA package (Smyth 2005; Ritchie *et al.* 2015) and corrected for multiple hypothesis testing by applying the false discovery rate method (Benjamini & Hochberg 1995). Sampling order was included as a random factor in the GLM. We searched for matches between each transcript in our assembly and the NCBI nonredundant protein database using BLAST+ (Camacho *et al.* 2009). For each hit that met a threshold *e*-value < 1e-5, we obtained corresponding Gene Ontology terms using BLAST2GO (Conesa *et al.* 2005) to assess the functional implications of any differentially expressed genes.

Locus assembly and SNP calling

We concatenated ddRAD-Seq reads from the forward and reverse direction for downward locus assembly with STACKS v1.20 (Catchen *et al.* 2011, 2013). We used fastx_trimmer (http://hannonlab.cshl.edu/fastx_toolkit/) to trim all reads to an equal length of 94 bp and then filtered out reads with any bases that fell below a

Phred score of 10. We also trimmed out any reads for which ≥5% of bases had a Phred score below 20. We then separated multiplexed libraries using the process_radtags function from the STACKS pipeline (Catchen *et al.* 2013). The final filtered, trimmed and demultiplexed data set contained 365 000 000 reads.

We pooled reads from all redpolls to perform de novo locus assembly for redpolls only using the denovo_map.pl script, which executes ustacks, cstacks and sstacks in succession and comes bundled with STACKS (Catchen *et al.* 2013). In brief, STACKS groups identical reads based on sequence similarity to form 'stacks', which can then be combined to form putative loci. We required a minimum of five reads for stack depth (-m), allowed five SNPs between any two stacks at a locus (-M) and five SNPs between any two loci when building catalogues (-n). These parameter settings performed well in a comparison of library assembly pipelines (Mastretta-Yanes *et al.* 2015). We allowed 20% missing data for each locus and extracted one locus per SNP using the –write_single_snp flag when running the populations program within STACKS. One individual had to be dropped from the ddRAD-Seq assembly pipeline due to poor coverage; therefore, the finalized ddRAD set included 76 individuals, including 9 of 10 individuals that comprise the RNA-Seq portion of this study.

Crossbill raw sequencing reads were processed in the same manner as redpolls. The final filtered, trimmed and demultiplexed crossbill data set contained 151 511 reads. We pooled all reads from both crossbills and redpolls to perform de novo crossbill and redpoll locus assembly using the denovo_map.pl script and used the same STACKS settings detailed above.

We also identified a separate panel of SNPs from the de novo transcriptome and 10 individual RNA-Seq libraries. We generated an index from our transcriptome and aligned each individual library to the reference using BWA under default settings (Li & Durbin 2009). We called SNPs from indexed alignments using the UnifiedGenotyper tool within Genome Analysis Toolkit under default settings (GATK; DePristo *et al.* 2011). As part of the SNP calling process, we filtered out sites with Phred quality scores <30 and filtered by mean depth, allele frequency and call rate and applied the BadCigarFilter using VCFTOOLS (version 3.0; Danecek *et al.* 2011), which removes malformed reads that start with spurious deletions. We retained a total of 215 825 out of 784 141 possible SNPs after filtering.

Population genetic analyses

We used the Bayesian clustering program STRUCTURE v 2.3.4 (Pritchard *et al.* 2000) to evaluate genetic differentiation among redpolls. We analysed two sets of ddRAD-Seq SNPs using the same analytical pipeline: (i) loci assembled from only redpoll data and (ii) loci assembled with data from redpolls and the crossbill out-group. For the redpoll data set, we ran three replicate analyses for 10 000 generations following 10 000 generations of burn-in, using the 'admixture' model across a range of K values from 1 to 5 (three replicates each), which were then averaged for population assignment scores. Because there were three putative species in this analysis, we paid specific attention to results from runs where the a priori constraint on the number of population clusters was $K = 3$ (redpolls only). Results from STRUCTURE were analysed using the Evanno *et al.* (2005) method in STRUCTURE HARVESTER (Earl & vonHoldt 2011). For the redpoll + crossbill data set, we ran STRUCTURE as above with an a priori constraint on the number of population clusters $K = 4$ (redpolls + crossbills). We did this to ensure that our SNP data could differentiate redpolls from the out-group taxon.

To corroborate our Bayesian clustering analyses, we performed principal component analyses (PCA) on the same two sets of loci using ADEGENET v1.4 (Jombart 2008; Jombart & Ahmed 2011) and performed an analysis of molecular variance (AMOVA; Excoffier *et al.* 1992) using PEGAS (Paradis 2010) to examine how genetic variation is partitioned among currently recognized species within the redpoll complex. We also tested for isolation by distance among redpolls through a partial Mantel test using the R package ADE4 (Chessel *et al.* 2004). We used BAYESCAN v2.1 (Foll & Gaggiotti 2008) with default settings to identify any outlier loci that are highly differentiated between currently recognized species and used a Friedman test to examine differences in observed heterozygosity between putative species. We also assessed population structure among the 10 individuals we collected for our RNA-Seq experiment (which span the phenotypic continuum—see Results) with the panel of 215 825 SNPs called from our de novo transcriptome by running a PCA using ADEGENET (Jombart 2008; Jombart & Ahmed 2011).

We performed Bayes factor delimitation (Grummer *et al.* 2014) using a subsample of SNPs from the ddRAD-Seq data set in combination with SNAPP (Bryant *et al.* 2012), which is a module of BEAST 2 (Bouckaert *et al.* 2014), to assess support for lumping or splitting redpoll species within a multispecies coalescent framework (Leaché *et al.* 2014). Due to computational constraints, we randomly sampled six individuals from each species and combined these with the two out-group individuals to construct two data sets. For these 20 individuals, one data set included 35 loci with no missing data; the other matrix included 200 randomly sampled loci with no missing data for the two out-group individuals and a maximum of 20% missing data

among redpolls. Although these loci represent a small portion entire genome, similar numbers of loci have successfully characterized general patterns of coalescence in other empirical studies (Grummer *et al.* 2014; Leaché *et al.* 2014). We conducted path sampling with 12 steps (100 000 MCMC steps; 100 000 pre-burn-in steps) to estimate the marginal likelihoods of models where redpolls are split or lumped (Leaché *et al.* 2014), which were later compared using Bayes factors (Kass & Raftery 1995).

Generalized linear models of phenotypic and genetic variation

To examine statistical associations between phenotypic variation (within the sample of 10 individuals that we collected in Cortland, NY) and different aspects of genetic variation, we constructed GLMs with either PC1 or PC2 from the plumage + bill morphology PCA as response variables and used various indices of genetic variation as predictor variables. First, we assessed whether phenotypic variation was associated with patterns of gene expression by including multidimensional scaling log-fold change dimension 1 and 2 scores (Ritchie *et al.* 2015) as predictor variables and PC1 and PC2 scores from the plumage + bill morphology PCA as response variables, respectively. We included processing order as a random factor in both of these models to account for the potentially confounding effect of time spent in captivity prior to tissue collection. To determine whether phenotypes were associated with variation in anonymous SNP loci, we quantified associations between PC1 and PC2 scores from the plumage + bill morphology PCA and PC1 and PC2 scores from the ddRAD-Seq PCA for the 10 individuals we collected. Because one of the individuals we collected was dropped from the ddRAD-Seq data set due to poor coverage, these comparisons were restricted to nine data points. Finally, we also assessed whether there were relationships between the PC1 and PC2 scores from the SNPs generated from the de novo transcriptome and PC1 and PC2 of the plumage + bill morphology PCA.

