A copy number variant is associated with a spectrum of pigmentation patterns in the rock pigeon (*Columba livia*)

Rebecca Bruders[1], Hannah Van Hollebeke[1], E.J. Osborne[2], Zev Kronenberg[2], Mark Yandell[2], Michael D. Shapiro[1,*]

[1] School of Biological Sciences, University of Utah, Salt Lake City, Utah, United States of America

[2] Department of Human Genetics, University of Utah, Salt Lake City, Utah, United States of America

[*] Corresponding author
Email: mike.shapiro@utah.edu

ABSTRACT

Rock pigeons (*Columba livia*) display an extraordinary array of pigment pattern variation. One such pattern, Almond, is characterized by a variegated patchwork of plumage colors that are distributed in an apparently random manner. Almond is a sex-linked, semi-dominant trait controlled by the classical *Stipper* (*St*) locus. Heterozygous males ($Z^{St}Z^{+}$ sex chromosomes) and hemizygous Almond females ($Z^{St}W$) are favored by breeders for their attractive plumage. In contrast, homozygous Almond males ($Z^{St}Z^{St}$) develop severe eye defects and lack all plumage pigmentation, suggesting that higher dosage of the mutant allele is deleterious. To determine the molecular basis of Almond, we compared the genomes of Almond pigeons to non-Almond pigeons and identified a candidate *St* locus on the Z chromosome. We found a copy number variant (CNV) within the differentiated region that captures complete or partial coding sequences of four genes, including the melanosome maturation gene *Mlana*. We did not find fixed coding changes in genes within the CNV, but all genes are misexpressed in regenerating feather bud collar cells of Almond birds. Notably, six other alleles at the *St* locus are associated with depigmentation phenotypes, and all exhibit expansion of the same CNV. Structural variation at *St* is linked to diversity in plumage pigmentation and gene expression, and thus provides a potential mode of rapid phenotypic evolution in pigeons.

AUTHOR SUMMARY

The genetic changes responsible for different animal color patterns are poorly understood, due in part to a paucity of research organisms that are both genetically tractable and phenotypically diverse. Domestic pigeons (*Columba livia*) have been artificially selected for many traits, including an enormous variety of color patterns that are variable both within and among different

breeds of this single species. We investigated the genetic basis of a sex-linked color pattern in pigeons called Almond that is characterized by a sprinkled pattern of plumage pigmentation. Pigeons with one copy of the Almond allele have desirable color pattern; however, male pigeons with two copies of the Almond mutation have severely depleted pigmentation and congenital eye defects. By comparing the genomes of Almond and non-Almond pigeons, we discovered that Almond pigeons have extra copies of a chromosome region that contains a gene that is critical for the formation of pigment granules. We also found that different numbers of copies of this region are associated with varying degrees of pigment reduction. The Almond phenotype in pigeons bears a remarkable resemblance to Merle coat color mutants in dogs, and our new results from pigeons suggest that similar genetic mechanisms underlie these traits in both species. Our work highlights the role of gene copy number variation as a potential driver of rapid phenotypic evolution.

INTRODUCTION

In natural populations of animals, pigment colors and patterns impact mate choice, signaling, mimicry, crypsis, and distraction of predators [1, 2]. In domestic animals, pigmentation traits are often selected by humans based on colors and patterns they find most attractive. Despite longstanding interest in the spectactular variation in color and pattern among animals, we know little about the molecular mechanisms that mediate color patterns. Understanding the genetic basis of the stunning array of animal color patterns benefits from the study of genetically tractable species; however, progress is hampered, in part, by a limited number of traditional model organisms that show limited variation in color and color patterning.

The domestic rock pigeon (*Columba livia*) is a striking example of variation shaped by artificial selection, with a multitude of colors and color patterns within and among more than 350

breeds. Because breeds of domestic pigeon belong to the same species and are interfertile, pigeons offer an exceptional opportunity to understand the genetic basis of pigmentation traits using laboratory crosses and genomic association studies [3]. Previously, we identified several genes involved in determining the type and intensity of plumage melanins in pigeons [4, 5], but considerably less is known about the molecular determinants of pattern deposition [6]. The molecular basis of pattern variation is an exciting frontier in pigmentation genetics, and recent work in other vertebrates reveals several genes that contribute to this process. Still, the genetic basis of pigment pattern is decidedly less well understood than the genes controlling pigment types [7–16].

The classical pigmentation pattern in *C. livia* known as Almond is caused by a semi-dominant mutation (*St* allele) at the sex-linked *Stipper* (*St*) locus [17] (Fig. 1). Unlike most other pigmentation pattern traits in pigeons, the variegated or sprinkled patchwork of plumage colors in Almond is apparently random within and among individuals [18]. Furthermore, the color pattern changes in an unpredictable manner with each molt [19–21]. The number of pigmented feathers in Almond pigeons also increases with each successive molt, and this effect is more pronounced in males [22, 23]. Notably, this phenomenon is the opposite of what is typically observed with pigmentation traits that change throughout the lifespan of an individual, such as vitiligo and graying, which result in a decrease in pigment over time [24–28]. In addition to Almond, at least six other alleles at *St* lead to varying degrees of depigmentation in pigeons, suggesting that the *St* locus might be a mutational hotspot [21, 29].

Heterozygous Almond males ($Z^{St}Z^{+}$) and hemizygous Almond females ($Z^{St}W$; males are the homogametic sex in birds), each of which have one copy of the *St* allele, are valued by breeders for their attractive color patterns. However, homozygous Almond males ($Z^{St}Z^{St}$) almost always

lack pigmentation in the first set of pennaceous feathers and have severe congenital eye defects [19, 30, 31] (Fig. 1B, C). The pattern of inheritance of Almond suggests that dosage of the mutant allele, rather than absence of the wild type allele, is responsible for the pigment and eye phenotypes in homozygous males. Eye defects are also associated with pigmentation traits in other vertebrate species, including dogs and horses, yet the molecular basis of these linked effects remains poorly understood [9, 32–37]. Therefore, Almond pigeons can illuminate links between pigmentation and eye defects, including whether pleiotropic effects of a single gene or linked genes with separate effects control these correlated traits.

In this study, we investigate the genomic identity of the *St* locus in domestic pigeons. Whole-genome sequence comparisons of Almond and non-Almond birds reveal a copy number variant (CNV) in Almond birds that includes the complete coding sequences of two genes, and partial coding sequences of two others. One of the complete genes, *Mlana*, plays a key role in the development of the melanosome (the organelle in which pigment granules are produced), making it a strong candidate for the pigmentation phenotype observed in Almond pigeons. We also find that different alleles at *St* are correlated with different degrees of expansion of the same CNV, thereby linking a spectrum of pigmentation variants to changes at one locus.

RESULTS

A sex-linked genomic region is associated with Almond pigmentation pattern

To determine the genomic location of the sex-linked *St* locus, we compared the genomes of 12 Almond pigeons to a panel of 109 non-Almond pigeons from a diverse set of breeds, using a probabilistic measure of allele frequency differentiation (pFst) [38] (see S1 Table for sample details). This whole-genome scan identified several significantly differentiated regions, but one

exceeded the others by several orders of magnitude and was located on a Z-chromosome scaffold (ScoHet5_227), as predicted from classical genetics studies (Fig. 2A). The differentiated region of ScoHet5_227 (position 5,186,219-5,545,482; peak SNP, *p*=1.1 e-16, genome wide significance threshold p = 5.5 e-10) contained eight annotated protein-coding genes, none of which had fixed coding changes in Almond compared to non-Almond genomes (VAAST [39]). Therefore, the Almond pigmentation pattern probably does not result from non-synonymous changes to protein-coding genes.

Several other scaffolds contained sequences that were significantly differentiated between Almond and non-Almond pigeons (Fig. 2A). All of these regions are autosomal, and we speculate that they are linked to other color traits that are often co-selected with Almond to give the most desirable sprinkled patchwork of colors, including T-check (a highly melanistic wing pattern), kite bronze (a deep reddening of the feathers), and recessive red (a pheomelanic color trait) [18, 21, 29]. However, because Almond is a sex-linked trait [17], we focused our attention on the Z-linked scaffold ScoHet5_227.

