

Epigenetics and the Evolution of Darwin's Finches

Michael K. Skinner^{1,*}, Carlos Gurerrero-Bosagna^{1,3}, M. Muksit Haque¹, Eric E. Nilsson¹, Jennifer A.H. Koop^{2,4}, Sarah A. Knutie², and Dale H. Clayton²

¹Center for Reproductive Biology, School of Biological Sciences, Washington State University

²Department of Biology, University of Utah

³Present address: Department of Physics, Biology and Chemistry (IFM), Linköping University, Sweden

⁴Present address: Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ

*Corresponding author: E-mail: skinner@wsu.edu.

Accepted: July 18, 2014

Data deposition: All DMR and CNV genomic data obtained in this study have been deposited in the NCBI public GEO database under the accession (GEO #: GSE58334).

Abstract

The prevailing theory for the molecular basis of evolution involves genetic mutations that ultimately generate the heritable phenotypic variation on which natural selection acts. However, epigenetic transgenerational inheritance of phenotypic variation may also play an important role in evolutionary change. A growing number of studies have demonstrated the presence of epigenetic inheritance in a variety of different organisms that can persist for hundreds of generations. The possibility that epigenetic changes can accumulate over longer periods of evolutionary time has seldom been tested empirically. This study was designed to compare epigenetic changes among several closely related species of Darwin's finches, a well-known example of adaptive radiation. Erythrocyte DNA was obtained from five species of sympatric Darwin's finches that vary in phylogenetic relatedness. Genome-wide alterations in genetic mutations using copy number variation (CNV) were compared with epigenetic alterations associated with differential DNA methylation regions (epimutations). Epimutations were more common than genetic CNV mutations among the five species; furthermore, the number of epimutations increased monotonically with phylogenetic distance. Interestingly, the number of genetic CNV mutations did not consistently increase with phylogenetic distance. The number, chromosomal locations, regional clustering, and lack of overlap of epimutations and genetic mutations suggest that epigenetic changes are distinct and that they correlate with the evolutionary history of Darwin's finches. The potential functional significance of the epimutations was explored by comparing their locations on the genome to the location of evolutionarily important genes and cellular pathways in birds. Specific epimutations were associated with genes related to the bone morphogenic protein, toll receptor, and melanogenesis signaling pathways. Species-specific epimutations were significantly overrepresented in these pathways. As environmental factors are known to result in heritable changes in the epigenome, it is possible that epigenetic changes contribute to the molecular basis of the evolution of Darwin's finches.

Key words: epimutations, DNA methylation, copy number variation, phylogeny, adaptive radiation, BMP, toll, melanogenesis.

Introduction

Epigenetic change has been postulated to play a role in the ecology and evolution of natural populations (Richards et al. 2010; Holeski et al. 2012; Liebl et al. 2013). Epigenetic changes are broadly defined as "molecular processes around DNA that regulate genome activity independent of DNA sequence and are mitotically stable" (Skinner et al. 2010). Some epigenetic processes are also meiotically stable and are transmitted through the germline (Anway et al. 2005; Jirtle and Skinner 2007). These epigenetic mechanisms, such as DNA methylation, can become programmed

(e.g., imprinted) and inherited over generations with potential evolutionary impacts. Environmental factors have been shown to promote the epigenetic transgenerational inheritance of phenotypic variants (Skinner et al. 2010). In recent years, the importance of environmental cues in the induction of such variation has been widely acknowledged (Bonduriansky 2012). Thus, like genetic change (Greenspan 2009), epigenetic change may also play an important role in evolution (Guerrero-Bosagna et al. 2005; Day and Bonduriansky 2011; Geoghegan and Spencer 2012, 2013a, 2013b, 2013c; Klironomos et al. 2013).

© The Author(s) 2014. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

In order for inherited epigenetic changes to play a significant role in microevolution, they must persist for tens of generations, or longer (Slatkin 2009). It is conceivable that epigenetic changes may also accumulate over longer periods of evolutionary time, contributing to processes such as adaptive radiation (Rebollo et al. 2010; Flatscher et al. 2012). This hypothesis assumes that epigenetic changes persist over thousands of generations. An initial step in testing this hypothesis would be to compare epigenetic differences among closely related species, and whether such changes accumulate over short spans of macroevolutionary time. For example, do epigenetic changes accumulate with phylogenetic distance? Addressing this question was the primary goal of this study.

The study was designed to explore the relationship between epigenetic changes and the evolutionary history of several species of Darwin's finches in the Galapagos Islands. This group of birds has been central to work on a variety of important topics in evolutionary biology, including adaptive radiation, character displacement, rapid evolution, hybridization between species, evolutionary developmental mechanisms, and the effect of invasive pathogens and parasites (Grant and Grant 2008; Huber et al. 2010; Donohue 2011). The adaptive radiation of Darwin's finches over a period of 2–3 Myr resulted in 14 extant species that fill distinct ecological niches. These species show striking variation in body size and the size and shape of their beaks (Grant and Grant 2008). Darwin's finches were selected for study because they are a well-studied example of the evolution of closely related species into different ecological niches (Grant and Grant 2008; Donohue 2011).

Natural selection is a process in which environmental factors influence the survival and reproductive success of individuals bearing different phenotypes. Only selection on phenotypic traits with a heritable basis can lead to evolutionary change (Endler 1986). Observations indicate that epigenetic mechanisms have a role in influencing genomic variability (Huttley 2004; Ying and Huttley 2011). As epigenetic changes are also influenced by environmental factors, and can be heritable across generations (Skinner et al. 2010), they provide another molecular mechanism that can influence evolutionary change. Although Lamarck (1802) proposed that environmental factors can influence inheritance directly, his mechanism has not been widely recognized as a component of modern evolutionary theory (Day and Bonduriansky 2011). Recent work in epigenetics shows that epigenetic changes can, in fact, increase the heritable phenotypic variation available to natural selection (Holeski et al. 2012; Liebl et al. 2013). Thus, epigenetics appears to provide a molecular mechanism that can increase phenotypic variation on which selection acts (Skinner 2011). The integration of genetic and epigenetic mechanisms has the potential to significantly expand our understanding of the origins of phenotypic variation and how environment can influence evolution.

For example, Crews et al. (2007) investigated the ability of an environmental factor (toxicant) to promote the epigenetic

transgenerational inheritance of alterations in the mate preferences of rats, with consequences for sexual selection. An F0 generation gestating female rat was exposed to the agricultural fungicide vinclozolin transiently. A dramatic alteration in the mate preferences of the F3 generation was observed (Crews et al. 2007) along with epigenetic alterations (termed epimutations) in the germline (sperm) (Guerrero-Bosagna et al. 2010). Transgenerational transcriptome changes in brain regions correlated with these alterations in mate preference behavior were also observed (Skinner et al. 2008, 2014). Thus, an environmental factor that altered mate preference was found to promote a transgenerational alteration in the sperm epigenome in an imprinted-like manner that was inherited for multiple generations (Crews et al. 2007; Skinner et al. 2010). Studies such as these suggest that environmental epigenetics may play a role in evolutionary changes through processes, such as sexual selection.

Recent reviews suggest a pervasive role for epigenetics in evolution (Rebollo et al. 2010; Day and Bonduriansky 2011; Kuzawa and Thayer 2011; Flatscher et al. 2012; Klironomos et al. 2013). The primary goal of this study was to test whether epigenetic changes accumulate over the long periods of evolutionary time required for speciation with adaptive radiation. Genome wide analyses were used to investigate changes in genetic and epigenetic variation among five species of Darwin's finches. The measure of genetic variation was copy number variation (CNV), which has been shown to provide useful and stable genetic markers with potentially more phenotypic functional links than point mutations such as single nucleotide polymorphisms (SNPs) (Lupski 2007; Sudmant et al. 2013). CNVs involve an increase or decrease in the number of copies of a repeat element at a specific genomic location. Recently, CNV changes in primates and other species have been shown to be very useful genetic measures for comparing evolutionary events (Nozawa et al. 2007; Gazave et al. 2011; Poptsova et al. 2013). CNV changes are involved in gene duplication and deletion phenomena, as well as repeat element phenomenon such as translocation events and can be influenced by DNA methylation (Skinner et al. 2010; Macia et al. 2011; Tang et al. 2012). The measure of epigenetic variation used was differential DNA methylation sites, which are known to be stable and heritable (Skinner et al. 2010). Comparing data for both genetic mutations (i.e., CNV) and epimutations (i.e., DNA methylation) allowed the relative magnitudes of these sources of variation to be compared across the five species included in the study.

Materials and Methods

Finch Field Work and Collection of Blood

Blood samples were collected from birds captured January–April 2009 at El Garrapatero, a lowland arid site on Santa Cruz Island, Galapagos Archipelago, Ecuador (Koop et al. 2011).

Birds were captured with mist nests and banded with numbered Monel bands to track recaptures. Birds were identified, aged, and sexed using size and plumage characteristics. A small blood sample (90 μ l) from each bird was collected in a microcapillary tube through brachial venipuncture. Samples were stored on wet ice in the field, then erythrocytes purified by centrifugation and cells stored in a -20°C freezer at a field station. Following the field season, samples were placed in a -80°C freezer for longer term storage. All procedures were approved by the University of Utah Institutional Animal Care and Use Committee (protocol #07-08004) and by the Galápagos National Park (PC-04-10: #0054411).

DNA Processing

Erythrocyte DNA was isolated with DNAeasy Blood and Tissue Kit (Qiagen, Valencia, CA) and then stored at -80°C prior to analysis. DNA was sonicated following a previously described protocol (without protease inhibitors) (Tateno et al. 2000) and then purified using a series of washes and centrifugations (Ward et al. 1999) from variable number of animals per species analyzed. The same concentrations of DNA from individual blood samples were then used to produce pools of DNA material. Two DNA pools were produced in total per species, each one containing the same amount of DNA from different animals. The number of individuals used per pool is shown in [supplementary table S6, Supplementary Material](#) online. These DNA pools were then used for chromosomal genomic hybridization (CGH) arrays or chromatin immunoprecipitation of methylated DNA fragments (MeDIP).