Niche modelling of Acanthis flammea and Acanthis hornemanni in North America

To compare the abiotic niches that constitute the breeding ranges of *A. flammea* and *A. hornemanni* in North America (*Acanthis cabaret* was not included here due to the difficulty of niche modelling for island populations and its limited range), we first gathered occurrence records through the Global Biodiversity Information Facility (GBIF; http://www.gbif.org/) via the DISMO package (Hijmans *et al.* 2013). We filtered occurrence data so that only records from the breeding season (i.e. June and July; Seutin *et al.* 1991) were included in our analyses. Because sampling bias is common among occurrence records (Hijmans *et al.* 2000), we subsampled our data set to include only one occurrence record for every cell within a 0.5° latitude × 0.5° longitude grid of North America. This resulted in 545 breeding occurrence records of *A. flammea* and 159 breeding occurrence records of *A. hornemanni* (Table S3, Supporting information).

After gathering occurrence records, we used the BioClim data set (Hijmans *et al.* 2005) with a resolution of 2.5 arc-minutes to extract 19 abiotic variables associated with each set of coordinates. Using these data, we generated ENMs using MAXENT 3.3 (Phillips *et al.* 2006; Elith *et al.* 2011), which performs well compared to alternative algorithms for generating ENMs (Elith *et al.* 2006). We generated pseudoabsence data by extracting bioclimatic data associated with 500 random points within North America (extent in unprojected coordinates: latitude 30.0° to 80.0° and longitude −50.0° to −174.0°; Fig. 1C). This extent is reasonable given that redpolls are highly vagile and have been recently reported as far south as Julian, CA (33.08°N, 116.60°W), making this geographic extent an approproiate approximation of the 'accessible' niche space (Barve *et al.* 2011). We used the default settings within Maxent and partitioned 20% of the occurrence data from both species to validate the performance of our models via *k*-fold cross-validation (Hastie *et al.* 2001). More specifically, we calculated the area under the curve (AUC) of the receiver operating characteristic score, which indicates the ability of our model to predict a designated subset of our occurrence data. This procedure was replicated 50 times to produce a distribution of AUC scores for each ENM. An AUC score of 1 indicates a model that perfectly predicts the testing data (i.e. species occurrences), whereas an AUC of 0.5 indicates a model that has no predictive power. Thus, we accepted a given ENM if the median AUC score across *k*-fold replicates was greater than an arbitrary cut-off value of 0.80.

We implemented two different statistical tests to compare the projected models of *A. flammea* and *A. hornemanni* in North America using the PHYLOCLIM package (Heibl & Calenge 2013). Following the methodology provided by Warren *et al.* (2008), we evaluated 'niche identity' and 'niche similarity' by calculating a modified version of Hellinger's distance (*I*; Legendre & Gallagher 2001), as well as Schoener's *D* (Schoener 1968), between the projected ENMs of the two species (Warren *et al.* 2010). To assess niche identity, we compared these niche similarity values to a null distribution of similarity measures that was built by comparing pseu-

doreplicated ENMs based on 50 pooled, randomized occurrence data from both species (Graham *et al.* 2004; Warren *et al.* 2008). This procedure tests niche conservatism in the strictest sense and determines whether the environmental tolerances for the two species are identical. We also quantified niche similarity by comparing similarity values between the projected ENMs of *A. flammea* and *A. hornemanni* to distributions of similarity values obtained by comparing ENMs built from occurrence data of each species to ENMs constructed with random points sampled throughout the geographic extent (i.e. extent pictured in Fig. 1C; Peterson *et al.* 1999; Warren *et al.* 2008). This procedure was repeated 50 times to generate null distributions of similarity values and determine whether the observed overlap between the niches of *A. flammea* and *A. hornemanni* is simply the product of regional similarities or the niche models of the two species are more similar or different than would be expected by chance.

Results

Population genetics using anonymous loci

Filtering and locus assembly protocols from the ddRAD data generated a redpoll data set consisting of 20 712 SNPs and a redpoll + out-group data set consisting of 1587 SNPs. A Friedman test revealed differences in observed heterozygosity between species ($\chi^2 = 4679.26$, d.f. = 2, $P < 0.05$; Fig. S4, Supporting information). Specifically, Wilcoxon signed-rank tests with Bonferroni correction indicated that observed heterozygosity was lower in *Acanthis cabaret* compared to *Acanthis flammea* ($P < 0.05$) and *Acanthis hornemanni* ($P < 0.05$). Observed heterozygosity did not differ between *A. hornemanni* and *A. flammea* ($P > 0.05$). These differences probably reflect variation in sample sizes because far fewer *A. cabaret* samples (6) were included in this study.

At $K = 3$ (the putative number of species in the redpoll analysis), results from STRUCTURE indicated that all redpolls clustered together (Fig. 2A) and the genetic PCA revealed low levels of differentiation among currently recognized redpoll species (Fig. 2B). By comparing changes in likelihood scores among different settings of the K parameter, we identified $K = 2$ as the preferred setting via the Evanno *et al.* (2005) method (Table S4, Supporting information); however, at every setting of K, all redpolls were assigned to the same population cluster. Our path-sampling analysis of different species delimitation models conducted using Bayes factor delimitation with SNAPP favoured a model with redpolls lumped as a single species within a coalescent framework (Fig. 2C); both the 35 locus data set with complete sampling (Bayes factor = 36.80) and the 200 locus data set (Bayes factor = 15.22) supported this finding.

When analysed with the out-group and a smaller SNP data set (1587 loci), redpolls were easily differentiated from the out-group, but exhibited low genetic

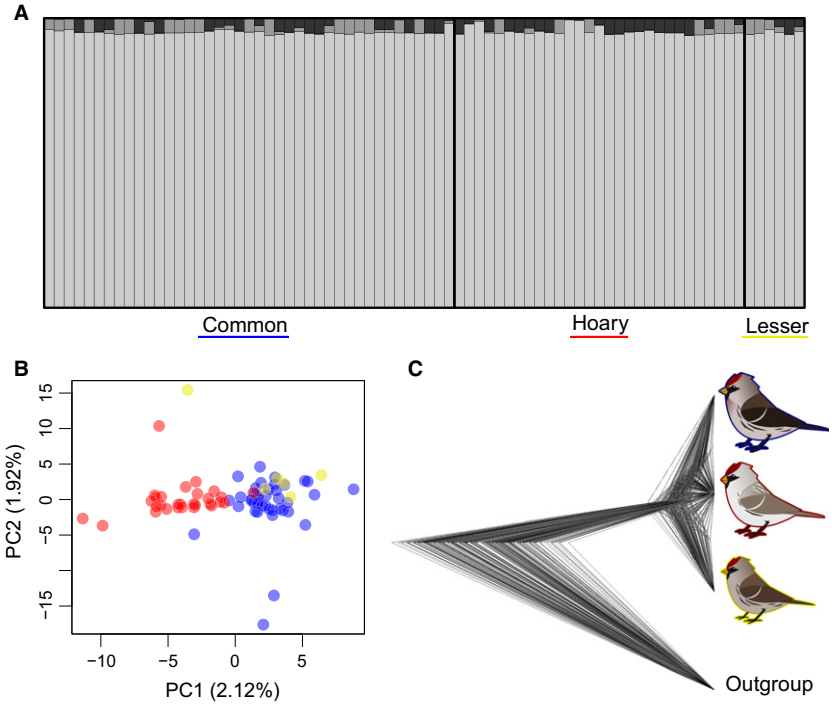

Fig. 2 Redpoll population genetic analyses. (A) Bayesian assignment probabilities from STRUCTURE showing lack of population clustering among currently recognized redpoll species using 20 721 SNPs. (B) Genetic PCA plot indicating weak population structure among currently recognized species of redpolls. Common redpoll is represented with blue, hoary redpoll is represented with red, and lesser redpoll is represented with yellow dots. (C) SNAPP tree using 1587 SNPs for common, hoary and lesser redpoll, and white-winged crossbill (grey). Bayes factor delimitation strongly favoured lumping redpolls into a single species (Bayes factor = 36.80).