A copy number variant is associated with the Almond pigment pattern

In the absence of fixed coding changes between Almond and non-Almond birds, we next asked if birds with different phenotypes had genomic structural differences in the candidate region. We examined sequencing coverage on ScoHet5_227 and found that all 12 Almond genomes had substantially higher coverage in the Almond candidate region relative to non-Almond genomes, indicating the presence of a copy number variant (CNV) (Fig. 2B, C). The CNV captures a 77-kb segment of the reference genome (ScoHet5_227: 5,181,467-5,259,256), with an additional increase in coverage in a nested 25-kb segment (ScoHet5_227: 5,201,091-5,226,635). Read-depth

133 analysis confirmed 7 copies of the outer 77-kb segment and 14 copies of the inner 25-kb segment
134 in the genomes of female ($Z^{St}W$) Almond pigeons, which have an *St* locus on only one
135 chromosome. We used PCR to amplify across the outer and inner CNV breakpoints of Almond
136 pigeons and determined that the CNV consists of tandem repeats of the 77-kb and nested 25-kb
137 segments (Fig. 3B).

138 We then genotyped the CNV region in a larger sample of Almond pigeons and found a
139 significant association between the number of tandem repeats and the Almond phenotype (TaqMan
140 assay; pairwise Wilcoxon test, p= 2.0 e-16). Almost all Almond birds have more than one copy of
141 the CNV per Z-chromosome (n=78 of 80) (Fig. 4). Conversely, nearly all non-Almond birds had
142 only one copy per Z-chromosome (n=55 of 57). The two non-Almond birds with >1 copy per
143 chromosome had a maximum of one additional copy of the CNV, indicating that small increases
144 in copy number do not necessarily cause the Almond phenotype. Overall, these analyses suggest
145 that expansion of CNV on ScoHet5_227 is associated with the Almond phenotype.
146
147 **Genes within the CNV are misexpressed in Almond feather buds**
148 We next asked if the CNV was associated with gene expression changes between
149 developing Almond and non-Almond feathers. To address this question, we compared expression
150 of genes in the CNV region among birds with ($Z^{St}Z^{+}$, $Z^{St}W$, $Z^{St}Z^{St}$) and without ($Z^{+}Z^{+}$ and $Z^{+}W$)
151 Almond alleles. We analyzed Almond feather buds with dark and light pigmentation separately to
152 assess whether expression differed between qualitatively different feather pigmentation types, both
153 of which are present in $Z^{St}Z^{+}$ and $Z^{St}W$ Almond individuals. The CNV contains the complete
154 coding sequences of two genes, *Mlana* and *Slc16a7,* and partial coding sequences of two additional
155 genes, *Ermp1* and *Kiaa2026* (Fig. 3A). *Mlana* is predicted to have up to 14 total copies per Z^{St}

chromosome based on sequencing coverage in $Z^{St}W$ Almond birds (Fig. 3B). *Mlana* is expressed almost exclusively in melanocytes (melanin-producing cells), and encodes a protein that is critical for melanosome maturation through interactions with the matrix-forming protein Pmel [40–42]. Thus, the combination of the biological role of *Mlana* and its location in the Almond CNV makes *Mlana* a strong candidate gene for the Almond phenotype.

Compared to non-Almond feather buds, *Mlana* expression is increased in dark feather buds, but not in light feather buds, from $Z^{St}Z^{+}$ and $Z^{St}W$ Almond birds or the unpigmented feather buds of homozygous Almond ($Z^{St}Z^{St}$) birds (Fig. 5A; see S2 Table and S3 Table for raw data for all qRT-PCR experiments). We noticed that the variance of expression observed for *Mlana* in both dark and light Almond feather buds, though not statistically significant (Kolmogorov-Smirnov test), trends higher than in non-Almond samples. This data distribution might reflect the variability of the phenotype itself, which is characterized by different quantities and intensities of feather pigmentation both within and between $Z^{St}Z^{+}$ and $Z^{St}W$ Almond pigeons.

Other genes completely or partially within the CNV show increased expression in feathers from birds with at least one Almond allele relative to non-Almond birds. *Slc16a7* encodes a monocarboxylate transporter, and is predicted to be amplified to six full-length copies in Almond pigeons (Fig. 3B). We observed a 40-fold increase in expression of *Slc16a7* in Almond feather buds compared to non-Almond (Fig. 5A). *Slc16a7* is not known to be important in pigmentation; however, this gene is expressed in the mammalian, where it is involved in lactic acid transport and osmotic balance [43–47].

In addition to the two genes fully contained within the CNV, a novel fusion of *Ermp1* (a metallopeptidase gene) and *Kiaa2026* (unknown function) is predicted to span the outer CNV breakpoints (Fig. 3C). Neither gene is known to play a role in pigmentation or eye development.

The predicted Ermp1/Kiaa2026 fusion protein is a truncated version of Ermp1, including the peptidase domain and 3 of the 6 transmembrane domains (Fig. 3B). The 22 amino acids from Kiaa2026 at the C-terminus of the fusion protein do not include a known protein domain [48]; thus, the fusion protein is unlikely to create a novel combination of functional domains. As expected, the *Ermp1/Kiaa2026* fusion gene is not expressed in feathers of non-Almond birds, but is expressed in birds with Almond alleles (Fig. 5A). When we analyzed the expression of the exons of *Kiaa2026* and *Ermp1* located outside the CNV, we did not observe expression differences among genotypes (Fig. 5B). Therefore, the Almond CNV is associated with expression of the novel fusion gene, but not with expression differences in the full-length transcripts of either contributing gene. Similarly, *Ric1*, a gene immediately outside the CNV, shows a modest (less than two-fold) expression increase in light Almond feathers relative to other feather types (Fig. 5B). In summary, genes inside CNV show variable or increased expression in feathers from Almond birds, whereas genes adjacent to the CNV show little or no expression change.

Gene expression changes suggest melanocyte dysfunction in Almond feather buds

Plumage pigmentation patterns in $Z^{St}Z^{+}$, $Z^{St}W$, and $Z^{St}Z^{St}$ Almond birds are radically different than non-Almond birds, which led us to predict that other components of the melanogenesis pathway might differ as well. The production of melanin by melanocytes is a multi-step process that begins with activation of several pathways, including Wnt and Mc1r signaling, via extracellular ligands and agonists [49–52]. Subsequently, expression of transcription factors, including Mitf, activates a genetic cascade that ultimately promotes the maturation of a functional melanocyte [53]. Within the melanocyte itself, a series of enzymatic reactions and assembly of the melanosome leads to the production and deposition of pigments. Melanosomes are then transferred

to skin cells and epidermal appendages, including feathers. In pigeons and other birds with melanin-based pigments, the balance of pheomelanin (reds, yellows) and eumelanin (blacks, browns) deposition determines plumage color [54].

To determine if pigment production signals diverge between Almond and non-Almond feather buds, we measured expression of several marker genes for melanocyte maturation and function by qRT-PCR. We first examined genes involved in melanocyte survival and differentiation, both of which are critical early events in melanin production. *Sox10*, which encodes a transcription factor that activates expression of many downstream genes including *Mitf, Tyrosinase,* and *Tyrp1* expression [55], is downregulated only in light Almond and homozygous Almond feather buds (Fig. 5B). Because *Sox10* regulates *Mitf* and other melanocyte genes, this result indicates that melanocyte dysfunction occurs early in the lightly pigmented Almond feathers, but not in dark Almond feathers. A second melanocyte differentiation and survival marker gene, *Mitf,* encodes a transcription factor that activates expression of *Tyrosinase, Tyrp1, Pmel,* and *Mlana* [40, 53, 56, 57]. Unlike *Sox10*, *Mitf* is not differentially expressed in any of the phenotypes we tested (Fig. 5C). This result suggests that melanocytes are present in the feathers of all phenotypes, even in severely depigmented feathers [9]. Two genes that activate *Mitf* expression (*Sox10* and *Mc1r*, see below; Fig. 5C) are downregulated, which implies that *Mitf* would be downregulated as well. However, the persistence of high *Mitf* expression could be the result of activation by other pathways such as Wnt and c-Kit signaling [56]. Together, our gene expression results indicate that melanocytes are present in all feather buds of Almond pigeons (*Mitf* is expressed), but decreased *Sox10* expression in light and homozygous Almond feathers suggests multiple copies of the Almond CNV are associated with dysfunction early in melanogenesis in light and homozygous Almond feathers (*Sox10* expression is decreased).

We next assayed genes involved in pigment production, an indicator of melanocyte function. *Mc1r*, which encodes a G-protein-coupled receptor necessary for eumelanin production [58], and *Tyrosinase,* which encodes a critical enzyme for both eumelanin and pheomelanin production, were downregulated only in homozygous Almond feather buds (Fig. 5C). Therefore, expression of two key determinants of pigment production is affected only in the most severe depigmentation phenotype. *Tyrp1*, which encodes another enzyme important for eumelanin but not phenomelanin production [59], was downregulated in all Almond feather buds, with the most severe effects in light and homozygous Almond feather buds (Fig. 5C). Thus, the eumelanin synthesis pathway is affected in all Almond feathers, but pigment generation and melanocyte function genes are more impacted in light Almond and homozygous Almond feather buds, with the most severe downregulation observed in homozygotes (Fig. 5C).