CNV Analysis

The array used for the CNV analysis was a CGH custom design by Roche Nimblegen that consisted of a whole-genome tiling array of zebra finch (*Taeniopygia guttata*) with 720,000 probes per array. The probe size ranged from 50 to 75 mer in length with median probe spacing of 1,395 bp. Two different comparative (CNV vs. CNV) hybridization experiments were performed (two subarrays) for each species in query (*Geospiza fuliginosa* [FUL], *G. scandens* [SCA], *Camarhynchus parvulus* [PAR], and *Platyspiza crassirostris* [CRA]) versus control *G. fortis* (FOR), with each subarray including hybridizations from DNA pools from these different species. Two DNA pools were built for each species ([supplementary table S6, Supplementary Material](#) online). For one subarray of each species, DNA samples from the experimental groups were labeled with Cy5 and DNA samples from the control lineage were labeled with Cy3. For the other subarray of each species, a dye swap was performed so that DNA samples from the experimental groups were labeled with Cy3 and DNA samples from the control lineage were labeled with Cy5.

For the CNV experiment raw data from the Cy3 and Cy5 channels were imported into R (R Development Core Team 2010), checked for quality, and converted to MA values

($M = \text{Cy5} - \text{Cy3}$; $A = [\text{Cy5} + \text{Cy3}]/2$). Within array and between array normalizations were performed as previously described (Manikkam et al. 2012). Following normalization, the average value of each probe was calculated and three different CNV algorithms were used on each of these probes including circular binary segmentation from the DNA copy (Olshen et al. 2004), CGHseg (Picard et al. 2005) and cghFlasso (Tibshirani and Wang 2008). These three algorithms were used with the default parameters. The average values from the output of these algorithms were obtained. A threshold of 0.04 as a cutoff was used on the summary (average of the log-ratio from the three algorithms) where gains are probes above the positive threshold and losses are probes below the negative threshold. Consecutive probes (≥ 3) of gains and losses were used to identify separate CNV regions. A cutoff of three-probe minimum was used and those regions were considered a valid CNV. The statistically significant CNVs were identified and *P* values associated with each region presented. A cutoff of $P < 10^{-5}$ was used to select the final regions of gains and losses.

Differential DNA Methylation Regions Analysis

MeDIP was performed as previously described (Guerrero-Bosagna et al. 2010) as follows: 6 μ g of genomic DNA was subjected to series of three 20-pulse sonications at 20% amplitude and the appropriate fragment size (200–1,000 ng) was verified through 2% agarose gels; the sonicated genomic DNA was resuspended in 350 μ l TE buffer and denatured for 10 min at 95°C and then immediately placed on ice for 5 min; 100 μ l of $5\times$ IP buffer (50 mM Na-phosphate pH 7, 700 mM NaCl (PBS), 0.25% Triton X-100) was added to the sonicated and denatured DNA. An overnight incubation of the DNA was performed with 5 μ g of antibody anti-5-methylCytidine monoclonal from Diagenode (Denville, NJ) at 4°C on a rotating platform. Protein A/G beads from Santa Cruz were prewashed on PBS-BSA (bovine serum albumin) 0.1% and resuspended in 40 μ l $1\times$ IP (immunoprecipitation) buffer. Beads were then added to the DNA-antibody complex and incubated 2 h at 4°C on a rotating platform. Beads bound to DNA-antibody complex were washed three times with 1 ml $1\times$ IP buffer; washes included incubation for 5 min at 4°C on a rotating platform and then centrifugation at 6,000 rpm for 2 min. Beads DNA-antibody complex were then resuspended in 250 μ l digestion buffer (50 mM Tris-HCl pH 8, 10 mM ethylenediaminetetraacetic acid, 0.5% SDS (sodium dodecyl sulfate) and 3.5 μ l of proteinase K (20 mg/ml) was added to each sample and then incubated overnight at 55°C on a rotating platform. DNA purification was performed first with phenol and then with chloroform:isoamyl alcohol. Two washes were then performed with 70% ethanol, 1 M NaCl, and glycogen. MeDIP-selected DNA was then resuspended in 30 μ l TE buffer.

The array used for the differential methylation analysis was a DNA-methylated custom array by Roche Nimblegen that

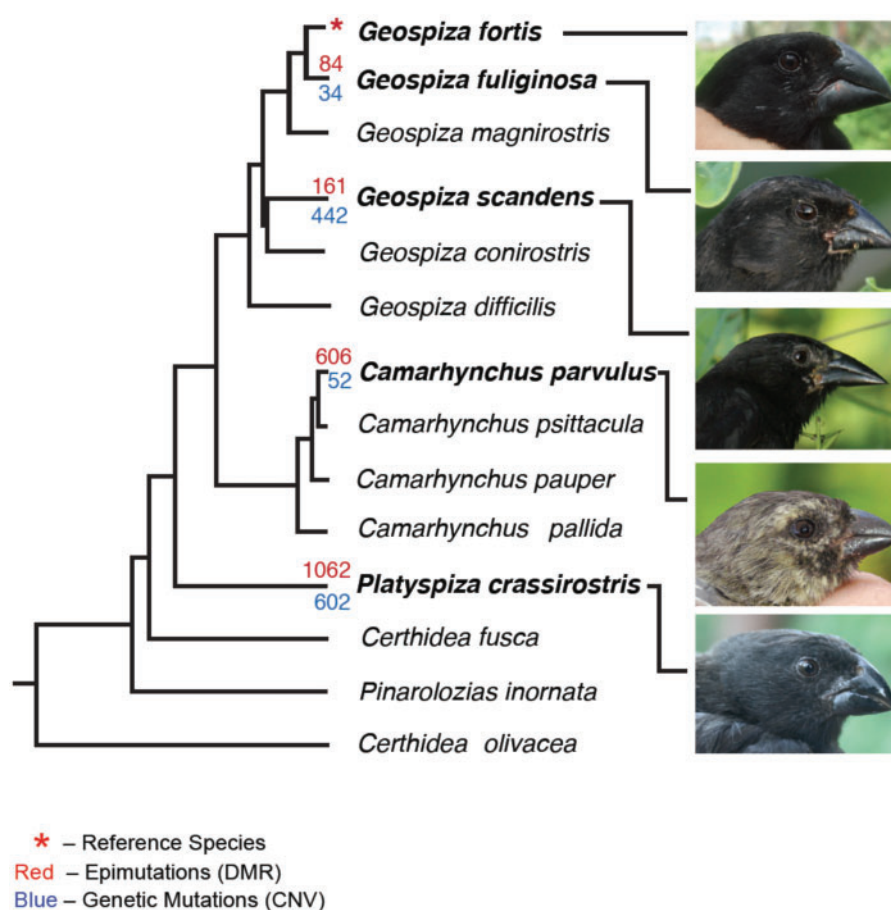


FIG. 1.—Number of epimutations and genetic mutations in relation to the phylogenetic relationships of five species of Darwin's finches. Photographs (by J.A.H.K. or S.A.K.) show variation in bill size and shape. Numbers on branches are the number of differences (three or more probes; table 1) in epimutations (DMR; in red) and genetic mutations (CNV; in blue) for each of four species, compared with a single reference species FOR (asterisk). The phylogram is based on allele length variation at 16 polymorphic microsatellite loci (from Petren et al. 1999). The topology of the tree is similar to that proposed by Lack (1947) on the basis of morphological traits.

consisted of a whole-genome tiling array of zebra finch (*Taeniopygia guttata*) made of four 2.1M and one 3x720k array with 8,539,570 probes per array. Probe sizes were 50–75 mer in length and median probe spacing was 200 bp. Two different comparative (MeDIP vs. MeDIP) hybridization experiments were performed (two subarrays) for each experimental species (FUL, SCA, PAR, CRA) versus control FOR, with each subarray including hybridizations from MeDIP DNA from DNA pools from these different species (supplementary table S6, Supplementary Material online). For one subarray of each species, MeDIP DNA samples from the experimental groups were labeled with Cy5 and MeDIP DNA samples from the control lineage were labeled with Cy3. For the other subarray of each species, a dye swap was performed so that MeDIP DNA samples from the experimental groups were labeled with Cy3 and MeDIP DNA samples from the control lineage were labeled with Cy5.

For each comparative hybridization experiment, raw data from both the Cy3 and Cy5 channels were imported into R, checked for quality, and converted into MA values. The normalization procedure is as previously described (Guerrero-Bosagna et al. 2010). Following normalization each adjacent ≥ 3 probe set value represents the median intensity difference between FUL, SCA, PAR and CRA and control FOR of a 600-bp window. Significance was assigned to probe differences between experimental species samples and reference FOR samples by calculating the median value of the intensity differences as compared with a normal distribution scaled to the experimental mean and standard deviation of the normalized data. A Z score and P value were computed for each probe from that distribution. The statistically significant differential DNA methylation regions (DMR) were identified and P values associated with each region represented, as previously described (Guerrero-Bosagna et al. 2010).