differentiation among currently recognized species or geographically isolated populations (Fig. S5A, Supporting information); the genetic PCA (Fig. S5B, Supporting information) also separated the out-group from redpolls but did not appreciably separate currently recognized redpoll species. Similarly, a PCA of the 215 825 locus data set generated from our RNA-Seq experiment (10 individuals that spanned the phenotypic continuum—see Table S2, Supporting information) did not differentiate currently described species (Fig. S6, Supporting information).

An AMOVA of the ddRAD-Seq SNP panel revealed that the overwhelming majority of genetic variation is partitioned within species (98.11%; Table 1) rather than between species (1.89%; Table 1). We also found that redpolls exhibit isolation by distance ($r = 0.12$, $P = 0.04$): individuals are more closely related to geographically proximate individuals, regardless of their phenotype. Our BAYESCAN analysis did not identify any outlier loci that were highly divergent between currently recognized species (Fig. S7, Supporting information).

Associations between phenotypes, differentially expressed genes and anonymous SNPs

We found strong associations between multidimensional scaling scores of differential gene expression and both principal component scores of the plumage + morphology PCA. The leading log-fold change dimension (LLFC) 1 of multidimensional scaling space was correlated with PC1 [$\beta = 2.882 \pm 0.307$ (SE), $P = 3.2e\text{-}05$, Table 2, Fig. 3A], and LLFC 2 was correlated with PC2 ($\beta = 0.059 \pm 0.068$, $P = 0.04$, Table 2, Fig. 3B). In contrast, LLFC 1 was not correlated with PC1 of SNP variation from the ddRAD-Seq data ($\beta = -0.229 \pm 0.495$, $P = 0.66$, Table 2, Fig. 3C) and LLFC 2 was not correlated with PC2 of ddRAD-Seq SNP variation ($\beta = 0.144 \pm 0.232$, $P = 0.55$, Table 2, Fig. 3D). Finally, PC1 and PC2 scores from the panel of SNPs called from the de novo transcriptome were not correlated with PC1 and PC2 scores, respectively, from the plumage + morphology PCA ($\beta = -0.812 \pm 0.726$, $P = 0.30$; $\beta = -0.005 \pm 0.007$, $P = 0.50$, Table 2, Fig. 3E).

Differential gene expression and gene annotations

We used BLAST+ (Camacho *et al*. 2009) to align our contigs against the UniProtKB/Swiss-Prot protein database and found that 29 850 of our contigs (24.5%) had a significant BLAST hit with an *e*-value of ≤1e-5. Our GLMs that quantified associations between phenotype and FPKM values did not recover any genes with a false discovery rate below 0.05 (Table S5, Supporting information). However, we did identify a number of candidate genes related to morphology that are worthy of further study under controlled conditions; for completeness, we present the 100 contigs with BLAST hits that were most strongly associated with phenotypic variation among redpolls (Table S5, Supporting information).

Niche modelling

Acanthis flammea and *A. hornemanni* demonstrate considerable overlap in suitable habitat in North America, although *A. hornemanni* does seem to prefer higher latitudes (Fig. 1C). The ENM for *A. flammea* (median AUC = 0.88, interquartile range = 0.87–0.90) predicted highly suitable habitat across much of northern Canada and Alaska, which reflects its widespread distribution throughout North America. In contrast, *A. hornemanni* (median AUC = 0.94, interquartile range = 0.93–0.95) prefers abiotic conditions associated with higher latitudes throughout Canada and Alaska. The variable that contributed most to the ENMs of both species was the maximum temperature of the warmest month (56.2% and 49.4% for *A. flammea* and *A. hornemanni*, respectively; Fig. S8, Supporting information). However, the response curves for certain BioClim variables differed substantially between the two species. For example, the ENM that we constructed for *A. hornemanni* indicated a higher probability of occurrence, compared to the ENM for *A. flammea*, among localities where the maximum temperature of the warmest month and the annual mean temperatures were lower (Fig. S9, Supporting information).

The niche equivalency test indicated that *A. flammea* and *A. hornemanni* do not occupy identical niches ($D = 0.579$, $P < 0.0001$; $I = 0.824$, $P < 0.0001$; Fig. S10,

Table 1 Results from analysis of molecular variance (AMOVA) using 20 712 SNPs, including the degrees of freedom (d.f.), the sum of squares (SS), mean squared deviation (MSD), variance (σ^2), the amount of total variation explained by hierarchical level and the estimate of population differentiation (Φ_{ST})

	d.f.	SS	MS	σ^2	% of total variation	Φ_{ST}
Between species	2	2143.93	1071.97	14.66	1.89	−0.01
Within species	73	55597.32	761.61	761.61	98.11	
Total	75	57741.25	769.88			

Table 2 Results from generalized linear models built to assess statistical associations between plumage and morphology principal component scores, multidimensional scaling leading log-fold change scores from RNA-Seq data, SNPs called from the ddRAD-Seq data set, and the transcriptome. Each model is separated with a horizontal rule: models with multidimensional scaling leading log-fold changes also included processing order as a possible predictor variable. The magnitude of each effect score with their corresponding standard error (SE) values is shown. All linear model terms that have a *P*-value lower than 0.05 are indicated with an asterisk

Response	Predictor	β ± SE	*T* value	*P*-value
PC1	mdsPC1	2.882 ± 0.307	9.399	3.20e-05*
PC1	Processing order	0.059 ± 0.068	0.865	0.416
PC2	mdsPC2	1.200 ± 0.477	2.517	0.04*
PC2	Processing order	0.151 ± 0.08	1.894	0.1
PC1	ddRAD-Seq SNP PC1	−0.229 ± 0.495	−0.463	0.657
PC2	ddRAD-Seq SNP PC2	0.144 ± 0.232	0.621	0.554
PC1	RNA-Seq SNP PC1	−0.812 ± 0.726	−1.119	0.296
PC2	RNA-Seq SNP PC2	−0.005 ± 0.007	−0.698	0.505

Supporting information). Yet, our background similarity test indicated that the abiotic conditions that characterize the distribution of *A. hornemanni* are more similar to those of *A. flammea* than would be expected based on the availability of habitat in North America (D = 0.579, 95% CI = 0.371–0.402; I = 0.824, 95% CI = 0.661–0.690; Fig. S10, Supporting information). Similarly, the observed similarity indices were higher than the confidence interval of the null distribution constructed using actual *A. flammea* occurrence data and randomly generated *A. hornemanni* data (D = 0.579, 95% CI = 0.502–0.558; I = 0.824, 95% CI = 0.773–0.813; Fig. S10, Supporting information). Thus, while the abiotic niches of *A. flammea* and *A. hornemanni* are not completely identical, they are more similar than comparisons of either species' ENMs with null models generated from random background points throughout North America.