Finally, we measured expression of the melanosome structure gene *Pmel*. Melanosome structure is thought to be necessary for eumelanin but not pheomelanin production [60]. Pmel is an amyloid protein that forms part of the melanosome matrix, an important structural component of the mature melanosome [60–62]. Our candidate gene *Mlana* encodes a protein that interacts with Pmel and is also critical for melanosome matrix formation. We found that *Pmel* is downregulated in all Almond feather buds, and most severely in the two most depigmented types, light Almond and homozygous Almond (Fig. 5C). As described above, *Mlana* expression increased in dark Almond feathers but was similar to non-Almond in light Almond and homozygous Almond feather buds. These results are difficult to reconcile because these two genes are regulated by Mitf. Nevertheless, our results show that even the pigmented feathers in Almond birds show altered expression of pigmentation genes.

In summary, in homozygous Almond feather buds, the pigmentation production pathway is altered at an early stage of eumelanogenesis. In birds with one copy of the Almond allele ($Z^{St}Z^{+}$ and $Z^{St}W$) light feathers show downregulation of more eumelanin production genes than do dark feathers. Thus, phenotypically different Almond feathers have distinct pigmentation gene expression profiles.

Other alleles at the *St* locus are copy number variants

Classical genetic studies point to multiple depigmentation alleles at the *St* locus [20, 29, 63, 64]. To determine if the Almond CNV is associated with these other alleles as well, we genotyped pigeons with other *St*-linked phenotypes and found significant increases in copy number in Qualmond (St^{Q}; N=10, p=8.3e-06), Sandy (St^{Sa}; N=3, p=3.2 e-02), Faded (St^{Fa}; N=11, p=5.0e-07), and Chalky (St^{C}; N=6, p=2.7e-04) pigeons compared to birds without *St*-linked phenotypes (Fig. 4, S4 Table). Another allele, Frosty (St^{fr}), showed a trend of copy number increase that did not reach significance (N=6, p=1). Together, these results demonstrate that copy number increase is associated with a variety of depigmentation alleles at the *St* locus.

We next asked whether different *St* alleles share the same CNV breakpoints. We amplified and sequenced across the Almond CNV breakpoints in Qualmond (N=4), Sandy (N=2), Faded (N=2), and Chalky (N=4) pigeons and found that the breakpoints are identical in all phenotypes tested. Therefore, a single initial mutational event was probably followed by different degrees of expansion in different *St* alleles. Notably, the breakpoints of the 77-kb segment (ScoHet5_227: 5,181,467 and 5,259,256) are enriched for CT repeats. These repeat sites could facilitate non-allelic homologous recombination, which could have generated the *St* allelic series [65].

DISCUSSION

Mlana is a strong candidate gene for the Almond phenotype

We identified a CNV associated with plumage pigmentation variation and an eye defect in domestic pigeons. Different numbers of copies of this structural variant are associated with a series of depigmentation alleles at the same locus. In the feathers of Almond birds, the CNV is associated with changes in the expression of genes within its bounds.

One of these genes, *Mlana,* is a strong candidate for Almond due to its role in melanosome maturation. *Mlana* and *Pmel* are co-regulated by Mitf and their protein products physically interact with each other during the process of matrix formation in the melanosome [41, 66]. Notably, *Pmel* mutations cause pigmentation phenotypes in cattle, chicken, and mouse [67–70]. *Pmel* mutations in horse, dog, and zebrafish result in both pigmentation phenotypes and eye defects, similar to Almond pigeons [37, 71–75]. For example, the merle coat pattern in dogs is associated with a transposon insertion in an intron of the *PMEL* gene, resulting in a non-functional PMEL protein and a phenotype that is remarkably similar to the Almond phenotype in pigeons [37, 72]. Dogs homozygous for the *PMEL* mutation, much like homozygous Almond pigeons, are severely hypopigmented. Additionally, homozygous *PMEL* mutant dogs have various eye defects, such as increased intraocular pressure, ametropia, microphthalmia, and coloboma [76]. The observation that Pmel, which interacts directly with Mlana, is repeatedly connected to both pigmentation and eye defects makes *Mlana* a strong candidate for similar correlated phenotypes in Almond pigeons. Likewise, in humans and mice, mutations in melanosome genes (e.g., *Oca2*, *Slc45a2*, *Slc24a5*) produce both epidermal depigmentation and eye defects, thereby further demonstrating a shared developmental link between these structures [77–79] .

292 The other full-length gene within the CNV, *Slc16a7,* does not have a known role in
293 pigmentation. However, this gene is a member of a class of monocarboxylate transporters that are
294 necessary to efficiently remove lactate from photoreceptor cells to prevent intracellular acidosis,
295 and to maintain a high glycolysis rate and proper cellular metabolism [44–46, 80, 81]. We
296 speculate that irregular expression of this gene could lead to cell death or dysfunction by causing
297 toxic lactic acid concentrations or by preventing lactic acid transport to nearby cells. In
298 regenerating Almond feathers, *Slc16a7* expression increases substantially (40-fold) relative to
299 non-Almond feathers, raising the possibility that this gene is somehow involved in pigmentation.
300 In short, changes in *Slc16a7* expression could drive components of the Almond phenotype in
301 feathers, eyes, or perhaps both. However, given the linked pigment and eye phenotypes observed
302 in *Pmel* mutants in other species, *Mlana* alone could be sufficient to induce both pigmentation and
303 eye defects in Almond pigeons. Future work will explore these various possibilities.
304
305 **Gene expression is altered in Almond birds**
306 In other organisms, copy number variation can result in gene expression changes in the same
307 direction as the copy number change (i.e., the presence of more copies is correlated with higher
308 expression) [82–84]. We observed a similar trend of higher expression of genes captured in the
309 Almond-linked CNV (Fig. 5A). In contrast to this trend, however, *Mlana* showed an increase in
310 expression in dark Almond feathers, but not in light Almond or homozygous Almond
311 (unpigmented) feathers. *Mlana* is also the gene with the greatest copy number increase, with up to
312 14 copies in hemizygous Almond genomes and 28 copies in the homozygous Almond genome.
313 With the above observations of gene expression in mind, why might homozygous Almond
314 birds lack *Mlana* expression in feather buds when they have 28 copies of the gene? One possibility

315 is epigenetic silencing. High copy numbers in tandem arrays induce gene silencing in several
316 organisms [85–89]. In fruit flies, for example, tandem arrays lead to variegated gene expression of
317 the white eye gene [86]. This change in expression, in turn, leads to mosaic eye color, a scenario
318 reminiscent of the color mosaicism in the feathers of Almond pigeons. In mouse, experimentally
319 reducing the number of copies of *lacZ* in a tandem array causes an increase in gene expression,
320 indicating that reducing copy number may relieve gene silencing [88]. Likewise, it is possible that
321 somatic copy number decrease could relieve gene silencing and restore higher expression of *Mlana*
322 in dark Almond feather buds.

323 Another potential explanation for the lack of *Mlana* expression in homozygous Almond
324 feathers is cell death or immunity-mediated destruction of melanocytes. Overexpression of *Mlana*
325 could have a toxic effect on cells, leading to cell death before melanocyte maturation. Similarly,
326 in humans, overexpression of genes is often associated with disease [90–92], and in yeast,
327 overexpression of genes can reduce growth rate [93]. Alternatively, Almond melanocytes might
328 elicit an autoimmune response, similar to the destruction of melanocytes in human pigmentation
329 disorders. MLANA is a dominant antigenic target for the T cell autoimmune response in human
330 skin affected by vitiligo [94, 95], and perhaps the presentation of Mlana antigens in Almond
331 pigeons elicits a response that depletes melanocytes in the developing feather buds. A potentially
332 analogous autoimmune response depletes the melanocyte population and mimics vitiligo in Smyth
333 line chickens [96].

334 If genes in the CNV are being randomly silenced in Almond pigeons, or cells with high
335 expression are escaping cell death in a random manner, then we might expect to see high variance
336 in gene expression among Almond feather samples. Consistent with this prediction, the variance
337 in expression of *Mlana* in both dark and light Almond feather buds trends higher than in non-

Almond samples (Fig. 5A). This variance might also explain the random pattern of pigmentation and de-pigmentation observed in the feathers of these birds. If each cell population is affected differently due to stochastic events resulting in differential expression, then random pigmentation patterns could be the outcome.