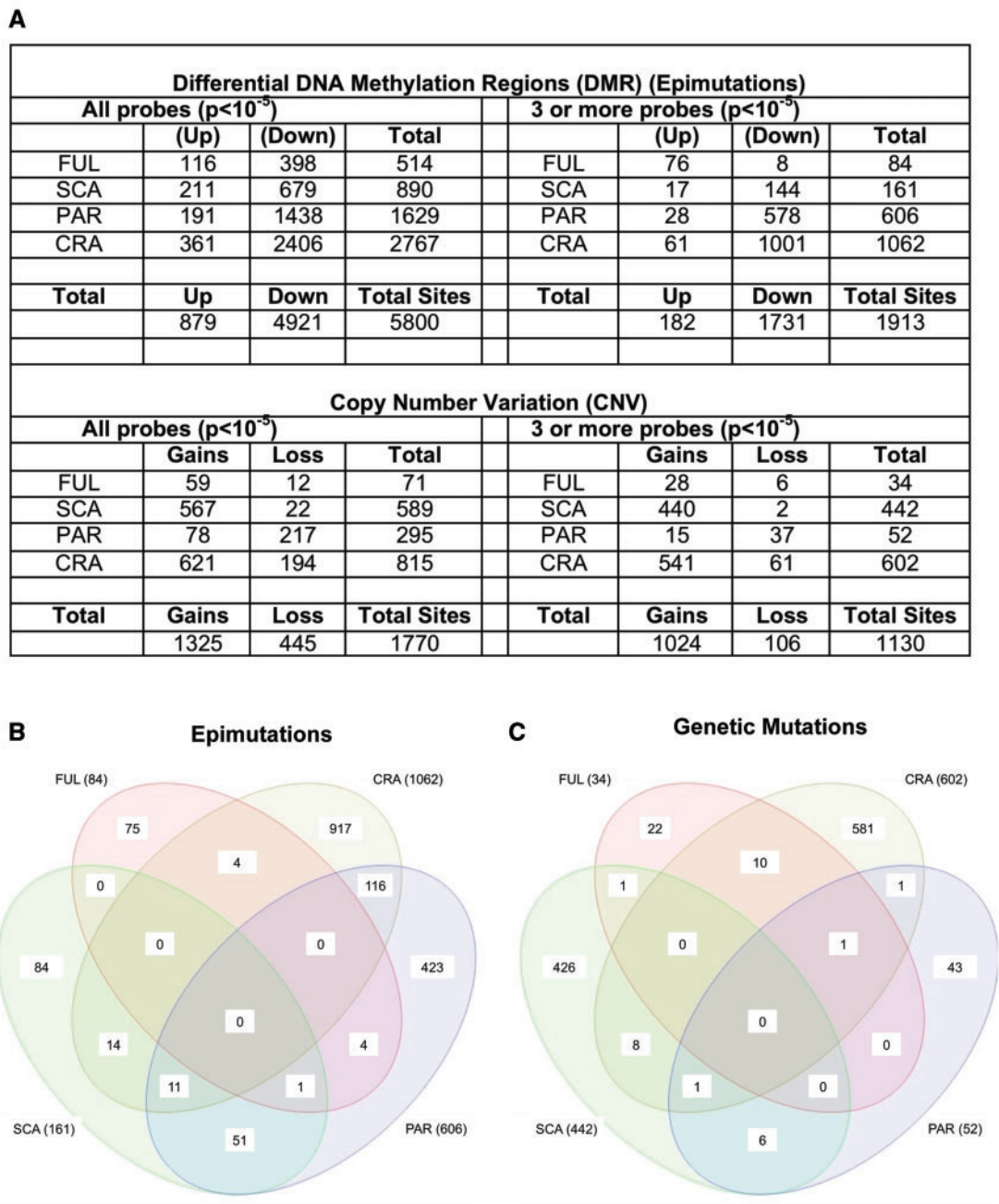


FIG. 2.—Number of epimutations and genetic mutations associated with Darwin’s finches. The number of differential DMR epimutations and CNV genetic mutations (A). DMR and CNV that differ significantly ($P < 10^{-5}$) from the reference species (FOR) are presented for all oligonucleotide probes, compared with peaks of three or more adjacent probes. The epimutations with an increase (Up) or decrease (Down) in DNA methylation are indicated. Those genetic mutations with an increase (Gain) or decrease (Loss) in CNV are indicated. Venn diagrams for epimutations (B) and genetic mutations (C) show overlaps between epimutations (DMR) and genetic mutations (CNV) among species. The species and total number of sites compared are listed on the outside of each colored elliptical.

Additional Bioinformatics and Statistics

The July 2008 assembly of the zebra finch genome (taeGut1, WUSTL v3.2.4) produced by the Genome Sequencing Center at the Washington University in St Louis (WUSTL) School of

Medicine was retrieved (WUSTL 2008). A seed file was constructed and a BSgenome package was forged for using the Finch DNA sequence in the R code (Herve Pages BSgenome: Infrastructure for Biostrings-based genome data packages. R

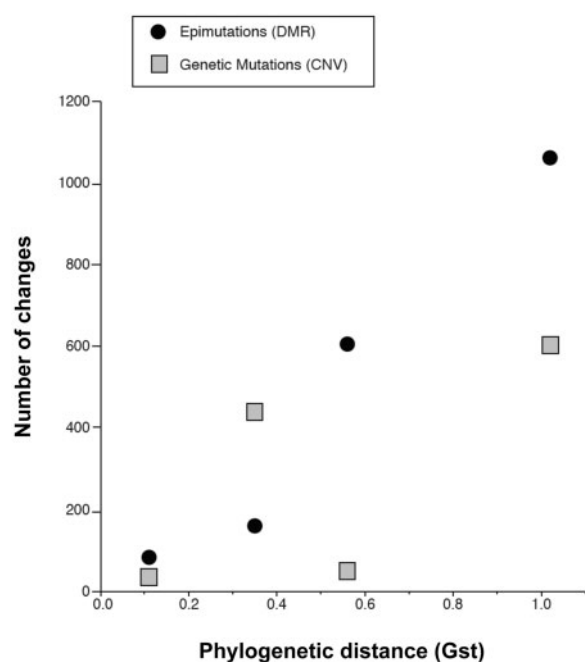


Fig. 3.—Phylogenetic distance is correlated with epigenetic changes, but not genetic changes. Branch lengths in figure 1 were used as measures of phylogenetic distance. The number of epimutations increased with phylogenetic distance (Spearman $Rho = 1.0$, $P < 0.0001$). In contrast, the number of genetic mutations did not increase with phylogenetic distance (Spearman $Rho = 0.8$, $P = 0.2$).

package version 1.24.0). This sequence was used to design the custom tiling arrays and to perform the bioinformatics.

The chromosomal location of CNV and DMR clusters used an R-code developed to find chromosomal locations of clusters (Skinner et al. 2012). A 2-Mb sliding window with 50,000 base intervals was used to find the associated CNV and DMR in each window. A Z-test statistical analysis with $P < 0.05$ was used on these windows to find the ones with overrepresented CNV and DMR were merged together to form clusters. A typical cluster region averaged approximately 3 Mb in size.

The DMR and CNV association with specific zebra finch genes and genome locations used the Gene NCBI database for zebra finch gene locations and correlated the epimutations associated (overlapped) with the genes. The three adjacent probes constituted approximately a 200-bp homology search. The KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway associations were identified as previously described (Skinner et al. 2012). Statistically significant overrepresentation uses a Fisher's exact analysis.

Spearman Rank correlation coefficients were used to test for a relationship between phylogenetic distance and epigenetic and genetic changes (Whitlock and Schluter 2009).

Results

Phylogenetic relationships of the five finch species in this study are shown in figure 1. The taxa chosen for this study included:

Two species of ground finches, FOR and FUL, which have crushing beaks with relatively deep bases; the cactus finch SCA, which has a long thin beak used for probing flowers; the small tree finch PAR, which has curved mandibles used for applying force at the tips; and the vegetarian finch CRA, which has a relatively short stubby bill used for crushing food along its entire length (Grant and Grant 2008; Donohue 2011; Rands et al. 2013). FOR was selected as a reference species for comparing genetic and epigenetic alterations among the remaining four species. Branch lengths in figure 1 were used as measures of phylogenetic distance.

The experimental design used purified erythrocytes from the different species. Although DNA sequences are the same for all cell types of an organism, the epigenome is distinct for each cell type, providing a molecular mechanism for the genome activity and functions that differ among different cell types (Skinner et al. 2010). Therefore, to investigate the overall epigenome requires a purified cell type. As birds have erythrocytes (red blood cells) that contain nuclei, samples of purified erythrocytes were collected from each of the Darwin's finch species to obtain DNA for molecular analysis.

The epigenetic alterations termed epimutations were assessed through the identification of differential DMR. The DMR were identified with the use of MeDIP with a methyl cytosine antibody, followed by a genome wide tiling array (Chip) for an MeDIP-Chip protocol (Guerrero-Bosagna et al. 2010). Although other epigenetic processes such as histone modifications, chromatin structure, and noncoding RNA are also important, DNA methylation is the best known epigenetic process associated with germline-mediated heritability and environmental manipulations (Skinner et al. 2010). Genetic variation was assessed using CNVs (i.e., amplifications and deletions of repeat elements) in the DNA using a CGH protocol (Pinkel and Albertson 2005; Gazave et al. 2011).

The reference genome used for the analysis was that of the zebra finch (*Taeniopygia guttata*) (Clayton et al. 2009), which had a preliminary estimate of greater than 83% similarity with a partial shotgun sequence of a Darwin's finch genome (Rands et al. 2013). This study actually suggests a much higher degree of identity. The zebra finch genome was tiled in a genome wide array with a 200-bp resolution and for a CGH array with a 1,500-bp resolution. These arrays were used in a competitive hybridization protocol between FOR (reference species) and the other four species (Guerrero-Bosagna et al. 2010). Differential hybridization using two different fluorescent DNA labeling tags identified the CNV with CGH using genomic DNA and the epimutation DMR with a MeDIP-Chip protocol. A statistical significance threshold of $P < 10^{-5}$ was set for the CNV or epimutation to be identified as a gain or loss compared with the reference species (fig. 2 and supplementary tables S1 and S2, Supplementary Material online). The data for all probes (oligonucleotides on the arrays) are presented. However, the criteria used to identify the CNV and DMR required the involvement of three or more adjacent