Discussion

Our findings expand upon previous studies of the evolutionary dynamics of redpolls that have relied on traditional molecular markers, such as mtDNA and microsatellites (Marten & Johnson 1986; Seutin *et al.* 1995; Ottvall *et al.* 2002; Marthinsen *et al.* 2008). Genome-wide panels of SNPs generated using ddRAD-Seq and a de novo transcriptome both support the hypothesis that redpolls comprise a single, Holarctic gene pool with remarkably little genetic differentiation between currently recognized species. The absence of outlier loci among a panel of tens of thousands of SNPs suggests that currently recognized redpoll species share very recent common ancestry; whole-genome sequencing will further clarify patterns of genomic differentiation between redpolls. Intriguingly, we found novel differences in gene expression that are correlated with redpoll phenotypes, suggesting that gene expression might play an important role in generating phenotypic diversity among redpolls. Finally, we demonstrated that *Acanthis flammea* and *Acanthis hornemanni* exhibit substantial overlap in suitable habitat in North America, with *A. hornemanni* typically occurring at higher latitudes than *A. flammea*.

Evolutionary history of redpolls

The consistent lack of genetic differentiation in two large panels of anonymous loci, including hundreds of thousands of SNPs within the transcriptome, is surprising given the geographic and phenotypic breadth of sampling included in this study. Recently, reduced-representation genomic approaches, like those used here, have provided unprecedented resolution in other lineages that exhibit marked phenotypic diversity on recent evolutionary timescales (e.g. Lake Victoria cichlids, Wagner *et al.* 2012; Nicaraguan crater lake cichlids, Elmer *et al.* 2014; American oaks, Hipp *et al.* 2014). Given the low levels of genome-wide genetic differentiation we detected in redpolls, it appears unlikely that currently recognized species underwent long periods of allopatric divergence and have since come back into secondary contact. Rather, our findings suggest that *A. flammea*, *A. hornemanni* and *Acanthis cabaret* have a predominantly shared evolutionary history and currently comprise a single gene pool distributed throughout the Holarctic, which may be undergoing contemporary differentiation via ecological selection.

Our ENMs demonstrate that the abiotic conditions that characterize *A. hornemanni* and *A. flammea* differ (i.e. *A. hornemanni* tend to occur at higher latitudes), but

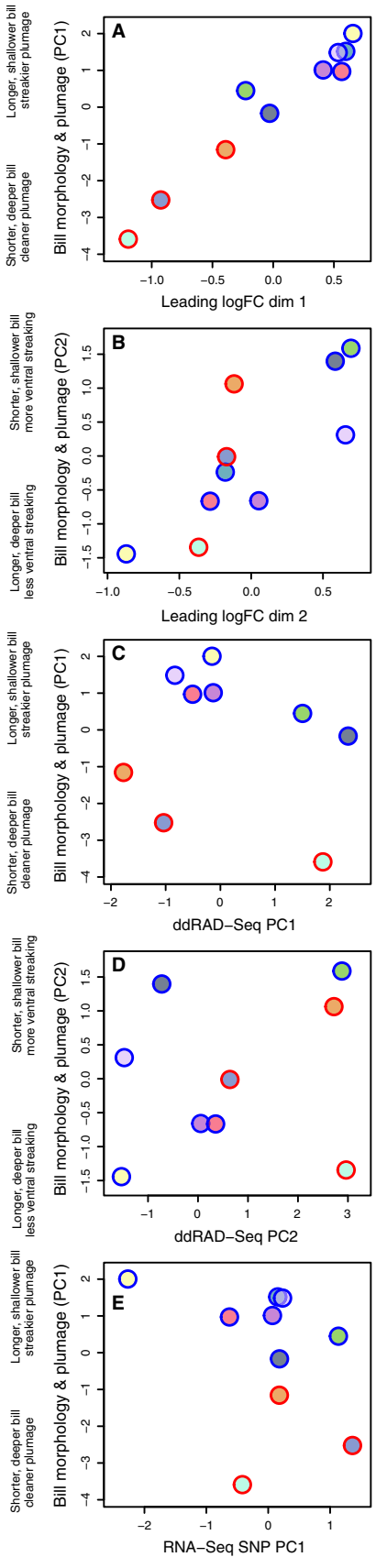

Fig. 3 Statistical associations between phenotypic diversity and genetic variation. Phenotypic variation is associated with variation in differences in gene expression profiles, but is uncorrelated with neutral variation from panels of SNPs constructed from ddRAD-Seq loci or the transcriptome. Each individual that was phenotyped and included in the RNA-Seq analyses is coded with a different colour that is consistent across panels. An outer blue circle indicates individuals that would be classified as *Acanthis flammea* according to traditional taxonomy, and an outer red circle indicates individuals that would be classified as *Acanthis hornemanni*. Scatter plots showing (A) PC1 of plumage and morphology variation and leading log-fold change dimension (LLFC) 1 of gene expression data; (B) PC2 of plumage and morphology variation and LLFC 2 of gene expression data; (C) PC1 of plumage and morphology variation and PC1 of SNP variation from ddRAD-Seq data; (D) PC2 of plumage and morphology variation and PC2 of SNP variation from ddRAD-Seq data; and (E) PC1 of plumage and morphology variation and PC1 scores from a genetic PCA of SNPs called from the de novo transcriptome.

are more similar than expected based on available conditions in North America. This pattern may be present due to the dichotomous nature of the current species classification scheme, which does not account for the continuous nature of phenotypic diversity in redpolls and could underemphasize perceived differences in abiotic conditions between taxa. Individuals with intermediate plumage are placed into one of the two species categories, which may influence how occurrence records are classified and the resulting species distribution models. Redpolls could potentially be experiencing different abiotic conditions in their contemporary distributions, which may have important implications for the patterns of differential gene expression we report (see below).

By comparing different models of species delimitation within a multispecies coalescent framework that accounts for incomplete lineage sorting, we find strong support that all currently recognized species of redpolls comprise a single coalescent lineage. In combination with the other population genetic analyses and ecological niche modelling results, our findings support the assertion that redpolls are part of a single, polymorphic metapopulation rather than distinct biological entities with separate evolutionary histories. Previous studies based on far fewer molecular markers reached similar conclusions (Marten & Johnson 1986; Seutin *et al.* 1995; Ottvall *et al.* 2002; Marthinsen *et al.* 2008). Given the low genetic differentiation we documented between currently recognized species and across large geographic expanses, redpolls might best be treated as a single species. Although the possibility persists that certain regions of the genome may be fixed or highly divergent between redpoll types, our findings do not support the assertion that multiple, separately evolving

metapopulations exist within the genus *Acanthis*. We also demonstrate that individual redpolls classified as different species span a phenotypic continuum, rather than discrete classes, which has been shown by previous studies (Troy 1985). Certain authorities, such as BirdLife International, already treat redpolls as a single species, and previous studies have arrived at similar conclusions (Troy 1985).