CNVs as mechanisms for the rapid generation of new phenotypes

In addition to finding a CNV at the *St* locus in Almond birds, we found quantitative variation in copy number among other alleles at this locus. Variation at this CNV may have a quantitative effect on de-pigmentation, with the degree of copy number increase correlating with degree of depigmentation and eye defects. For example, pigeon breeders report that Sandy and Whiteout – two phenotypes with among the highest numbers of copies of the CNV (Fig. 4) – have associated eye defects similar to Almond (Tim Kvidera, personal communication) [29, 63]. Although we currently have a small sample size of other *St*-linked phenotypes, we see a trend that other alleles produce milder pigment phenotypes and have less CNV expansion than the Almond allele. Similar quantitative effects of CNVs occur in other organisms as well, including a correlation between comb size and copy number of *Sox5* intron 1 in chickens [97].

Pigeon breeders have reported that parents with one *St*-linked phenotype can produce offspring of another phenotype in the *St* series [29, 98]. Specifically, Faded, Qualmond, and Hickory pigeons have produced Almond offspring. These classical breeding studies suggest that allelic conversion can occur rapidly and, based on our finding of copy number variation among *St* alleles, may result from simple expansion or contraction of a CNV. In another striking similarity between Merle dogs and Almond pigeons, germline expansions or contractions of the *Merle* allele of *PMEL* result in a spectrum of coat pattern phenotypes that can differ between parents and

361 offspring [72, 99]. Thus, unstable CNVs like the one we found at the *St* locus may provide a
362 mechanism for extraordinarily rapid phenotypic diversification in pigeons and other organisms
363 [100–103].
364
365 **MATERIALS & METHODS**
366 **Animal husbandry**
367 Animal husbandry and experimental procedures were performed in accordance with
368 protocols approved by the University of Utah Institutional Animal Care and Use Committee
369 (protocols 10-05007, 13-04012, and 16-03010).
370
371 **DNA sample collection and extraction**
372 Blood samples were collected in Utah at local pigeon shows, at the homes of local pigeon
373 breeders, and from pigeons in the Shapiro lab. Photos of each bird were taken upon sample
374 collection for our records and for phenotype verification. Breeders outside of Utah were contacted
375 by email to obtain feather samples. Breeders were sent feather collection packets and instructions,
376 and feather samples were sent back to the University of Utah along with detailed phenotypic
377 information and genetic relatedness. DNA was then extracted from blood, as previously described
378 [4]. DNA from feathers was extracted using the user developed protocol for Purification of total
379 DNA from nails, hair, or feathers using the DNeasy Blood & Tissue Kit (Qiagen Sciences,
380 Germantown, MD).
381
382 **Genomic analyses**

383 BAM files from a panel of previously resequenced birds were combined with BAM files
384 derived from new sequences from 11 Almond females and 16 non-Almond birds aligned to the
385 Cliv_2.1 genome assembly [104] (new sequence accessions: SRA SRP176668, accessions
386 SRR8420387-SRR8420407 and SRR9003406-SRR9003411; BAM files created as described
387 previously [6]). SNVs and small indels were called using the Genome Analysis Toolkit (Unified
388 Genotyper and LeftAlign and TrimVariants functions, default settings [105]). Variants were
389 filtered as described previously [38] and the subsequent variant call format (VCF) file was used
390 for downstream analyses.

391 Whole genomes of 12 Almond and 96 non-Almond birds were tested for allele frequency
392 differentiation using pFst (VCFLIB software library, https://github.com/vcflib; see S1 Table for
393 sample information) [38]. For analysis of fixed coding changes, VAAST 2.0 [39] was used to
394 conduct an association test and to search for putative disease-causing genetic variants common to
395 all Almond individuals but absent from non-Almonds. Annotated variants from affected
396 individuals were merged by simple union into a target file. The background file included variants
397 from 66 non-Almond birds, while the target file contained variants from the 12 Almond birds. This
398 VAAST analysis revealed that there were no fixed genetic variants among the Almond individuals
399 that were absent in the background dataset.

400

401 **CNV breakpoint identification and read-depth analysis**

402 Read depth in the CNV-containing region was analyzed in 12 Almond and 118 non-
403 Almond resequenced whole genomes. Scaffold ScoHet5_227 gdepth files were generated using
404 VCFtools [108]. Read depth was normalized using a region (scaffold ScoHet5_227: 1-5,000,000)
405 that did not show an increase in sequencing coverage in Almond genomes.

406 To determine the CNV breakpoints, we first identified the region of increased sequencing
407 coverage in Almond genomes using the depth function in VCFtools [108]. Next, we examined
408 BAM files of Almond genomes in IGV [110] in the region of coverage increase, and identified
409 locations at which reads were consistently split (did not map contiguously). These locations were
410 the putative breakpoints. We then designed PCR primers that amplify 1-kb products spanning the
411 putative breakpoints (see S5 Table for primer sequences). Finally, we used PCR to amplify across
412 the putative breakpoints. PCR products were purified and sequenced, and aligned to the pigeon
413 genome assembly using Blast+ version 2.7.1 [109]. The CNV breakpoint primers (see Fig. 3B)
414 successfully amplified products in 40 of 43 Almond pigeons tested.
415
416 **Fusion gene analysis**
417 The putative mRNA sequence of the *Ermp1/Kiaa2026* fusion gene was determined by
418 concatenating the mRNA sequence of the exons on one side of the outer breakpoint with the exons
419 that map to the outer breakpoint. The fusion of these exons was confirmed using exon spanning
420 primers and qPCR (See S5 Table for primer sequences). The putative mRNA sequence was
421 translated, and then analyzed for domains using HMMER searches in SMART (Simple Modular
422 Architecture Research Tool) [48]. We searched for domains in the SMART database, and also
423 searched for outlier homologs, PFAM domains, signal peptides, and internal repeats.
424
425 **Taqman assay for copy number estimates**
426 Copy number variation was estimated using a custom Taqman Copy Number Assay
427 targeted to the *Mlana* region (MLANA_CCWR201) for 150 Almond, 9 Qualmond, 3 Sandy, 14
428 Faded, and 6 Chalky, 5 Frosty, and 56 individuals without *St*-linked phenotypes. Following DNA

extraction, samples were diluted to 5 ng/uL and run in quadruplicate according to manufacturer's protocol. Copy number was determined using CopyCaller Software v2.1 (ThermoFisher Scientific, Waltham, MA). An intron in *RNaseP* was used for normalization of copy number.

RNA isolation and cDNA synthesis

To assay gene expression, secondary covert wing feathers were plucked to stimulate regeneration and allowed to regenerate for 9 days (see S2 Table for sample details). Nine-day regenerating feather buds were plucked, then the proximal 5 mm was cut and stored in RNA later at 4°C overnight. Feather buds were then dissected and collar cells removed, and stored at -80°C until RNA isolation. RNA was then isolated and reverse transcribed to cDNA as described previously [4].

qRT-PCR analysis

cDNA was amplified using intron-spanning primers for the appropriate targets using a CFX96 qPCR instrument and iTaq Universal Sybr Green Supermix (Bio-Rad, Hercules, CA) (S5 Table). Samples were run in duplicate and normalized to β-actin (see S3 Table for raw results). Results were compared in R [111] using ANOVA, followed by a Tukey post hoc test to determine differences between phenotypic groups. Differences were considered statistically significant if $p < 0.05$. Primers used for each gene are included in S5 Table.

ACKNOWLEDGEMENTS

We thank past and present members of the Shapiro lab for assistance with sample collection and processing; members of the Utah Pigeon Club and National Pigeon Association for sample

452 contributions; and Ken Davis and Tim Kvidera for critical discussions and advice. We thank Tim
453 Kvidera for photographs of Whiteout, Sandy, Frosty, Faded, and Chalky pigeons in Figure 4. We
454 thank Anna Vickrey, Max Sidesinger, Elena Boer, Emily Maclary, Sara Young, and Robert
455 Greenhalgh, for technical assistance and advice. This work was supported by the National Science
456 Foundation (CAREER DEB-1149160 to M.D.S.; GRF 1256065 to R.B.) and the National
457 Institutes of Health (R01GM115996 and R35GM131787 to M.D.S., R01GM104390 to M.Y.,
458 fellowship T32GM007464 to Z.K.). The funders had no role in study design, data collection and
459 analysis, decision to publish, or preparation of the manuscript. We acknowledge a computer time
460 allocation from the Center for High Performance Computing at the University of Utah.