Darwin Finch Copy Number Variation (CNV) Against FOR Reference

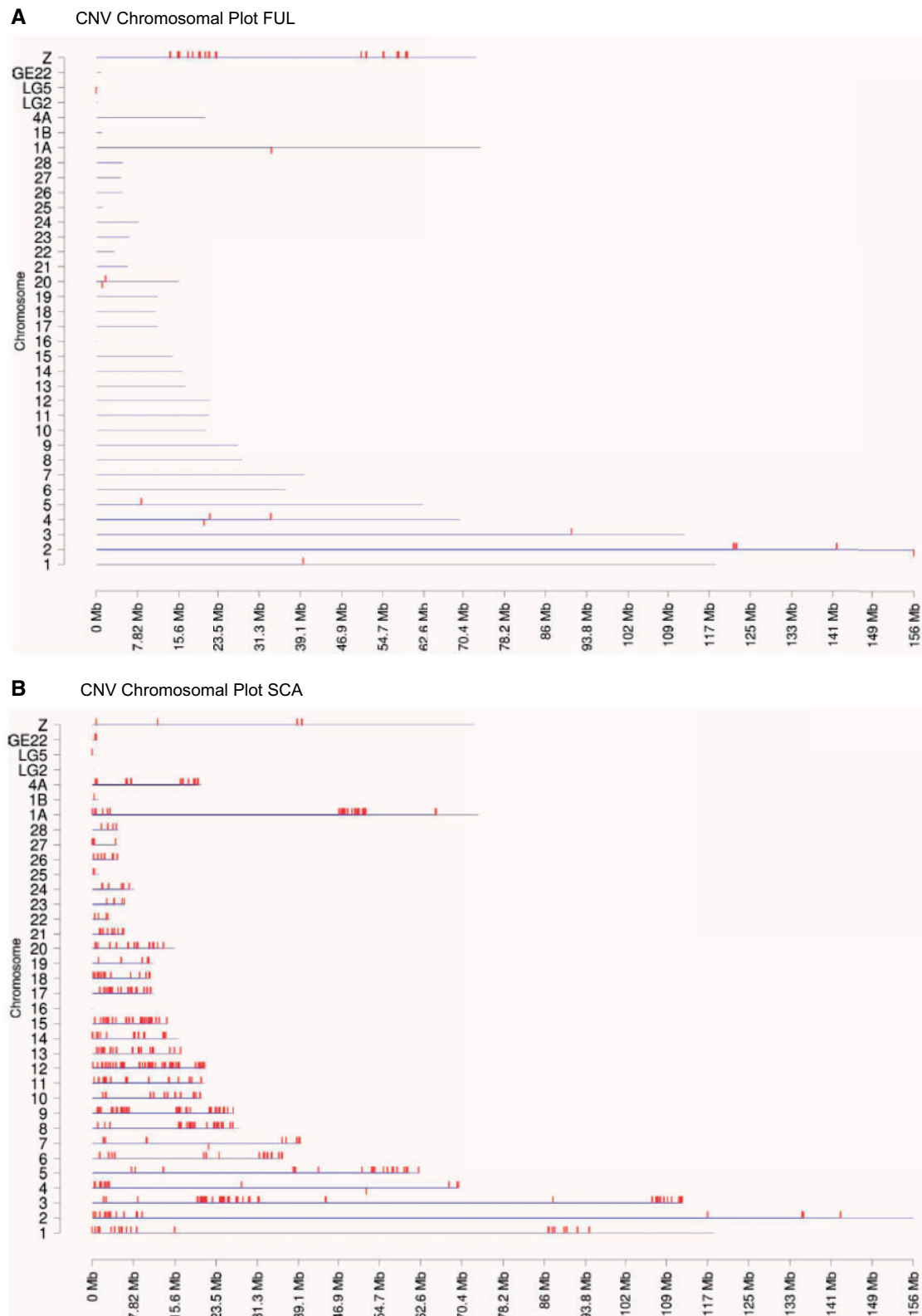


Fig. 4.—Chromosomal locations of the CNVs for each species. The chromosome number and size are presented in reference to the zebra finch genome. The chromosomal location of each CNV is marked with a red tick for FUL (A), SCA (B), PAR (C), and CRA (D).

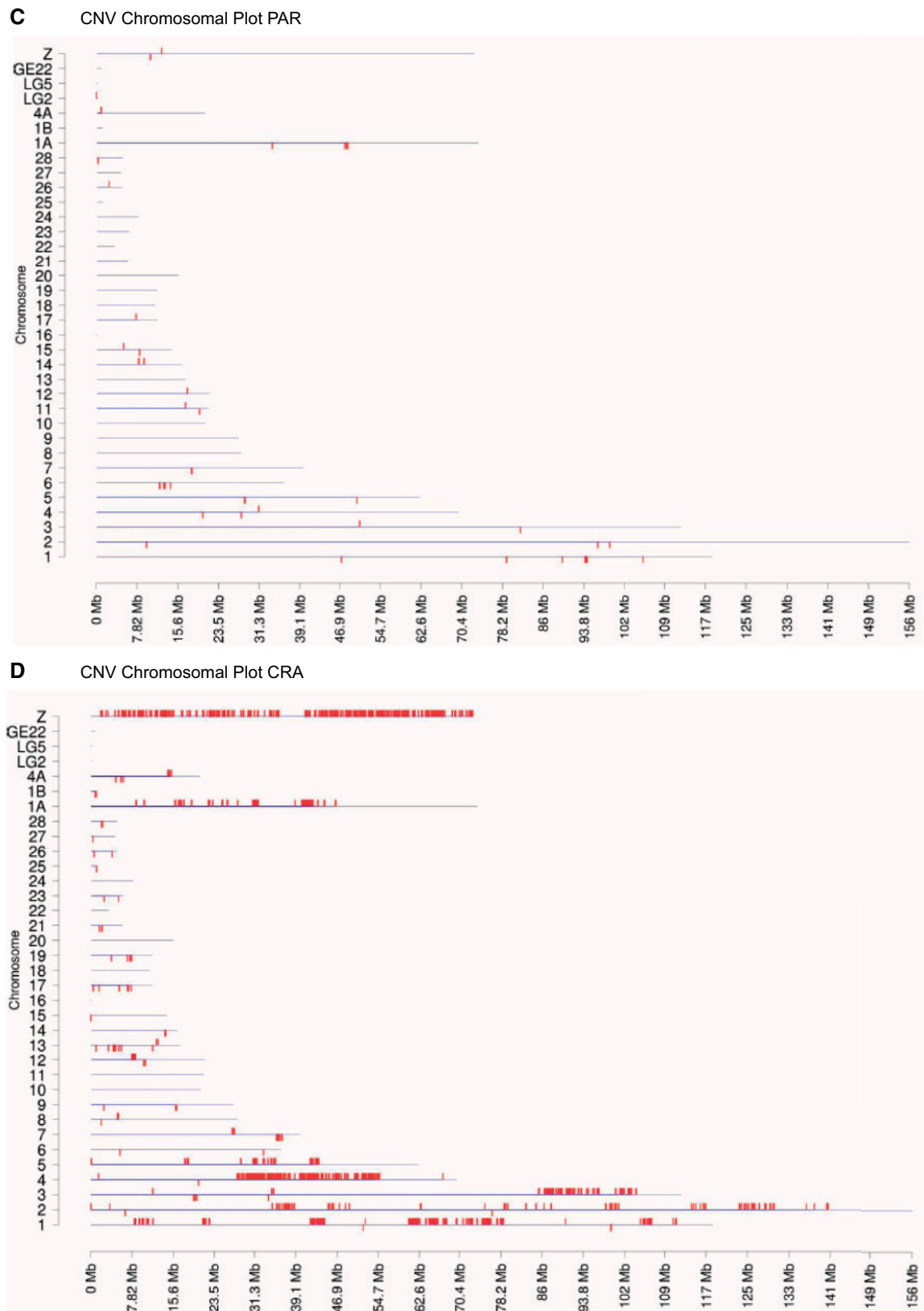


FIG. 4.—Continued.

Darwin Finch Differential DNA Methylation Regions (DMR) Epimutations Against FOR Reference

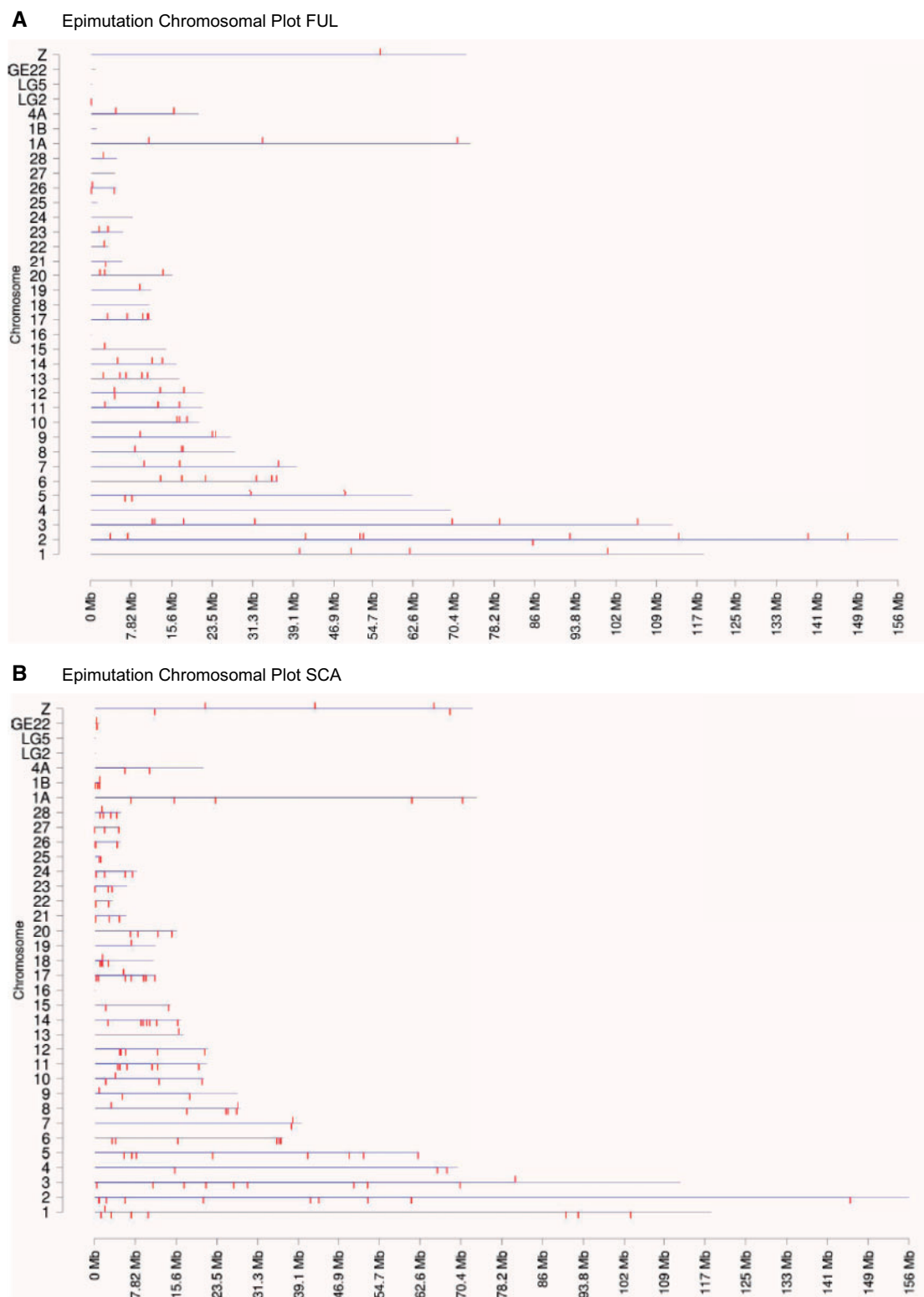
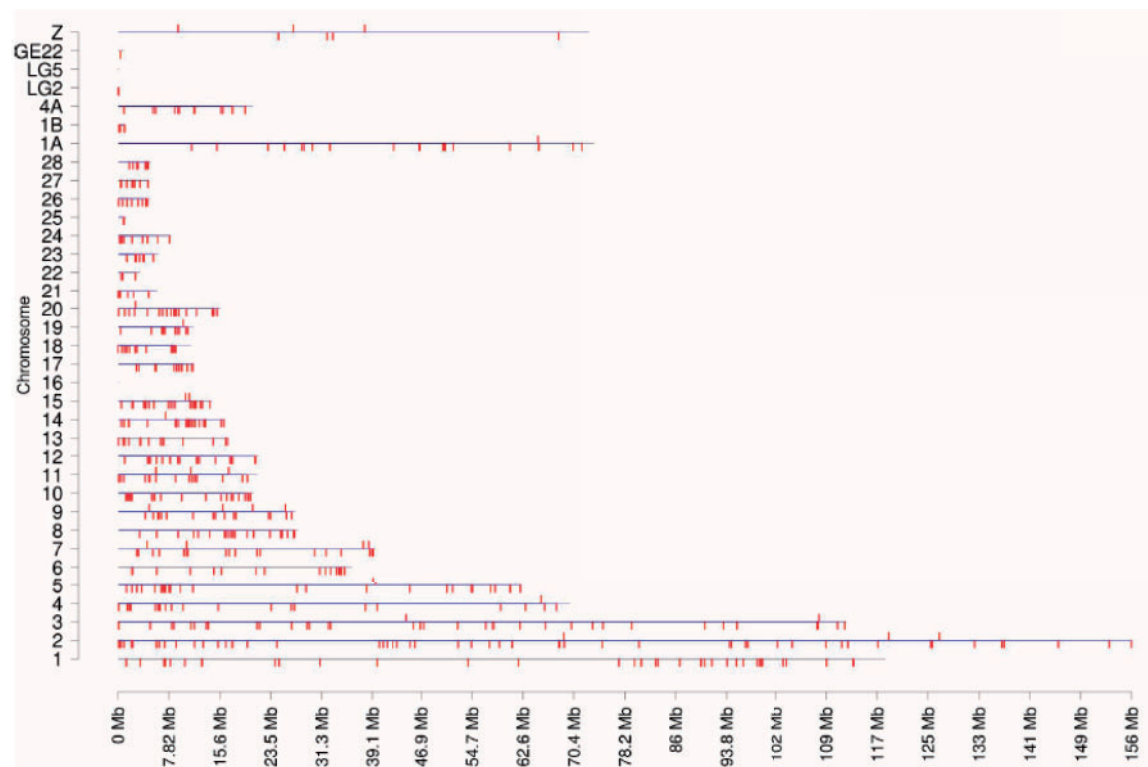