Redpolls are primarily granivorous and rely on patchily distributed tree seed crops that can vary substantially in abundance from year to year (Clement 2010a,b). As a result, redpolls are highly nomadic during the nonbreeding season and travel great distances to find food within flocks that regularly contain individuals that vary in phenotype; anecdotes from bird banding recoveries suggest that individual redpolls frequently travel thousands of miles (e.g. recoveries between Scandinavia and China). Widespread annual movements could contribute to genetic connectivity among geographically disjunct populations if redpolls show low breeding site fidelity and pair with individuals possessing dissimilar phenotypes. Indeed, there is little evidence for assortative mating in redpolls (Troy 1985; but see Lifjeld & Bjerke 1996) and a hybrid zone between the phenotypes, which would indicate restricted gene flow between the species, has never been reported. Finally, in addition to the continuous variation in phenotypic characters used to identify species (Troy 1985), redpolls possess very similar vocal repertoires and have been shown to adopt identical breeding calls in mixed pairs (Molau 1985) and match flock-mate calls (Mundinger 1979). The possibility exists that redpolls may develop flock-specific call repertoires that are independent of variation in plumage and morphology (N. Pieplow, personal communication), which could facilitate gene flow between currently recognized redpoll species.

Although our data suggest the presence of a single, weakly differentiated gene pool within *Acanthis*, the observed patterns could be the result of extremely recent and ongoing speciation, perhaps via ecological selection (Marthinsen *et al.* 2008). The slight differences we detected in abiotic niches between *A. flammea* and *A. hornemanni* in North America could play a role in the divergence process. Indeed, the plumage and morphological characters we measured may have adaptive significance: lighter plumage may improve camouflage at higher latitudes, and smaller bills (possessed by the higher latitude *A. hornemanni*) are known to reduce heat loss in birds (Symonds and Tattersall 2010; Greenberg *et al.* 2012). If redpolls have recently experienced divergent selection, we would expect to have detected some level of genomic heterogeneity rather than the widespread lack of differentiation that we document here; however, the number of loci analysed here represents only a fraction of the redpoll genome and denser sampling will likely be necessary to detect loci under divergent selection, if they exist (Michel *et al.* 2010; Tiffin & Ross-Ibarra 2014). Given that redpolls are abundant, incomplete lineage sorting due to large ancestral population sizes may be obscuring our ability to detect independently evolving lineages.

Differential gene expression and phenotypic diversity

Although the tissues sampled were not optimal for looking at expression differences related to the phenotypic traits we measured, we found a strong correlation between overall levels of differential gene expression and phenotypic variation among individuals. This suggests that gene expression variation could be playing an important role in generating phenotypic diversity within redpolls. Controlled aviary experiments that sample multiple stages of development and tissues more relevant to plumage streaking and bill size and shape (e.g. feather follicle, embryonic bill) will be important extensions of this work. Additionally, it is unclear whether gene expression varies seasonally in adult redpolls. It is possible that gene expression during the breeding season could be different than what we detected in this sample of wintering birds. This will be explored in future studies.

Gene expression can play an important role in the evolution of phenotypic differences and local adaptations, as shown in humans (Fraser 2013), sticklebacks (Shapiro *et al.* 2004), *Peromyscus* mice (Manceau *et al.* 2010) and *Heliconius* butterflies (Reed *et al.* 2011). *Cis*-regulatory modifications can involve few or many upstream promoters or enhancers with large cascading effects on gene expression and resulting phenotypes (Romero *et al.* 2012). Associations between our multivariate measures of gene expression and phenotypic diversity indicate broad, multigenic patterns of differential gene expression among individuals. Thus, we are currently unable to pinpoint the causal genes underlying differences in bill morphology or plumage patterning among redpolls, but have reason to believe that these differences may be adaptive and potentially related to the slight niche differences that we detected (see above). Given that phenotypic variation is continuous within the genus (Fig. 3; Troy 1985), there are probably many loci that contribute to the differences in morphology and plumage discussed here.

Although none of the genes identified in this study were significantly associated with phenotypic variation following a correction for multiple hypothesis testing, we did find multiple candidate genes worthy of further study. For example, our list of genes that were most

strongly associated with phenotypic variation included multiple genes involved in the *Wnt* signalling pathway (e.g. *tsukushin* and *frizzled*-3; Table S5, Supporting information). Expression levels of multiple genes involved in *Wnt* signalling appear to play a role in developing different bill morphologies among birds (Brugmann *et al.* 2010) and facial developmental pathways of various vertebrates (Brugmann *et al.* 2007). Additionally, the *Wnt* signalling pathway regulates *Bmp* pathway activity (Tzahor *et al.* 2003), a well-studied developmental pathway that affects bill morphology in birds (Abzhanov *et al.* 2004). With respect to plumage variation among redpolls, the *MC1R* pathway has been implicated in melanin-based phenotypic variation in many vertebrate systems (Mundy 2005; Hubbard *et al.* 2010). We found that higher expression levels of *MC5R*, which also plays a role in regulating cyclic AMP levels and the melanogenesis pathways, are associated with increased ventral and dorsal streaking (Table S5, Supporting information).

The presence of differential gene expression despite low levels of genomic differentiation suggests that phenotypic plasticity could also play an important role in generating phenotypic diversity in redpolls. Identical genotypes can produce variable phenotypes under differing environmental conditions, but examinations of phenotypic plasticity usually focus on intraspecific comparisons rather than interspecific comparisons (West-Eberhard 1989, 2005; Nijhout 2003). Environmental cues could produce different phenotypes among redpolls if they act early in development as is commonly observed in insects and plants (Nijhout 2003; West-Eberhard 2005). The potential for environmentally induced phenotypic plasticity to cause geographic variation in morphology in vertebrates is less well known (but see James 1983) and is an important avenue of future inquiry in redpolls.

Conclusions

Our ability to understand the evolutionary context that facilitates rapid phenotypic evolution has been greatly improved by the adoption of high-throughput sequencing technologies and reduced-representation genomic approaches (e.g. Wagner *et al.* 2012; Cui *et al.* 2013; Jones *et al.* 2013; Nadeau *et al.* 2013; Elmer *et al.* 2014). However, these approaches still generate data sets that represent a small portion of the genome. It is becoming increasingly clear that such approaches are not a panacea for resolving species boundaries and the genetic architecture of phenotypic diversity within rapidly diversifying lineages. As demonstrated here, even lineages that have traditionally been classified as separate species with pronounced phenotypic variation cannot always be differentiated via high-throughput sequencing. This has important implications for taxonomic revisions, conservation and the designation of conservation units (see, Funk *et al.* 2012; McCormack & Maley 2015): such phenotypically diverse lineages may represent single evolutionary units, as appears to be the case in the redpoll finches, for which thousands of anonymous loci provided no evidence of population divergence, or they may not. Even with whole-genome data, identifying loci that contribute to phenotypic diversity is challenging, particularly for quantitative traits that probably involve many loci (Tiffin & Ross-Ibarra 2014). Intriguingly, our results indicate that gene expression information may reveal how phenotypic differences arise among taxa; however, controlled conditions are required to remove environmentally induced variation as a source of bias.