461

462 **REFERENCES**

463 1. Protas ME, Patel NH. (2008) Evolution of Coloration Patterns. Annual Review of Cell and
464 Developmental Biology 24:425–471. doi:10.1146/annurev.cellbio.24.110707.175302.
465 2. Roberts NW, Mappes J, Arbuckle K, et al. (2017) The biology of color. Science
466 357:eaan0221. doi:10.1126/science.aan0221.
467 3. Domyan ET, Shapiro MD. (2017) Pigeonetics takes flight: Evolution, development, and
468 genetics of intraspecific variation. Developmental Biology 427:241–250.
469 doi:10.1016/J.YDBIO.2016.11.008.
470 4. Domyan ET, Guernsey MW, Kronenberg Z, et al. (2014) Epistatic and combinatorial
471 effects of pigmentary gene mutations in the domestic pigeon. Current Biology 24:459–
472 464. doi:10.1016/j.cub.2014.01.020.
473 5. Guernsey MW, Ritscher L, Miller MA, Smith DA, Schöneberg T, Shapiro MD. (2013) A
474 Val85Met Mutation in Melanocortin-1 Receptor Is Associated with Reductions in

Eumelanic Pigmentation and Cell Surface Expression in Domestic Rock Pigeons (Columba livia). PLoS ONE 8:1–9. doi:10.1371/journal.pone.0074475.

6. Vickrey AI, Bruders R, Kronenberg Z, et al. (2018) Introgression of regulatory alleles and a missense coding mutation drive plumage pattern diversity in the rock pigeon. eLife 7:e34803. doi:10.7554/eLife.34803.

7. 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 365:2439–2450. doi:10.1098/rstb.2010.0104.

8. Rosenblum EB, Parent CE, Brandt EE. (2014) The Molecular Basis of Phenotypic Convergence. Annual Review of Ecology, Evolution, and Systematics 45:203–226. doi:10.1146/annurev-ecolsys-120213-091851.

9. Cieslak M, Reissmann M, Hofreiter M, Ludwig A. (2011) Colours of domestication. Biological Reviews 86:885–899. doi:10.1111/j.1469-185X.2011.00177.x.

10. Hoekstra HE. (2006) Genetics, development and evolution of adaptive pigmentation in vertebrates. Heredity 97:222–234. doi:10.1038/sj.hdy.6800861.

11. Mallarino R, Henegar C, Mirasierra M, Manceau M, Schradin C, Vallejo M, Beronja S, Barsh GS, Hoekstra HE. (2016) Developmental mechanisms of stripe patterns in rodents. Nature 539:518–523. doi:10.1038/nature20109.

12. Parichy DM. (2003) Pigment patterns: fish in stripes and spots. Current Biology 13:R947-50. doi:10.1016/J.CUB.2003.11.038.

13. Salmon Hillbertz NHC, Isaksson M, Karlsson EK, et al. (2007) Duplication of FGF3, FGF4, FGF19 and ORAOV1 causes hair ridge and predisposition to dermoid sinus in

Ridgeback dogs. Nature Genetics 39:1318–1320. doi:10.1038/ng.2007.4.

14. Brunberg E, Andersson L, Cothran G, Sandberg K, Mikko S, Lindgren G. (2006) A missense mutation in PMEL17 is associated with the Silver coat color in the horse. BMC Genetics 7:46. doi:10.1186/1471-2156-7-46.

15. Roberts RB, Moore EC, Kocher TD. (2017) An allelic series at pax7a is associated with colour polymorphism diversity in Lake Malawi cichlid fish. Molecular Ecology 26:2625–2639. doi:10.1111/mec.13975.

16. Parichy DM, Spiewak JE. (2015) Origins of adult pigmentation: diversity in pigment stem cell lineages and implications for pattern evolution. Pigment Cell and Melanoma Research Melanoma Research 28:31–50. doi:10.1111/pcmr.12332.

17. Wriedt C, Christie W. (1925) Zur Genetik der gesprenkelten Haustaube. Induktive Abstammungs- und Vererbungslehre 38:271–306. doi:10.1007/BF02118234.

18. Sell A. (2012) Pigeon genetics: applied genetics in the domestic pigeon. Sell Publishing, Achim, Germany

19. Hollander WF, Cole LJ. (1940) Somatic Mosaics in the Domestic Pigeon. Genetics 25:16–40.

20. Hollander WF. (1942) Auto-sexing in the domestic pigeon. Journal of Heredity 33:135–140. doi:10.1093/oxfordjournals.jhered.a105150.

21. Quinn JW. (1971) The Pigeon Breeders Notebook An Introduction to Pigeon Science. Published by Author, Atwater, Ohio

22. Ghigi A. (1908) Sviluppo e comparsa di caratteri sessuli secondari in alcuni ucelli. Reale Accademia delle Scienze dell'Istituto di Bologna 15 Mar:3–23.

23. Moore J. (1735) Columbarium: or, the pigeon-house; being an introduction to a natural

history of tame pigeons. London

24. Ezzedine K, Eleftheriadou V, Whitton M, Van Geel N. (2015) Vitiligo. The Lancet 386:74–84. doi:10.1016/S0140-6736(14)60763-7.

25. Yaghoobi R, Omidian M, Bagherani N. (2011) Vitiligo: A review of the published work. Journal of Dermatology 38:419–431. doi:10.1111/j.1346-8138.2010.01139.x.

26. Njoo MD, Westerhof W. (2001) Vitiligo Pathogenesis and treatment. American Journal of Clinical Dermatology 2:167–81. doi:10.2165/00128071-200102030-00006.

27. Tobin DJ, Paus R. (2001) Graying: Gerontobiology of the hair follicle pigmentary unit. Experimental Gerontology 36:29–54. doi:10.1016/S0531-5565(00)00210-2.

28. Endou M, Aoki H, Kobayashi T, Kunisada T. (2014) Prevention of hair graying by factors that promote the growth and differentiation of melanocytes. The Journal of Dermatology 41:716–723. doi:10.1111/1346-8138.12570.

29. Peter J, Rodgers R. (2015) The Pigeon Genetics Newsletter. The Pigeon Genetics Newsletter 10:1–25.

30. Hollander WF. (1944) Mosaic Effects in Domestic Birds. The Quarterly Review of Biology 19:285–307.

31. Wright L, Lumley RW, Ludlow JW, Lydon AF. (1895) Fulton's Book of Pigeons.

32. Bellone RR. (2010) Pleiotropic effects of pigmentation genes in horses. Animal Genetics 41:100–110. doi:10.1111/j.1365-2052.2010.02116.x.

33. Andersson LS, Wilbe M, Viluma A, Cothran G, Ekesten B, Ewart S, Lindgren G. (2013) Equine Multiple Congenital Ocular Anomalies and Silver Coat Colour Result from the Pleiotropic Effects of Mutant PMEL. PLoS ONE 8:e75639. doi:10.1371/journal.pone.0075639.

34. Osinchuk S, Grahn B. (2018) Diagnostic ophthalmology. Canadian Veterinary Journal 59:315–316.

35. Ségard EM, Depecker MC, Lang J, Gemperli A, Cadoré J-L. (2013) Ultrasonographic features of PMEL17 (Silver) mutant gene-associated multiple congenital ocular anomalies (MCOA) in Comtois and Rocky Mountain horses. Veterinary Ophthalmology 16:429–435. doi:10.1111/vop.12021.

36. Ramsey DT, Ewart SL, Render JA, Cook CS, Latimer CA. (1999) Congenital ocular abnormalities of Rocky Mountain Horses. Veterinary Ophthalmology 2:47–59. doi:10.1046/j.1463-5224.1999.00050.x.

37. Clark LA, Wahl JM, Rees CA, Murphy KE. (2006) Retrotransposon insertion in SILV is responsible for merle patterning of the domestic dog. PNAS 103:1376–81. doi:10.1073/pnas.0506940103.

38. Domyan ET, Kronenberg Z, Infante CR, et al. (2016) Molecular shifts in limb identity underlie development of feathered feet in two domestic avian species. eLife 5:e12115. doi:10.7554/eLife.12115.

39. Hu H, Huff CD, Moore B, Flygare S, Reese MG, Yandell M. (2013) VAAST 2.0: Improved Variant Classification and Disease-Gene Identification Using a Conservation-Controlled Amino Acid Substitution Matrix. Genetic Epidemiology 37:622–634. doi:10.1002/gepi.21743.

40. Widlund HR, Fisher DE, Ramaswamy S, Miller AJ, Du J, Horstmann MA. (2011) MLANA/MART1 and SILV/PMEL17/GP100 Are Transcriptionally Regulated by MITF in Melanocytes and Melanoma. The American Journal of Pathology 163:333–343. doi:10.1016/s0002-9440(10)63657-7.

41. Hoashi T, Watabe H, Muller J, Yamaguchi Y, Vieira WD, Hearing VJ. (2005) MART-1 is required for the function of the melanosomal matrix protein PMEL17/GP100 and the maturation of melanosomes. Journal of Biological Chemistry 280:14006–14016. doi:10.1074/jbc.M413692200.