Fig. 5.—Chromosomal locations of the epimutations for each species. The chromosome number and size are presented in reference to the zebra finch genome. The chromosomal location of each DMR is marked with a red tick for FUL (A), SCA (B), PAR (C), and CRA (D).

C Epimutation Chromosomal Plot PAR



D Epimutation Chromosomal Plot CRA

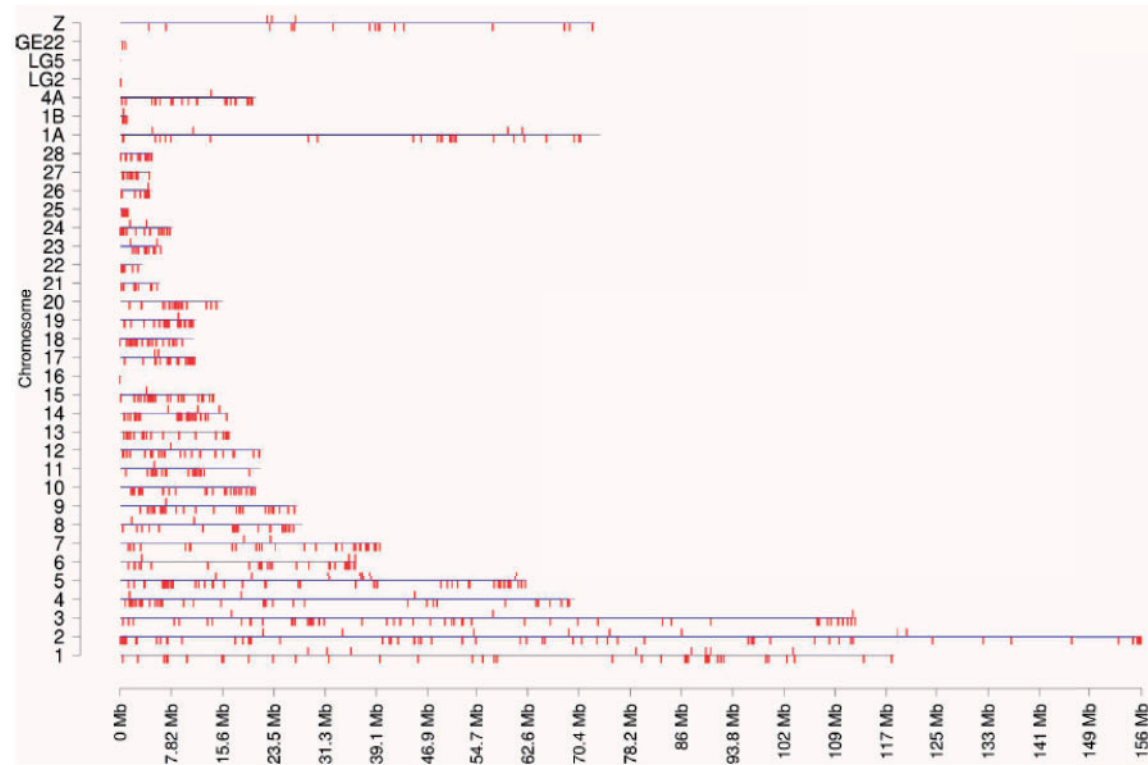


FIG. 5.—Continued.

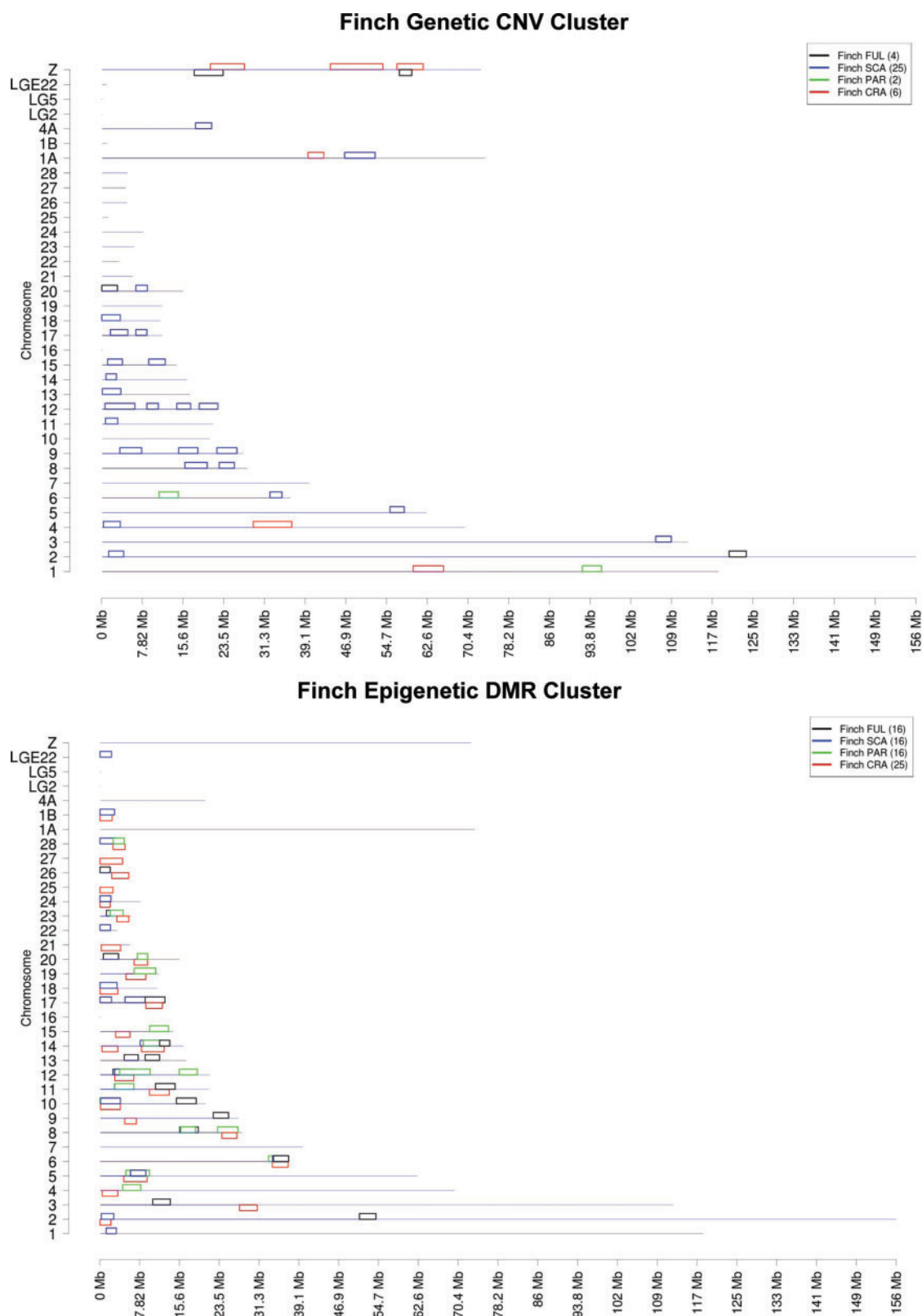


Fig. 6.—Chromosomal locations for clusters of CNV and DMR. The chromosome number and size are presented in reference to the zebra finch genome. The chromosomal location of statistically significant ($P < 10^{-5}$) overrepresented clusters of CNV (A) and DMR (B). The legend shows species and total number of clusters.

probes in the genome sequence having significant differential hybridization. These selection criteria reduce the number of false positives and provide a more reliable comparison (fig. 2). Therefore, the data presented used stringent criteria and represent the most reproducible epimutations and genetic CNV mutations among all three different experiments.