Rapid bouts of phenotypic diversification and speciation have provided seminal examples of evolution in action, yet also present theoretical, computational and conservation challenges. Despite technological advances that have characterized the genomic era, these challenges remain in many systems. Continued persistence and analytical innovation will reward molecular ecologists with acute knowledge regarding the evolutionary and ecological processes that comprise lineage diversification and phenotypic differentiation in rapidly evolving lineages.

Acknowledgements

M. Young, A. Cibois, D. Bonter, J. Barry, L. Seitz and S. Birks at The University of Washington Burke Museum helped gather genetic resources. E. Bondra and S. Bogdanowicz assisted with library prep. We thank I. Lovette, R. Harrison and K. Wagner for comments on the manuscript. G. Bradburd, P. Title, L. Campagna, P. Deane-Coe, A. Ellison, C. Balakrishnan, Z. Cheviron and M. Hare provided useful insight and discussions. W. Jeyasingham drew redpoll portraits for our figures. We thank three anonymous reviewers for helpful feedback that improved the manuscript. We are grateful towards all institutions that have provided occurrence data through the GBIF portal. This study was funded in part by a Keickheffer Adirondack Fellowship, the Cornell Lab of Ornithology Athena Fund and the Fuller Evolutionary Biology Program at the Cornell Lab of Ornithology. SAT was supported in part by the Cornell Center for Comparative and Population Genomics as well as a Banting Fellowship from the Natural Sciences and Engineering Research Council of Canada. NAM was supported in part by an EPA STAR Graduate Fellowship (F13F21201).

References

Abràmoff MD, Magalhães PJ, Ram SJ (2004) Image processing with ImageJ. *Biophotonics International*, **11**, 36–43.

Abzhanov A, Protas M, Grant BR, Grant PR, Tabin CJ (2004) Bmp4 and morphological variation of beaks in Darwin's finches. *Science*, **305**, 1462–1465.

Barton NH, Hewitt GM (1981) Hybrid zones and speciation. In: *Evolution and Speciation* (eds Atchley WR, Woodruff DS), pp. 109–145. Cambridge University Press, Cambridge, UK.

Barve N, Barve V, Jiménez-Valverde A *et al.* (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, **222**, 1810–1819.

Bazykin AD (1969) Hypothetical mechanism of speciation. *Evolution*, **23**, 685–687.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Methodological*, **57**, 289–300.

Bouckaert R, Heled J, Kühnert D *et al.* (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, **10**, e1003537.

Brawand D, Soumillon M, Necsulea A *et al.* (2012) The evolution of gene expression levels in mammalian organs. *Nature*, **478**, 343–348.

Brooks WS (1968) Comparative adaptations of the Alaskan redpolls to the arctic environment. *Wilson Bulletin*, **80**, 253–280.

Brugmann SA, Goodnough LH, Gregorieff A *et al.* (2007) Wnt signaling mediates regional specification in the vertebrate face. *Development*, **134**, 3283–3295.

Brugmann SA, Powder KE, Young NM *et al.* (2010) Comparative gene expression analysis of avian embryonic facial structures reveals new candidates for human craniofacial disorders. *Human Molecular Genetics*, **19**, 920–930.

Bryant D, Bouckaert R, Felsenstein J *et al.* (2012) Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution*, **29**, 1917–1932.

Camacho C, Coulouris G, Avagyan V *et al.* (2009) BLAST+: architecture and applications. *BMC Bioinformatics*, **10**, 421.

Catchen JM, Amores A, Hohenlohe P *et al.* (2011) Stacks: building and genotyping Loci de novo from short-read sequences. *G3 (Bethesda, Md.)*, **1**, 171–182.

Catchen J, Hohenlohe PA, Bassham S *et al.* (2013) Stacks: an analysis tool set for population genomics. *Molecular Ecology*, **22**, 3124–3140.

Chessel D, Dufour AB, Dray S (2004) *CRAN Package ade4. Analysis of Ecological Data. v1.6-2*. R Foundation for Statistical Computing, Vienna, Austria.

Clement P (2010a) Common Redpoll (*Carduelis flammea*). In: *Handbook of the Birds of the World* (eds del Hoyo J, Elliott A, Christie DA), pp. 564–565. Lynx Edicions, Barcelona, Spain.

Clement P (2010b) Arctic Redpoll (*Carduelis hornemanni*). In: *Handbook of the Birds of the World* (eds del Hoyo J, Elliott A, Christie DA), pp. 565–566. Lynx Edicions, Barcelona, Spain.

Clements JF, Schulenberg TS, Iliff MJ *et al.* (2014) *The Clements Checklist of Birds of the World: Version 6.9*. Cornell University Press, Ithaca, New York.

Conesa A, Gotz S, Garcia-Gomez JM *et al.* (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, **21**, 3674–3676.

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.

Cui R, Schumer M, Kruesi K, Walter R, Andolfatto P, Rosenthal GG (2013) Phylogenomics reveals extensive reticulate evolution in *Xiphophorus* fishes. *Evolution*, **67**, 2166–2179.

Danecek P, Auton A, Abecasis G *et al.* (2011) The variant call format and VCFtools. *Bioinformatics*, **27**, 2156–2158.

DePristo MA, Banks E, Poplin R *et al.* (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, **43**, 491–498.

Drovetski SV, Zink RM, Rohwer S *et al.* (2004) Complex biogeographic history of a Holarctic passerine. *Proceedings of the Royal Society B: Biological Sciences*, **271**, 545–551.

Drovetski SV, Zink RM, Mode NA (2009) Patchy distribution belie morphological and genetic homogeneity in rosy-finches. *Molecular Phylogenetics and Evolution*, **50**, 437–445.

Earl DA, vonHoldt BM (2011) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

Ekblom R, Farrell LL, Lank DB, Burke T (2012) Gene expression divergence and nucleotide differentiation between males of different color morphs and mating strategies in the ruff. *Ecology and Evolution*, **2**, 2485–2505.

Elith J, Graham CH, Anderson RP *et al.* (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.

Elith J, Phillips SJ, Hastie T *et al.* (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, **17**, 43–57.

Elmer KR, Fan S, Kusche H *et al.* (2014) Parallel evolution of Nicaraguan crater lake cichlid fishes via non-parallel routes. *Nature Communications*, **5**, 5168.

Escudero M, Eaton DAR, Hahn M, Hipp AL (2014) Genotyping-by-sequencing as a tool to infer phylogeny and ancestral hybridization: a case study in *Carex* (Cyperaceae). *Molecular Phylogenetics and Evolution*, **79**, 359–367.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Filteau M, Pavey SA, St-Cyr J, Bernatchez L (2013) Gene coexpression networks reveal key drivers of phenotypic divergence in lake whitefish. *Molecular Biology and Evolution*, **30**, 1384–1396.

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**, 977–993.

Fraser HB (2013) Gene expression drives local adaptation in humans. *Genome Research*, **23**, 1089–1096.

Funk WC, McKay JK, Hohenlohe PA, Allendorf FW (2012) Harnessing genomics for delineating conservation units. *Trends in Ecology and Evolution*, **27**, 489–496.

Grabherr MG, Haas BJ, Yassour M *et al.* (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, **29**, 644–652.

Graham CH, Ron SR, Santos JC *et al.* (2004) Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. *Evolution*, **58**, 1781.

Greenberg R, Cadena V, Danner RM, Tattersall G (2012) Heat loss may explain bill size differences between birds occupying different habitats. *PLoS ONE*, **7**, e40933.