42. Mazie AM De, Donselaar E Van, Salvi S, Davoust J, Cerottini J, Slot JW. (2002) The Melanocytic Protein Melan-A / MART-1 Has a Subcellular Localization Distinct from Typical Melanosomal Proteins. Traffic 3:678–693.

43. Zeuthen T, Hamann S, La Cour M. (1996) Cotransport of H+, lactate and H2O by membrane proteins in retinal pigment epithelium of bullfrog. Journal of Physiology 497:3–17. doi:10.1113/jphysiol.1996.sp021745.

44. Halestrap AP, Meredith D. (2004) The SLC16 gene family - From monocarboxylate transporters (MCTs) to aromatic amino acid transporters and beyond. Pflugers Archiv European Journal of Physiology 447:619–628. doi:10.1007/s00424-003-1067-2.

45. Meredith D, Christian HC. (2008) The SLC16 monocaboxylate transporter family. Xenobiotica 38(7-8):1072–1106. doi:10.1080/00498250802010868.

46. Chidlow G, Wood JPM, Graham M, Osborne NN. (2004) Expression of monocarboxylate transporters in rat ocular tissues. American Journal of Physiology-Cell Physiology 288:C416–C428. doi:10.1152/ajpcell.00037.2004.

47. Jackson VN, Price NT, Carpenter L, Halestrap AP. (1997) Cloning of the monocarboxylate transporter isoform MCT2 from rat testis provides evidence that expression in tissues is species-specific and may involve post-transcriptional regulation. The Biochemical Journal 324:447–453. doi:10.1042/bj3240447.

48. Letunic I, Bork P. (2017) 20 years of the SMART protein domain annotation resource.

590 Nucleic Acids Research 46:D493–D496. doi:10.1093/nar/gkx922.

591 49. Price ER, Horstmann MA, Wells AG, Weilbaecher KN, Takemoto CM, Landis MW,
592 Fisher DE. (1998) α-melanocyte-stimulating hormone signaling regulates expression of
593 Microphthalmia, a gene deficient in Waardenburg syndrome. Journal of Biological
594 Chemistry 273:33042–33047. doi:10.1074/jbc.273.49.33042.

595 50. Aberdam E, Bertolotto C, Sviderskaya E V, de Thillot V, Hemesath TJ, Fisher DE,
596 Bennett DC, Ortonne JP, Ballotti R. (1998) Involvement of microphthalmia in the
597 inhibition of melanocyte lineage differentiation and of melanogenesis by agouti signal
598 protein. Journal of Biological Chemistry 273:19560–19565.
599 doi:10.1074/JBC.273.31.19560.

600 51. Dunn KJ, Brady M, Ochsenbauer-Jambor C, Snyder S, Incao A, Pavan WJ. (2005) WNT1
601 and WNT3a promote expansion of melanocytes through distinct modes of action. Pigment
602 Cell Research 18:167–180. doi:10.1111/j.1600-0749.2005.00226.x.

603 52. Takeda K, Yasumoto K, Takada R, Takada S, Watanabe K, Udono T, Saito H, Takahashi
604 K, Shibahara S. (2000) Induction of melanocyte-specific microphthalmia-associated
605 transcription factor by Wnt-3a. Journal of Biological Chemistry 275:14013–14016.
606 doi:10.1074/ jbc.C000113200.

607 53. Levy C, Khaled M, Fisher DE. (2006) MITF: master regulator of melanocyte development
608 and melanoma oncogene. Trends in Molecular Medicine 12:406–414.
609 doi:10.1016/j.molmed.2006.07.008.

610 54. Haase E, Ito S, Sell A, Wakamatsu K. (1992) Feathers from Wild and Domestic Pigeons.
611 The Journal of Heredity 83:4–7. doi:10.1093/oxfordjournals.jhered.a111160.

612 55. Harris ML, Baxter LL, Loftus SK, Pavan WJ. (2010) Sox proteins in melanocyte

development and melanoma. Pigment Cell and Melanoma Research 23:496–513. doi:10.1111/j.1755-148X.2010.00711.x.

56. Steingrímsson E, Copeland NG, Jenkins NA. (2004) Melanocytes and the Microphthalmia Transcription Factor Network. Annual Review of Genetics 38:365–411. doi:10.1146/annurev.genet.38.072902.092717.

57. Yasumoto K, Yokoyama K, Takahashi K, Tomita Y, Shibahara S. (1997) Functional analysis of microphthalmia-associated transcription factor in pigment cell-specific transcription of the human tyrosinase family genes. Journal of Biological Chemistry 272:503–9. doi:10.1074/jbc.272.1.503.

58. Mountjoy KG, Robbins LS, Mortrud MT, Cone RD. (1992) The cloning of a family of genes that encode the melanocortin receptors. Science 257:1248–1251. doi:10.1126/science.1325670.

59. Zdarsky E, Favor J, Jackson IJ. (1990) The molecular basis of brown, an old mouse mutation, and of an induced revertant to wild type. Genetics 126:443–449.

60. Watt B, Van Niel G, Raposo G, Marks MS. (2013) PMEL: A pigment cell-specific model for functional amyloid formation. Pigment Cell and Melanoma Research 26:300–315. doi:10.1111/pcmr.12067.

61. Hellström AR, Watt B, Fard SS, et al. (2011) Inactivation of PMEL alters melanosome shape but has only a subtle effect on visible pigmentation. PLoS Genetics. doi: 10.1371/journal.pgen.1002285

62. Kobayashi T, Urabe K, Orlow SJ, Higashi K, Imokawa G, Kwon BS, Potterf B, Hearing VJ. (1994) The Pmel 17/silver locus protein. Characterization and investigation of its melanogenic function. Journal of Biological Chemistry 269:29198–29205.

636 63. Huntley R. (1999) The Almond Family.
637 http://www.angelfire.com/ga/huntleyloft/qualmond.html. Accessed 1 Jan 2017
638 64. Hollander WF. (1975) Sectorial Mosaics in the Domestic Pigeon : 25 More Years. Journal
639 of Heredity 66:177–202.
640 65. Colnaghi R, Carpenter G, Volker M, O'Driscoll M. (2011) The consequences of structural
641 genomic alterations in humans: Genomic Disorders, genomic instability and cancer.
642 Seminars in Cell and Developmental Biology 22:875–885.
643 doi:10.1016/j.semcdb.2011.07.010.
644 66. Du J, Miller AJ, Widlund HR, Horstmann MA, Ramaswamy S, Fisher DE. (2003)
645 MLANA/MART1 and SILV/PMEL17/GP100 Are Transcriptionally Regulated by MITF
646 in Melanocytes and Melanoma. American Journal of Pathology 163:333–343.
647 doi:10.1016/S0002-9440(10)63657-7.
648 67. Kerje S, Sharma P, Gunnarsson U, et al. (2004) The Dominant white, Dun and Smoky
649 color variants in chicken are associated with insertion/deletion polymorphisms in the
650 PMEL17 gene. Genetics 168:1507–18. doi:10.1534/genetics.104.027995.
651 68. Kuehn C, Weikard R. (2007) Multiple splice variants within the bovine silver homologue
652 (SILV) gene affecting coat color in cattle indicate a function additional to fibril formation
653 in melanophores. BMC Genomics 8:335. doi:10.1186/1471-2164-8-335.
654 69. Schmutz SM, Dreger DL. (2013) Interaction of *MC1R* and *PMEL* alleles on solid coat
655 colors in Highland cattle. Animal Genetics 44:9–13. doi:10.1111/j.1365-
656 2052.2012.02361.x.
657 70. Kwon BS, Halaban R, Ponnazhagan S, Kim K, Chintamaneni C, Bennett D, Pickard RT.
658 (1995) Mouse silver mutation is caused by a single base insertion in the putative

659 cytoplasmic domain of Pmel 17. Nucleic Acids Research 23:154–8.

660 71. Schonthaler HB, Lampert JM, Von Lintig J, Schwarz H, Geisler R, Neuhauss SCF. (2005)
661 A mutation in the silver gene leads to defects in melanosome biogenesis and alterations in
662 the visual system in the zebrafish mutant fading vision. Developmental Biology 284:421–
663 436. doi:10.1016/j.ydbio.2005.06.001.

664 72. Murphy SC, Evans JM, Tsai KL, Clark LA. (2018) Length variations within the Merle
665 retrotransposon of canine PMEL: correlating genotype with phenotype. Mobile DNA
666 9:26. doi:10.1186/s13100-018-0131-6.