The increases or decreases in DNA methylation for the DMR are presented, along with the total number of epimutations in figure 2. The majority of epimutations for all species but FUL involves a decrease in DNA methylation (fig. 2A). The gains or losses in CNV are also presented, along with the total number of genetic alterations. The majority of genetic mutations for all species but PAR involves an increase in CNV number. Interestingly, the number of epimutations observed was generally higher, using the criteria selected, than the number of genetic alterations (fig. 2). However, the overall magnitude of epigenetic change was comparable to that of genetic change. Data for the five different species are shown in figure 1 for both epimutations (red) and genetic alterations (blue). The number of epimutations was significantly correlated with phylogenetic distance, whereas the number of genetic mutations was not (fig. 3).

The chromosomal locations of the CNV for the different finch species are shown in figure 4. CNVs were found on most chromosomes, with FUL having the least and CRA having the most. The chromosomal locations of the DMR epimutations for the different finch species are shown in figure 5. All chromosomes were found to have epimutations, with CRA having the highest number. These chromosomal plots suggested that some of the species might have clusters of CNV and/or DMR on some of the chromosomes (figs. 3 and 4). Therefore, a cluster analysis previously described (Skinner et al. 2012) was used to examine 50-kb regions throughout the genome to test for statistically significant ($P < 10^{-5}$) overrepresentation of CNV or DMR (fig. 6). Clusters, which have an average size of 3 Mb, are shown as species-specific boxes for CNV (fig. 6A) and for DMR (fig. 6B). Cluster characteristics and overlap are presented in [supplementary table S3, Supplementary Material](#) online. Clusters were obtained for all species, with a higher number of DMR clusters than CNV clusters. The highest number of CNV clusters was in SCA, with more than a 4-fold increase over CRA (fig. 6). Therefore, in addition to having more CNV than expected (assuming an increasing number with phylogenetic distance), SCA showed more CNV clusters than other species (fig. 2). Genome instability in these cluster regions may influence the increased numbers of CNV in SCA, which increases the presence of CNV clusters. In contrast, SCA did not show more DMR numbers or clusters than expected, assuming an increasing number with phylogenetic distance. Epimutation cluster overlap was more common among species (fig. 6 and table 1), suggesting that specific regions of the chromosomes were more susceptible to epigenetic alterations. Altered DNA methylation states have been experimentally shown to be stable for hundreds of

Table 1

Cluster Overlap between Species

CNVs				
CNV				
	FUL	SCA	PAR	CRA
FUL	4	0	0	2
SCA	0	25	0	0
PAR	0	0	2	0
CRA	2	0	0	6
Epimutations				
DMR				
	FUL	SCA	PAR	CRA
FUL	16	5	6	7
SCA	5	16	8	11
PAR	6	8	16	11
CRA	7	11	11	25

NOTE.—The overlap of CNV or DMR clusters between species is presented for the CNVs and epimutations.

generations (Cubas et al. 1999; Akimoto et al. 2007; Skinner et al. 2010).

The potential overlaps in specific CNV or DMR sites among species were examined. The overlap in genetic mutations among the four species is shown in a Venn diagram in figure 2C, whereas the overlap in epimutations is shown in figure 2B. No overlap in specific CNV or DMR sites was observed among all species, and less than 10% overlap was generally observed between any two species. Interestingly, the CNV overlap between FUL and CRA was higher than for the other species (fig. 2C). Generally, genetic and epigenetic alterations were distinct between species, with the majority being species specific. The epimutations showed more overlap between species than the genetic CNV mutations (fig. 2B and table 1). In considering within species overlap between the CNV and epimutations, less than 3% had common genomic locations. Therefore, the epimutations do not appear to be linked to the genetic CNV mutations, but are distinct.

The final analysis examined the potential functional significance of the epimutations by examining DMR and genes known to be associated with avian evolution. Several gene families and cellular signaling pathways have previously been shown to be involved in bird evolution, including the bone morphogenic protein (BMP) family and pathway (Abzhanov et al. 2004; Badyaev et al. 2008), the toll receptor family and signaling pathway (Alcaide and Edwards 2011), and the melanins family and pathway (Mundy 2005). All the genes associated with these signaling pathways were localized on the finch genome and compared with the genomic locations of the epimutations and CNV. Epimutation-associated genes within the BMP pathway (fig. 7), toll pathway (fig. 8), and

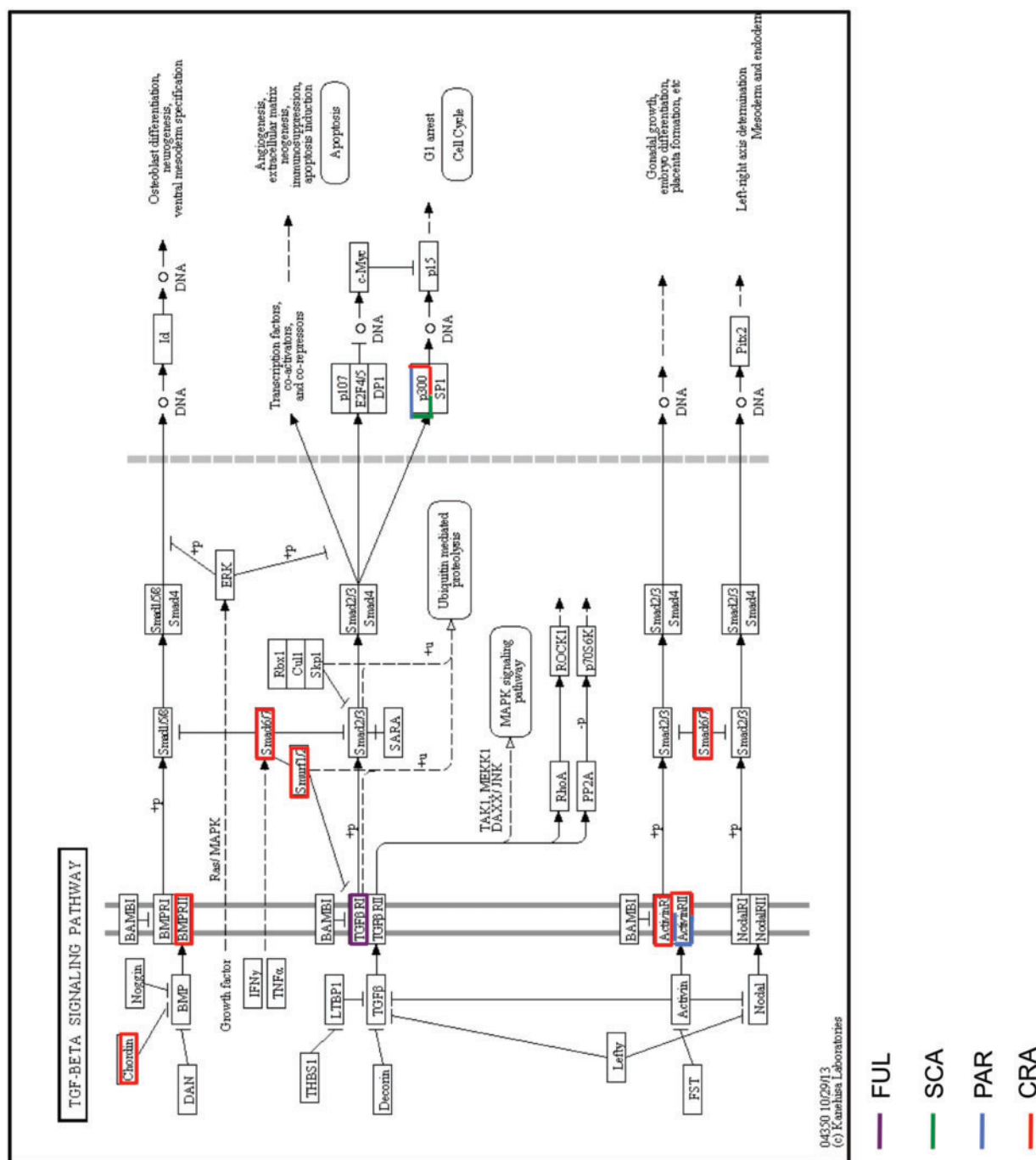


FIG. 7.—Epimutation-associated genes and correlated BMP pathway. The genes having associated epimutations in the signaling pathway presented for the different species are identified as FUL (purple), SCA (green), PAR (blue), and CRA (red) colored boxed genes.

melanin's pathway (fig. 9) are shown. Epimutations were overrepresented in all of these pathways (Fisher's exact test: BMP/TGFbeta (transforming growth factor) pathway, $P < 1 \times 10^{-6}$; toll pathway, $P < 5.7 \times 10^{-4}$; melanogenesis pathway, $P < 2.5 \times 10^{-13}$). Interestingly, the BMP pathway involved in beak development and shape had a statistically significant overrepresentation of CRA-associated epimutations

when examined independently ($P < 2.7 \times 10^{-5}$) (fig. 7). In addition, the toll receptor pathway involved in immune response had a statistically significant overrepresentation of PAR-associated epimutations when examined independently ($P < 7.7 \times 10^{-4}$) (fig. 8). The melanogenesis pathway involved in color had a mixture of epimutations from most of the species when examined independently ($P < 7 \times 10^{-5}$) (fig. 9).