Grummer JA, Bryson RW, Reeder TW (2014) Species delimitation using bayes factors: simulations and application to the sceloporus scalaris species group (Squamata: Phrynosomatidae). *Systematic Biology*, **63**, 119–133.

Haas BJ, Papanicolaou A, Yassour M *et al.* (2013) De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, **8**, 1494–1512.

Harris MP, Norman FJ, McColl R (1965) A mixed population of redpolls in northern Norway. *British Birds*, **58**, 288–294.

Harrison RG (1990) Hybrid zones: windows on evolutionary process. *Oxford Surveys in Evolutionary Biology*, **7**, 69–128.

Harrison RG, Rand DM (1989) Mosaic hybrid zone and the nature of species boundaries. In: *Speciation and its Consequences* (eds Otte D, Ender JA), pp. 111–133. Sinauer Associates Inc, Sunderland, Massachusetts.

Harrison PW, Wright AE, Mank JE (2012) The evolution of gene expression and the transcriptome–phenotype relationship. *Seminars in Cell & Developmental Biology*, **23**, 222–229.

Harvey MG, Brumfield RT (2014) Genomic variation in a widespread Neotropical bird (*Xenops minutus*) reveals divergence, population expansion, and gene flow. *Molecular Phylogenetics and Evolution*, **83**, 305–316.

Hastie T, Tibshirani R, Friedman JH (2001) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer-Verlag, New York, New York.

Heibl C, Calenge C (2013) phyloclim: Integrating phylogenetics and climatic niche modeling. R package version 0.9-4.

Herremans M (1989) Taxonomy and evolution in redpolls *Carduelis flammea*–hornemanni; a multivariate study of their biometry. *Ardea*, **78**, 441–458.

Hijmans RJ, Garrett KA, Huaman Z *et al.* (2000) Assessing the geographic representativeness of genebank collections: the case of Bolivian wild potatoes. *Conservation Biology*, **14**, 1755–1765.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hijmans RJ, Phillips S, Leathwick J, Elith J (2013) dismo: Species distribution modeling.

Hipp AL, Eaton DAR, Cavender-Bares J, Fitzek E, Nipper R, Manos PS (2014) A framework phylogeny of the American oak clade based on sequenced RAD data. *PLoS ONE*, **9**, e93975.

vonHoldt BM, Pollinger JP, Earl DA *et al.* (2011) A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids. *Genome Research*, **21**, 1294–1305.

Hubbard JK, Uy JAC, Hauber ME, Hoekstra HE, Safran RJ (2010) Vertebrate pigmentation: from underlying genes to adaptive function. *Trends in Genetics*, **26**, 231–239.

James FC (1983) Environmental component of morphological differentiation in birds. *Science*, **221**, 184–186.

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403–1405.

Jombart T, Ahmed I (2011) adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, **27**, 3070–3071.

Jones JC, Fan S, Franchini P, Schartl M, Meyer A (2013) The evolutionary history of *Xiphophorus* fish and their sexually selected sword: a genome-wide approach to using restriction site-associated DNA sequencing. *Molecular Ecology*, **22**, 2986–3001.

Kass RE, Raftery AE (1995) Bayes factors. *Journal of the American Statistical Association*, **90**, 773–795.

Key K (1968) The concept of stasipatric speciation. *Systematic Biology*, **17**, 14–22.

Knox A, Helbig A, Parkin D, Sangster G (2001) The taxonomic status of Lesser Redpoll. *British Birds*, **94**, 260–267.

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, **10**, R25.

Larson WA, Seeb LW, Everett MV, Waples RK, Templin WD, Seeb JE (2013) Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (*Oncorhynchus tshawytscha*). *Evolutionary Applications*, **7**, 355–369.

Leaché AD, Fujita MK, Minin VN, Bouckaert RR (2014) Species delimitation using genome-wide SNP data. *Systematic Biology*, **63**, 534–542.

Legendre P, Gallagher E (2001) Ecologically meaningful transformations for ordination of species data. *Oecologia*, **129**, 271–280.

Lemmon EM, Lemmon AR (2013) High-throughput genomic data in systematics and phylogenetics. *Annual Review of Ecology Evolution and Systematics*, **44**, 99–121.

Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference. *BMC Bioinformatics*, **12**, 323.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754–1760.

Lifjeld JT, Bjerke BA (1996) Evidence for assortative pairing by the *cabaret* and *flammea* subspecies of the common redpoll *Carduelis flammea* in SE Norway. *Fauna Norvegica Series C, Cinclus*, **19**, 1–8.

Lohse M, Bolger AM, Nagel A *et al.* (2012) RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. *Nucleic Acids Research*, **40**, W622–W627.

Maddison W, Knowles L (2006) Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology*, **55**, 21–30.

Manceau M, Domingues VS, Linnen CR, Rosenblum EB, Hoekstra HE (2010) Convergence in pigmentation at multiple levels: mutations, genes and function. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 2439–2450.

Marten JA, Johnson NK (1986) Genetic relationships of North American cardueline finches. *Condor*, **88**, 409–420.

Marthinsen G, Wennerberg L, Lifjeld JT (2008) 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.

Mastretta-Yanes A, Arrigo N, Alvarez N, Jorgensen TH, Piñero D, Emerson BC (2015) Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Molecular Ecology*, **15**, 28–41.

McCormack JE, Maley JM (2015) Interpreting negative results with taxonomic and conservation implications: another look