667 73. Brunberg E, Andersson L, Cothran G, Sandberg K, Mikko S, Lindgren G. (2006) A
668 missense mutation in PMEL17 is associated with the Silver coat color in the horse. BMC
669 Genetics. doi: 10.1186/1471-2156-7-46

670 74. Komáromy AM, Rowlan JS, La Croix NC, Mangan BG. (2011) Equine Multiple
671 Congenital Ocular Anomalies (MCOA) syndrome in PMEL17 (Silver) mutant ponies:
672 Five cases. Veterinary Ophthalmology 14:313–320. doi:10.1111/j.1463-
673 5224.2011.00878.x.

674 75. Hellström AR, Watt B, Fard SS, et al. (2011) Inactivation of Pmel Alters Melanosome
675 Shape But Has Only a Subtle Effect on Visible Pigmentation. PLoS Genetics 7:e1002285.
676 doi:10.1371/journal.pgen.1002285.

677 76. Gelatt KN, Powell NG, Huston K. (1981) Inheritance of microphthalmia with coloboma in
678 the Australian shepherd dog. American Journal of Veterinary Research 42:1686–1690.

679 77. Vogel P, Read RW, Vance RB, Platt KA, Troughton K, Rice DS. (2008) Ocular albinism
680 and hypopigmentation defects in Slc24a5-/- mice. Veterinary Pathology 45:264–279.
681 doi:10.1354/vp.45-2-264.

682 78. Brooks BP, Larson DM, Chan C-C, et al. (2007) Analysis of Ocular Hypopigmentation in
683 Rab38 cht/cht Mice. Investigative Opthalmology & Visual Science 48:3905.
684 doi:10.1167/iovs.06-1464.
685 79. Gronskov K, Ek J, Brondum-Nielsen K. (2007) Oculocutaneous albinism. Orphanet
686 Journal of Rare Diseases 2, 43. doi:10.1186/1750-1172-2-43.
687 80. Mannermaa E, Vellonen KS, Urtti A. (2006) Drug transport in corneal epithelium and
688 blood-retina barrier: Emerging role of transporters in ocular pharmacokinetics. Advanced
689 Drug Delivery Reviews 58:1136–1163. doi:10.1016/j.addr.2006.07.024.
690 81. Nguyen TT, Bonanno JA. (2012) Lactate-H + transport is a significant component of the
691 in vivo corneal endothelial pump. Investigative Ophthalmology & Visual Science
692 53:2020–2029. doi:10.1167/iovs.12-9475.
693 82. Stranger BE, Forrest MS, Dunning M, et al. (2007) Relative Impact of Nucleotide and
694 Copy Number Variation on Gene Expression Phenotypes. Science 315:848–853.
695 doi:10.1126/science.1137223.
696 83. Zhou J, Lemos B, Dopman EB, Hartl DL. (2011) Copy-number variation: The balance
697 between gene dosage and expression in Drosophila melanogaster. Genome Biology and
698 Evolution 3:1014–1024. doi:10.1093/gbe/evr023.
699 84. Freeman JL, Perry GH, Feuk L, et al. (2006) Copy number variation: New insights in
700 genome diversity. Genome Research 949–961. doi:10.1101/gr.3677206.16.
701 85. Cruz C, Houseley J. (2014) Endogenous RNA interference is driven by copy number.
702 eLife 3:e01581. doi:10.7554/eLife.01581.
703 86. Dorer DR, Henikoff S. (1994) Expansions of transgene repeats cause heterochromatin
704 formation and gene silencing in Drosophila. Cell 77:993–1002. doi:10.1016/0092-

8674(94)90439-1.

87. Assaad FF, Tucker KL, Signer ER. (1993) Epigenetic repeat-induced gene silencing (RIGS) in Arabidopsis. Plant Molecular Biology 22:1067–1085. doi:10.1007/BF00028978.

88. Garrick D, Fiering S, Martin DIK, Whitelaw E. (1998) Repeat-induced gene silencing in mammals. Nature Genetics 18:56–59. doi:10.1038/ng0198-56.

89. Henikoff S. (1998) Conspiracy of silence among repeated transgenes. BioEssays 20:532–535. doi:10.1002/(SICI)1521-1878(199807)20:7<532::AID-BIES3>3.0.CO;2-M.

90. Girirajan S, Campbell CD, Eichler EE. (2011) Human Copy Number Variation and Complex Genetic Disease. Annual Review of Genetics 45:203–226. doi:10.1146/annurev-genet-102209-163544.

91. Wain L V., Armour J AL, Tobin MD. (2009) Genomic copy number variation, human health, and disease. The Lancet 374:340–350. doi:10.1016/S0140-6736(09)60249-X.

92. Zhang F, Gu W, Hurles ME, Lupski JR. (2009) Copy Number Variation in Human Health, Disease, and Evolution. Annual Review of Genomics and Human Genetics 10:451–81. doi:10.1146/annurev.genom.9.081307.164217.

93. Sopko R, Huang D, Preston N, et al. (2006) Mapping Pathways and Phenotypes by Systematic Gene Overexpression. Molecular Cell 21:319–330. doi:10.1016/j.molcel.2005.12.011.

94. Lang KS, Muhm A, Moris A, Stevanovic S, Rammensee H-G, Caroli CC, Wernet D, Schittek B, Knauss-Scherwitz E, Garbe C. (2001) HLA-A2 Restricted, Melanocyte-Specific CD8+ T Lymphocytes Detected in Vitiligo Patients are Related to Disease Activity and are Predominantly Directed Against MelanA/MART1. Journal of

728 Investigative Dermatology 116:891–897. doi:10.1046/j.1523-1747.2001.01363.x.

729 95. Byrne KT, Turk MJ. (2011) New perspectives on the role of vitiligo in immune responses
730 to melanoma. Oncotarget 2:684–94. doi:10.18632/oncotarget.323.

731 96. Shi F, Kong B-W, Song J, Lee J, Dienglewicz RL, Erf GF. (2012) Understanding
732 mechanisms of vitiligo development in Smyth line of chickens by transcriptomic
733 microarray analysis of evolving autoimmune lesions. BMC Immunology 13:18.
734 doi:10.1186/1471-2172-13-18.

735 97. Wright D, Boije H, Meadows JRS, et al. (2009) Copy Number Variation in Intron 1 of
736 SOX5 Causes the Pea-comb Phenotype in Chickens. PLoS Genetics 5:e1000512.
737 doi:10.1371/journal.pgen.1000512.

738 98. Hollander WF. (1982) Origins and excursions in pigeon genetics- flecks and sex. The Ink
739 Spot INC, Burton, Kansas

740 99. Sponenberg DP. (1984) Germinal reversion of the merle allele In Australian shepherd
741 dogs. Journal of Heredity 75:78–78. doi:10.1093/oxfordjournals.jhered.a109874.

742 100. Farslow JC, Lipinski KJ, Packard LB, Edgley ML, Taylor J, Flibotte S, Moerman DG,
743 Katju V, Bergthorsson U. (2015) Rapid Increase in frequency of gene copy-number
744 variants during experimental evolution in Caenorhabditis elegans. BMC Genomics
745 16:1044. doi:10.1186/s12864-015-2253-2.

746 101. Bryk J, Tautz D. (2014) Copy number variants and selective sweeps in natural populations
747 of the house mouse (Mus musculus domesticus). Frontiers in Genetics 5:153.
748 doi:10.3389/fgene.2014.00153.

749 102. Axelsson E, Ratnakumar A, Arendt ML, Maqbool K, Webster MT, Perloski M, Liberg O,
750 Arnemo JM, Hedhammar Å, Lindblad-Toh K. (2013) The genomic signature of dog

751 domestication reveals adaptation to a starch-rich diet. Nature 495:360–364.
752 doi:10.1038/nature11837.
753 103. Cone KR, Kronenberg ZN, Yandell M, Elde NC. (2017) Emergence of a Viral RNA
754 Polymerase Variant during Gene Copy Number Amplification Promotes Rapid Evolution
755 of Vaccinia Virus. Journal of Virology 91:e01428-16. doi:10.1128/jvi.01428-16.
756 104. Holt C, Campbell M, Keays DA, et al. (2018) Improved Genome Assembly and
757 Annotation for the Rock Pigeon (Columba livia). G3 8:1391–1398.
758 doi:10.1534/g3.117.300443.
759 105. McKenna A, Hanna M, Banks E, et al. (2010) The Genome Analysis Toolkit: A
760 MapReduce framework for analyzing next-generation DNA sequencing data. Genome
761 Research 20:1297–1303. doi:10.1101/gr.107524.110.
762 106. Yandell M, Huff C, Hu H, Singleton M, Moore B, Xing J, Jorde LB, Reese MG. (2011) A
763 probabilistic disease-gene finder for personal genomes. Genome Research 21:1529–1542.
764 doi:10.1101/gr.123158.111.
765 107. Shapiro MD, Domyan ET. (2013) Quick guide: Domestic Pigeons. Current Biology
766 23:R302–R303. doi:10.1016/j.cub.2013.01.063.
767 108. Danecek P, Auton A, Abecasis G, et al. (2011) The variant call format and VCFtools.
768 Bioinformatics 27:2156–2158. doi:10.1093/bioinformatics/btr330.
769 109. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL.
770 (2009) BLAST+: architecture and applications. BMC Bioinformatics 10:421.
771 doi:10.1186/1471-2105-10-421.
772 110. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov
773 JP. (2011) Integrative genomics viewer. Nature Biotechnology 29:24–26.