1985

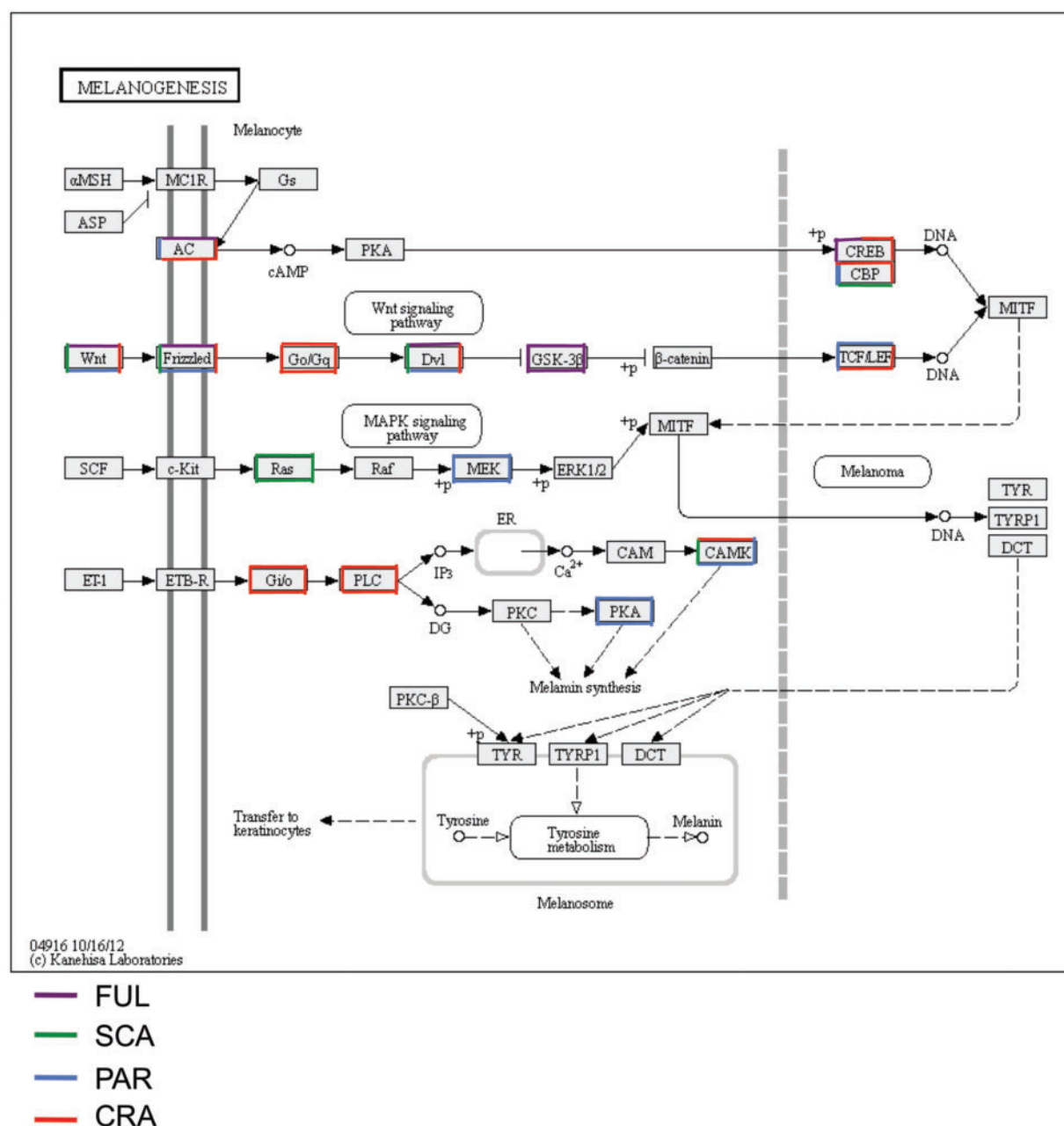


Fig. 9.—Epimutation-associated genes and correlated melanogenesis pathway. The genes having associated epimutations in the signaling pathway presented for the different species are identified as FUL (purple), SCA (green), PAR (blue), and CRA (red) colored boxed genes.

In addition to the pathway-specific genes, the total number of epimutations and CNV associated with genes are presented in table 2, with full lists in [supplementary tables S4 and S5](#), [Supplementary Material](#) online. The epimutations and CNV for single probe and ≥ 3 probe identification are presented in table 2. Observations indicate that approximately half of the epimutations and CNV identified were associated with genes. Therefore, a high percentage of the epimutations and CNV identified were associated with genes and were statistically overrepresented in several gene pathways

previously shown to be involved in particular aspects of avian evolution. Although this gene association analysis demonstrates that epimutations correlate with genes and important pathways, the functional or causal link to specific evolutionary processes remains to be investigated.

Discussion

This study provides one of the first genome-wide comparisons of genetic and epigenetic mutations among related species of

Table 2

Epimutation and CNV Gene Associations

CNVs				
	Total CNV 1+ Probes	Total CNV 3+ Probes	CNV Association with 14K Genes 1+ Probes	CNV Association with 14K Genes 3+ Probes
FUL	71	34	40	24
SCA	589	442	363	350
PAR	295	52	136	37
CRA	815	602	437	345
Epimutations				
	Total Epimutations 1+ Probes	Total Epimutations 3+ Probes	Epimutation Association with 14K Genes 1+ Probes	Epimutation Association with 14K Genes 3+ Probes
FUL	514	84	295	48
SCA	890	161	558	115
PAR	1,629	606	996	407
CRA	2,767	1,062	1,611	639

NOTE.—The 14,000 zebra finch genes annotated having epimutation or CNV associations are presented for the total number of associations (overlaps) for both regions identified with single (1+ probes) and adjacent (3+ probes) data sets.

organisms. There were relatively more epimutations than genetic CNV mutations among the five species of Darwin's finches, which suggests that epimutations are a major component of genome variation during evolutionary change. There was also a statistically significant correlation between the number of epigenetic differences and phylogenetic distance between finches (figs. 1 and 3), indicating that the number of epigenetic changes continues to accumulate over long periods of evolutionary time (2–3 Myr). In contrast, there was no significant relationship between the number of genetic CNV changes and phylogenetic distance.

The zebra finch genome was used as a reference for this study because a complete Darwin's finch genome is not yet available. The zebra finch genome showed hybridization with all probes on the array for each of the Darwin's finch species, suggesting that the genomes appear to be extremely similar. Loss of heterozygosity (absence of genomic regions, resulting in lack of probe hybridization) was not identified in any of the analyses. This suggests a high level of conservation and identity between the species' genomes. In the event the Darwin's finch genome has additional DNA sequence that is not present in the zebra finch genome, we would not have detected this DNA. Therefore, our data may be an underestimate of the Darwin's finch genome. Another technical limitation of our study was that we only considered genetic CNV (amplifications and deletions of repeat elements), but not other genetic variants such as point mutations or translocations. Although CNV frequency is higher than other mutations (e.g., SNPs) and stable in the genome (Gazave et al. 2011), this study's focus on CNV should be kept in mind. The epimutations examined are

differential DMR that have previously been shown to be frequent and transgenerationally stable (Anway et al. 2005; Guerrero-Bosagna et al. 2010; Skinner et al. 2010). Although other epigenetic processes such as histone modification, altered chromatin structure, and noncoding RNA may also be important, DNA methylation is the most established heritable epigenetic mark. This aspect of the experimental design should be kept in mind.

Among the five species of finches there were fewer genetic mutations (CNV) than epigenetic mutations. However, the cactus finch SCA showed a surprisingly large number of genetic CNV mutations than expected when compared with the reference species (FOR). The SCA mutations also clustered to similar locations on the genome to a greater extent than in the other species (fig. 6A). The reason for the disproportionately large number of CNV in the SCA comparison is unclear.

In contrast to the genetic mutation (CNV) analysis, the number of epimutations increased monotonically with phylogenetic distance (figs. 1 and 3). Overlap of specific epigenetic sites among species was minimal, including those for SCA (fig. 2B). An interesting possibility is that the epigenome may alter genome stability and generate genetic variation within species. A similar phenomenon has been shown for cancer, in which epigenetic alterations may precede genetic changes and alter genomic stability (Feinberg 2004). A decrease in the DNA methylation of specific repeat elements has previously been shown to correlate with an increase in CNV (Macia et al. 2011; Tang et al. 2012). Therefore, environmentally induced abnormal epigenetic shifts may influence genetic

mutations, such that a combination of epigenetics and genetics promotes phenotypic variation. Our observations demonstrate a relationship between the number of epigenetic changes and phylogenetic distance.

A comparison of the positions of epimutations and known gene families was also carried out. These gene families included those involved in the BMP pathway, which is related to beak shape (Badyaev et al. 2008), the toll receptor pathway, which is involved in immunological function (Alcaide and Edwards 2011), and the melanogenesis pathway, which affects color (Mundy 2005). Genes in all three of these families and signaling pathways were found to have species-specific epimutations (figs. 7–9). Future studies should focus on the causal relationship between epigenetic alterations and phenotypic traits.

Genetic mutations are postulated to provide much of the variation upon which natural selection acts (Gazave et al. 2011; Stoltzfus 2012). However, genetic changes alone are limited in their ability to explain phenomena ranging from the molecular basis of disease etiology to aspects of evolution (Skinner et al. 2010; Day and Bonduriansky 2011; Longo et al. 2012; Klironomos et al. 2013). Therefore, genetic mutations may not be the only molecular factors to consider (Richards 2006, 2009). Indeed, epigenetic and genetic changes may jointly regulate genome activity and evolution, as recent evolutionary biology modeling suggests (Day and Bonduriansky 2011; Klironomos et al. 2013). This integration of genetics and epigenetics may improve our understanding of the molecular control of many aspects of biology, including evolution.

Supplementary Material

Supplementary tables S1–S6 are available at *Genome Biology and Evolution* online (<http://www.gbe.oxfordjournals.org/>).

Acknowledgments

The authors acknowledge the advice and critical reviews of Dr Jeb Owen (WSU), Dr Hubert Schwabl (WSU), Dr David Crews (U Texas Austin), Dr Kevin P. Johnson (U Illinois) and Dr Sarah Bush (Utah). They thank Ms Sean Leonard, Ms Shelby Weeks, Dr C. Le Bohec, Mr O. Tiselma and Mr R. Clayton for technical assistance and Ms Heather Johnson for assistance in preparation of the manuscript. The research was supported by the National Institute of Health grants to M.K.S. and National Science Foundation grants to D.H.C. M.K.S. conceived the study. M.K.S. and D.H.C. designed the study. C.G.B., M.M.H., E.E.N., J.A.H.K., S.A.K., and D.H.C. performed the experiments and acquired the data. All authors analyzed the data. M.K.S. and D.H.C. wrote the manuscript. All authors edited and approved the manuscript. The authors declare no competing financial interests.