at the distinctness of coastal California Gnatcatchers. *Auk*, **132**, 380–388.
McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT (2013) Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution*, **66**, 526–538.
Michel AP, Sim S, Powell TH *et al.* (2010) Widespread genomic divergence during sympatric speciation. *Proceedings of the National Academy of Sciences of the USA*, **107**, 9724–9729.
Molau M (1985) Gråsiskkomplexet i Sverige (The redpoll complex in Sweden). *Vår Fågelvärld*, **44**, 5–20.
Mundinger PC (1979) Call learning in the carduelinae: ethological and systematic considerations. *Systematic Biology*, **28**, 270–283.
Mundy NI (2005) A window on the genetics of evolution: MC1R and plumage colouration in birds. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 1633–1640.
Nadeau NJ, Martin SH, Kozak KM *et al.* (2013) Genome-wide patterns of divergence and gene flow across a butterfly radiation. *Molecular Ecology*, **22**, 814–816.
Nice CC, Gomp[ert Z, Fordyce JA, Forister ML, Lucas LK, Buerkle CA (2012) Hybrid speciation and independent evolution in lineages of alpine butterflies. *Evolution*, **67**, 1055–1068.
Nijhout HF (2003) Development and evolution of adaptive polyphenisms. *Evolution and Development*, **5**, 9–18.
Nosil P, Feder JL (2012) Genomic divergence during speciation: causes and consequences. *Philosophical Transactions of the Royal Society B*, **367**, 332–342.
Ottvall R, Bensch S, Walinder G, Lifjeld JT (2002) No evidence of genetic differentiation between lesser redpolls *Carduelis flammea cabaret* and common redpolls *Carduelis f. flammea*. *Avian Science*, **2**, 237–244.
Paradis E (2010) pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics*, **26**, 419–420.
Pavey SA, Collin H, Nosil P, Rogers S (2010) The role of gene expression in ecological speciation. *Annals of the New York Academy of Sciences*, **1206**, 110–129.
Peterson AT, Soberon J, Sánchez-Cordero V (1999) Conservatism of ecological niches in evolutionary time. *Science*, **285**, 1265–1267.
Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, **7**, e37135.
Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.
Poelstra JW, Vijay N, Bossu CM *et al.* (2014) The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science*, **344**, 1410–1414.
Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
Questiau S, Eybert MC, Gaginskaya AR, Gielly L, Taberlet P (1998) Recent divergence between two morphologically differentiated subspecies of bluethroat (Aves: Muscicapidae: *Luscinia svecica*) inferred from mitochondrial DNA sequence variation. *Molecular Ecology*, **7**, 239–245.
R Core Team (2014) *R: A Language and Environment for Statistical Computing*. R Core Team, Vienna, Austria.
Rand DM, Harrison RG (1989) Ecological genetics of a mosaic hybrid zone: mitochondrial, nuclear, and reproductive differentiation of crickets by soil type. *Evolution*, **43**, 432–449.
Reed RD, Papa R, Martin A *et al.* (2011) Optix drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science*, **333**, 1137–1141.
Ritchie ME, Phipson B, Wu D *et al.* (2015) Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, **43**, 1–13.
Robinson MD, McCarthy DJ, Smyth GK (2009) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139–140.
Romero IG, Ruvinsky I, Gilad Y (2012) Comparative studies of gene expression and the evolution of gene regulation. *Nature Reviews Genetics*, **13**, 505–516.
Sangster G, Knox AG, Helbig AJ, Parkin DT (2002) Taxonomic recommendations for European birds. *Ibis*, **144**, 153–159.
Schoener TW (1968) The Anolis lizards of Bimini: resource partitioning in a complex fauna. *Ecology*, **49**, 704–726.
Seehausen O, Butlin RK, Keller I *et al.* (2014) Genomics and the origin of species. *Nature Reviews Genetics*, **15**, 176–192.
Seutin G, Boag PT, White BN, Ratcliffe LM (1991) Sequential polyandry in the Common Redpoll (*Carduelis flammea*). *Auk*, **108**, 166–170.
Seutin G, Boag PT, Ratcliffe LM (1992) Plumage variability in redpolls from Churchill, Manitoba. *Auk*, **109**, 771–785.
Seutin G, Ratcliffe LM, Boag PT (1995) Mitochondrial DNA homogeneity in the phenotypically diverse redpoll finch complex (Aves: Carduelinae: *Carduelis flammea-hornemanni*). *Evolution*, **49**, 962–973.
Shapiro MD, Marks ME, Peichel CL *et al.* (2004) Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature*, **428**, 717–723.
Smyth GK (2005) Limma: linear models for microarray data. In: *Bioinformatics and Computational Biology Solutions Using R and Bioconductor* (eds Gentleman R, Carey V, Huber W, Irizarry R, Dudoit S), pp. 397–420. Springer, New York, NY.
Symonds MRE, Tattersall GJ (2010) Geographical variation in bill size across bird species provides evidence for allen's rule. *The American Naturalist*, **176**, 188–197.
Tiffin P, Ross-Ibarra J (2014) Advances and limits of using population genetics to understand local adaptation. *Trends in Ecology and Evolution*, **29**, 673–680.
Troy DM (1985) A phenetic analysis of the redpolls *Carduelis flammea flammea* and *C. hornemanni exilipes*. *Auk*, **102**, 82–96.
Tzahor E, Kempf H, Mootoosamy RC *et al.* (2003) Antagonists of Wnt and BMP signaling promote the formation of vertebrate head muscle. *Genes and Development*, **17**, 3087–3099.
Wagner CE, Keller I, Wittwer S *et al.* (2012) Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology*, **22**, 787–798.
Warren DL, Glor RE, Turelli M (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*, **62**, 2868–2883.

Warren DL, Glor RE, Turelli M (2010) ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography*, **33**, 607–611.

West-Eberhard MJ (1989) Phenotypic plasticity and the origins of diversity. *Annual Review of Ecology and Systematics*, **20**, 249–278.

West-Eberhard MJ (2005) Developmental plasticity and the origin of species differences. *Proceedings of the National Academy of Sciences of the USA*, **102**, 6543–6549.

Wolf JBW, Lindell J, Backstrom N (2010) Speciation genetics: current status and evolving approaches. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 1717–1733.

Wu CI (2001) The genic view of the process of speciation. *Journal of Evolutionary Biology*, **14**, 851–865.

Zuccon D, Prŷs-Jones R, Rasmussen PC, Ericson PGP (2012) The phylogenetic relationships and generic limits of finches (Fringillidae). *Molecular Phylogenetics and Evolution*, **62**, 581–596.

S.A.T. and N.A.M. conceived the study. N.A.M. and S.A.T. conducted field work. S.A.T. generated ddRAD-Seq libraries. N.A.M. performed transcriptome assembly and differential gene expression analyses. N.A.M. and S.A.T. annotated the transcriptome. N.A.M. ran the population genetic analyses. N.A.M. generated figures with help from S.A.T. N.A.M. and S.A.T. wrote the manuscript.

Data accessibility

The de novo transcriptome assembly and raw reads of RNA-Seq data are available through BioProject number PRJNA256306. Raw, demultiplexed reads from the ddRAD-Seq portion of this study are available through the NCBI Short Read Archive (SRP052607). Phenotypic measurements and occurrence data used in this study are available through the Supporting Information. Panels of SNP data generated from the ddRAD-Seq pipelines and the transcriptome, FPKM read counts for gene expression analyses, input files for SNAPP, a posterior distribution of trees generated from SNAPP, and R and shell scripts used to run bioinformatics pipelines and analyse data are available through Dryad (http://dx.doi.org/10.5061/dryad.15rk0).

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Sampling information for 77 ingroup and two outgroup individuals for population genetic analyses presented in this study.

Table S2 Phenotype data for 10 individuals collected for RNA-Seq experiment.

Table S3 Breeding occurrence data for *A. flammea* and *A. hornemanni* in North America used in generating ecological niche models.

Table S4 Results of Structure with different settings of the *K* parameter.

Table S5 Results of differential gene expression analysis using generalized linear models and morphology + plumage PC1 as the response variable.

Fig. S1 Loadings for principal component axes of plumage and morphology characters.

Fig. S2 Scatterplot of first two principal component axes of plumage and morphology.

Fig. S3 Summary statistics from Trinity de novo assembly.

Fig. S4 Comparison of observed heterozygosity among 20 712 loci sampled from *A. flammea*, *A. hornemanni*, and *A. cabaret*.

Fig. S5 Results of Structure and genetic PCA analyses using a locus assembly pipeline that includes outgroup individuals.

Fig. S6 Results of PCA using SNPs called from the transcriptome.

Fig. S7 Output of BAYESCAN run on ddRAD-Seq loci, which did not detect any outlier loci between currently recognized species of redpolls.

Fig. S8 Percent contribution of each BioClim variable included in constructing ecological niche models for *A. hornemanni* and *A. flammea*.

Fig. S9 Response curves of each BioClim variable included in constructing ecological niche models for *A. hornemanni* and *A. flammea*.

Fig. S10 Statistical tests for niche identity (top) and background similarity (bottom) for *A. hornemanni* and *A. flammea*.