Literature Cited

- Abzhanov A, Protas M, Grant BR, Grant PR, Tabin CJ. 2004. Bmp4 and morphological variation of beaks in Darwin's finches. *Science* 305(5689):1462–1465.
- Akimoto K, et al. 2007. Epigenetic inheritance in rice plants. *Ann Bot*. 100(2):205–217.
- Alcaide M, Edwards SV. 2011. Molecular evolution of the toll-like receptor multigene family in birds. *Mol Biol Evol*. 28(5):1703–1715.
- Anway MD, Cupp AS, Uzumcu M, Skinner MK. 2005. Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science* 308(5727):1466–1469.
- Badyaev AV, Young RL, Oh KP, Addison C. 2008. Evolution on a local scale: developmental, functional, and genetic bases of divergence in bill form and associated changes in song structure between adjacent habitats. *Evolution* 62(8):1951–1964.
- Bonduriansky R. 2012. Rethinking heredity, again. *Trends Ecol Evol*. 27(6):330–336.
- Clayton DF, Balakrishnan CN, London SE. 2009. Integrating genomes, brain and behavior in the study of songbirds. *Curr Biol*. 19(18):R865–R873.
- Crews D, et al. 2007. Transgenerational epigenetic imprints on mate preference. *Proc Natl Acad Sci U S A*. 104(14):5942–5946.
- Cubas P, Vincent C, Coen E. 1999. An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* 401(6749):157–161.
- Day T, Bonduriansky R. 2011. A unified approach to the evolutionary consequences of genetic and nongenetic inheritance. *Am Nat*. 178(2):E18–E36.
- Donohue K. 2011. Darwin's finches: readings in the evolution of a scientific paradigm. Chicago (IL): University of Chicago Press, p. 492.
- Endler J. 1986. Natural selection in the wild. Princeton (NJ): Princeton University Press.
- Feinberg AP. 2004. The epigenetics of cancer etiology. *Semin Cancer Biol*. 14(6):427–432.
- Flatscher R, Frajman B, Schönschetter P, Paun O. 2012. Environmental heterogeneity and phenotypic divergence: can heritable epigenetic variation aid speciation? *Genet Res Int*. 2012:698421.
- Gazave E, et al. 2011. Copy number variation analysis in the great apes reveals species-specific patterns of structural variation. *Genome Res*. 21(10):1626–1639.
- Geoghegan JL, Spencer HG. 2012. Population-epigenetic models of selection. *Theor Popul Biol*. 81(3):232–242.
- Geoghegan JL, Spencer HG. 2013a. Exploring epiallele stability in a population-epigenetic model. *Theor Popul Biol*. 83:136–144.
- Geoghegan JL, Spencer HG. 2013b. The adaptive invasion of epialleles in a heterogeneous environment. *Theor Popul Biol*. 88:1–8.
- Geoghegan JL, Spencer HG. 2013c. The evolutionary potential of paramutation: a population-epigenetic model. *Theor Popul Biol*. 88:9–19.
- Grant P, Grant R. 2008. How and why species multiply: the radiation of Darwin's finches. Princeton (NJ): Princeton University Press.
- Greenspan RJ. 2009. Selection, gene interaction, and flexible gene networks. *Cold Spring Harb Symp Quant Biol*. 74:131–138.
- Guerrero-Bosagna C, Sabat P, Valladares L. 2005. Environmental signaling and evolutionary change: can exposure of pregnant mammals to environmental estrogens lead to epigenetically induced evolutionary changes in embryos? *Evol Dev*. 7(4):341–350.
- Guerrero-Bosagna C, Settles M, Lucker B, Skinner MK. 2010. Epigenetic transgenerational actions of vinclozolin on promoter regions of the sperm epigenome. *PLoS One* 5(9):e13100.
- Holeski LM, Jander G, Agrawal AA. 2012. Transgenerational defense induction and epigenetic inheritance in plants. *Trends Ecol Evol*. 27:618–626.
- Huber SK, et al. 2010. Ecoimmunity in Darwin's finches: invasive parasites trigger acquired immunity in the medium ground finch (*Geospiza fortis*). *PLoS One* 5(1):e8605.

- Huttley GA. 2004. Modeling the impact of DNA methylation on the evolution of BRCA1 in mammals. *Mol Biol Evol.* 21(9):1760–1768.
- Jirtle RL, Skinner MK. 2007. Environmental epigenomics and disease susceptibility. *Nat Rev Genet.* 8(4):253–262.
- Klironomos FD, Berg J, Collins S. 2013. How epigenetic mutations can affect genetic evolution: model and mechanism. *Bioessays* 35(6): 571–578.
- Koop JA, Huber SK, Laverty SM, Clayton DH. 2011. Experimental demonstration of the fitness consequences of an introduced parasite of Darwin's finches. *PLoS One* 6(5):e19706.
- Kuzawa CW, Thayer ZM. 2011. Timescales of human adaptation: the role of epigenetic processes. *Epigenomics* 3(2):221–234.
- Lack D. 1947. Darwin's finches. Cambridge University Press.
- Lamarck JB. 1802. Recherches sur l'organisation des corps vivans. Paris: Chez L'auteur, Maillard.
- Liebl AL, Schrey AW, Richards CL, Martin LB. 2013. Patterns of DNA methylation throughout a range expansion of an introduced songbird. *Integr Comp Biol.* 53(2):351–358.
- Longo G, Miquel PA, Sonnenschein C, Soto AM. 2012. Is information a proper observable for biological organization? *Prog Biophys Mol Biol.* 109(3):108–114.
- Lupski JR. 2007. An evolution revolution provides further revelation. *Bioessays* 29(12):1182–1184.
- Macia A, et al. 2011. Epigenetic control of retrotransposon expression in human embryonic stem cells. *Mol Cell Biol.* 31(2):300–316.
- Manikkam M, Guerrero-Bosagna C, Tracey R, Haque MM, Skinner MK. 2012. Transgenerational actions of environmental compounds on reproductive disease and epigenetic biomarkers of ancestral exposures. *PLoS One* 7(2):e31901.
- Mundy NI. 2005. A window on the genetics of evolution: MC1R and plumage colouration in birds. *Proc Biol Sci.* 272(1573):1633–1640.
- Nozawa M, Kawahara Y, Nei M. 2007. Genomic drift and copy number variation of sensory receptor genes in humans. *Proc Natl Acad Sci U S A.* 104(51):20421–20426.
- Olshen AB, Venkatraman ES, Lucito R, Wigler M. 2004. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics* 5(4):557–572.
- Petren K, Grand BR, Grant PR. 1999. A phylogeny of Darwin's finches based on microsatellite DNA length variation. *Proc R Soc Lond B.* 266(1417):321–329.
- Picard F, Robin S, Lavielle M, Vaisse C, Daudin J-J. 2005. A statistical approach for array CGH data analysis. *BMC Bioinformatics* 6:27.
- Pinkel D, Albertson DG. 2005. Comparative genomic hybridization. *Annu Rev Genomics Hum Genet.* 6:331–354.
- Poptsova M, Banerjee S, Gokcumen O, Rubin MA, Demichelis F. 2013. Impact of constitutional copy number variants on biological pathway evolution. *BMC Evol Biol.* 13:19.
- R Development Core Team. 2010. R: a language for statistical computing. Vienna (Austria): R Foundation for Statistical Computing. Available from: <http://www.R-project.org>.
- Rands CM, et al. 2013. Insights into the evolution of Darwin's finches from comparative analysis of the *Geospiza magnirostris* genome sequence. *BMC Genomics* 14:95.
- Rebollo R, Horard B, Hubert B, Vieira C. 2010. Jumping genes and epigenetics: towards new species. *Gene* 454(1–2):1–7.
- Richards CL, Bossdorf O, Pigliucci M. 2010. What role does heritable epigenetic variation play in phenotypic evolution? *BioScience* 60: 232–237.
- Richards EJ. 2006. Inherited epigenetic variation—revisiting soft inheritance. *Nat Rev Genet.* 7(5):395–401.
- Richards EJ. 2009. Quantitative epigenetics: DNA sequence variation need not apply. *Genes Dev.* 23(14):1601–1605.
- Skinner MK. 2011. Environmental epigenetic transgenerational inheritance and somatic epigenetic mitotic stability. *Epigenetics* 6(7): 838–842.
- Skinner MK, Anway MD, Savenkova MI, Gore AC, Crews D. 2008. Transgenerational epigenetic programming of the brain transcriptome and anxiety behavior. *PLoS One* 3(11):e3745.
- Skinner MK, Manikkam M, Guerrero-Bosagna C. 2010. Epigenetic transgenerational actions of environmental factors in disease etiology. *Trends Endocrinol Metab.* 21(4):214–222.
- Skinner MK, Mohan M, Haque MM, Zhang B, Savenkova MI. 2012. Epigenetic transgenerational inheritance of somatic transcriptomes and epigenetic control regions. *Genome Biol.* 13(10):R91.
- Skinner MK, Savenkova MI, Zhang B, Gore AC, Crews D. 2014. Gene bionetworks involved in epigenetic transgenerational inheritance of altered mate preference: environmental epigenetics and evolutionary biology. *BMC Genomics* 15:377.
- Slatkin M. 2009. Epigenetic inheritance and the missing heritability problem. *Genetics* 182(3):845–850.
- Stoltzfus A. 2012. Constructive neutral evolution: exploring evolutionary theory's curious disconnect. *Biol Direct.* 7:35.
- Sudmant PH, et al. 2013. Evolution and diversity of copy number variation in the great ape lineage. *Genome Res.* 23:1373–1382.
- Tang MH, et al. 2012. Major chromosomal breakpoint intervals in breast cancer co-localize with differentially methylated regions. *Front Oncol.* 2:197.
- Tateno H, Kimura Y, Yanagimachi R. 2000. Sonication per se is not as deleterious to sperm chromosomes as previously inferred. *Biol Reprod.* 63(1):341–346.
- Tibshirani R, Wang P. 2008. Spatial smoothing and hot spot detection for CGH data using the fused lasso. *Biostatistics* 9(1):18–29.
- Ward WS, Kimura Y, Yanagimachi R. 1999. An intact sperm nuclear matrix may be necessary for the mouse paternal genome to participate in embryonic development. *Biol Reprod.* 60(3):702–706.
- Whitlock MC, Schluter D. 2009. The analysis of biological data. Greenwood Village (CO): Roberts and Company Publishers.
- WUSTL. 2008. Jul. 2008 assembly of the zebra finch genome (taeGut1, WUSTL v3.2.4), as well as repeat annotations and GenBank sequences, Database Provider, NCBI.
- Ying H, Huttley G. 2011. Exploiting CpG hypermutability to identify phenotypically significant variation within human protein-coding genes. *Genome Biol Evol.* 3:938–949.

Associate editor: Bill Martin