774 doi:10.1038/nbt.1754.

775 111. R Development Core Team. (2018) R: A language and environment for statistical

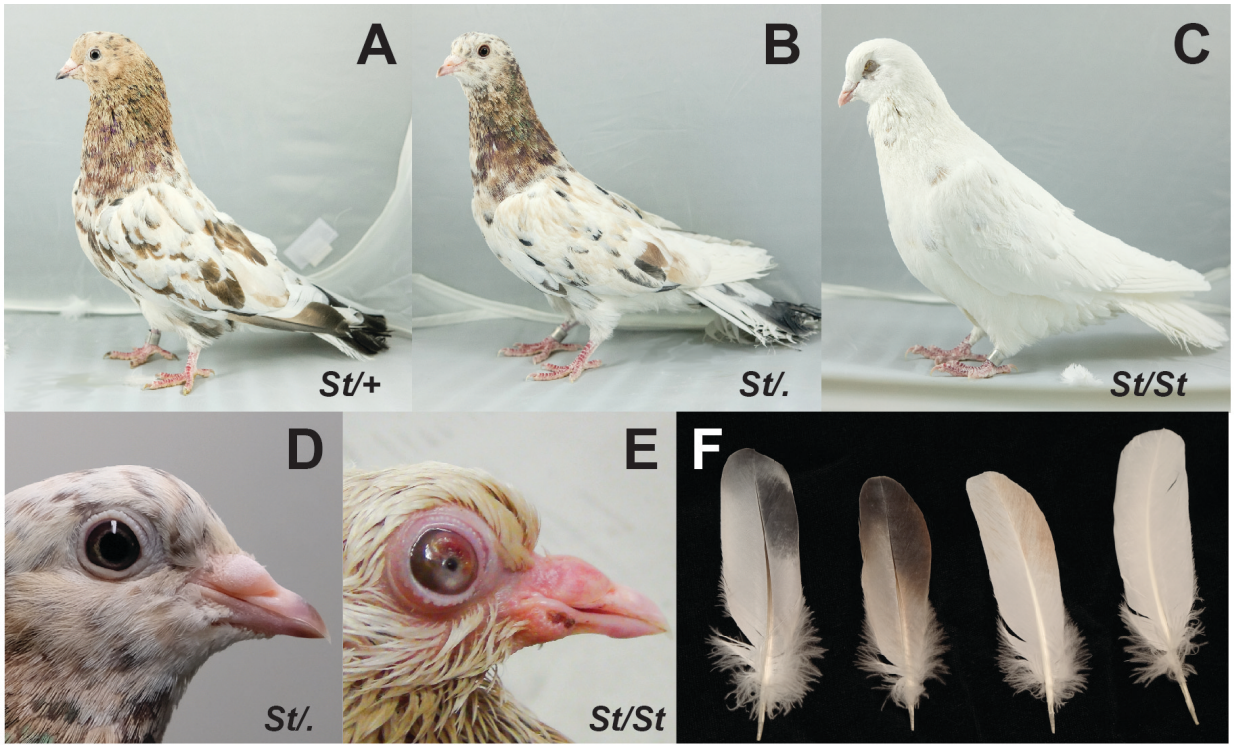
776 computing. In: R Foundation for Statistical Computing Statistical Computing.

777 https://www.r-project.org.

778

779

780 **FIGURES AND FIGURE LEGENDS**
781



782 **Figure 1.** Phenotypes of pigeons carrying Almond alleles (*St*, Almond allele; +, wild type allele).
783 (A) Heterozygous Almond male. (B) Hemizygous Almond female. (C) Homozygous Almond
784 male. (D) Almond females have no observable eye defects. (E) Homozygous Almond males often
785 show severe eye defects. Defects pictured in this juvenile include bloated eyelid and anterior
786 opacity. (F) Wing feathers from different phenotypes, left to right: non-Almond, dark Almond,
787 light Almond, homozygous Almond.

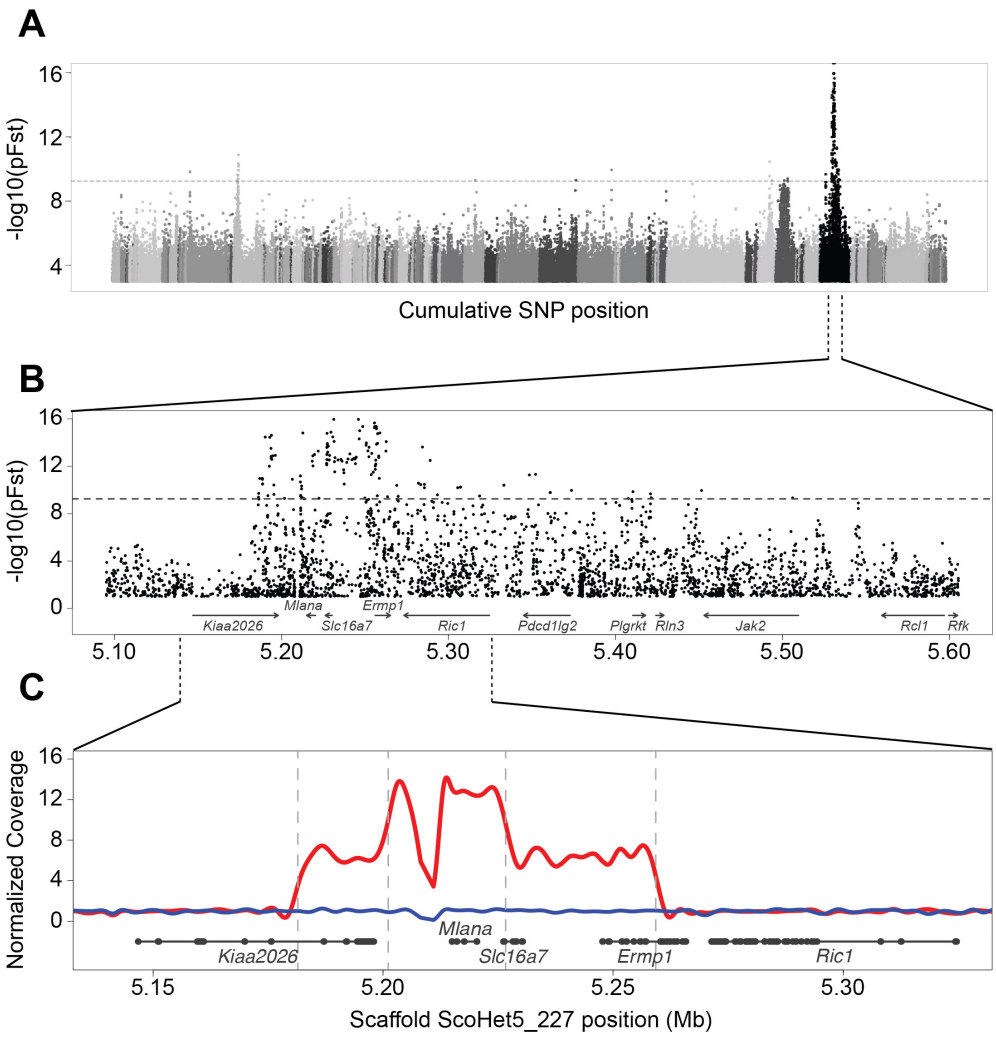


Figure 2. Almond is associated with a CNV on a sex-linked genomic scaffold. (A) Whole-genome pFst comparisons between Almond and non-Almond pigeons. Each dot represents a SNP position, with shades of gray indicating different genomic scaffolds. The horizontal dashed grey line indicates genome-wide significance threshold. (B) Detail of pFst plot for candidate region on ScoHet5_227, a sex-linked scaffold [104]. Gene models are depicted at the bottom of the plot. (C) Detail view of the CNV region. Solid red line represents the mean normalized read depth for 10 female Almond birds in this region. The blue line is a single representative of non-Almond female coverage. Vertical dashed lines indicate positions of CNV breakpoints. Gene models are depicted below the coverage plot in grey (thick lines, exons; thin lines, introns).

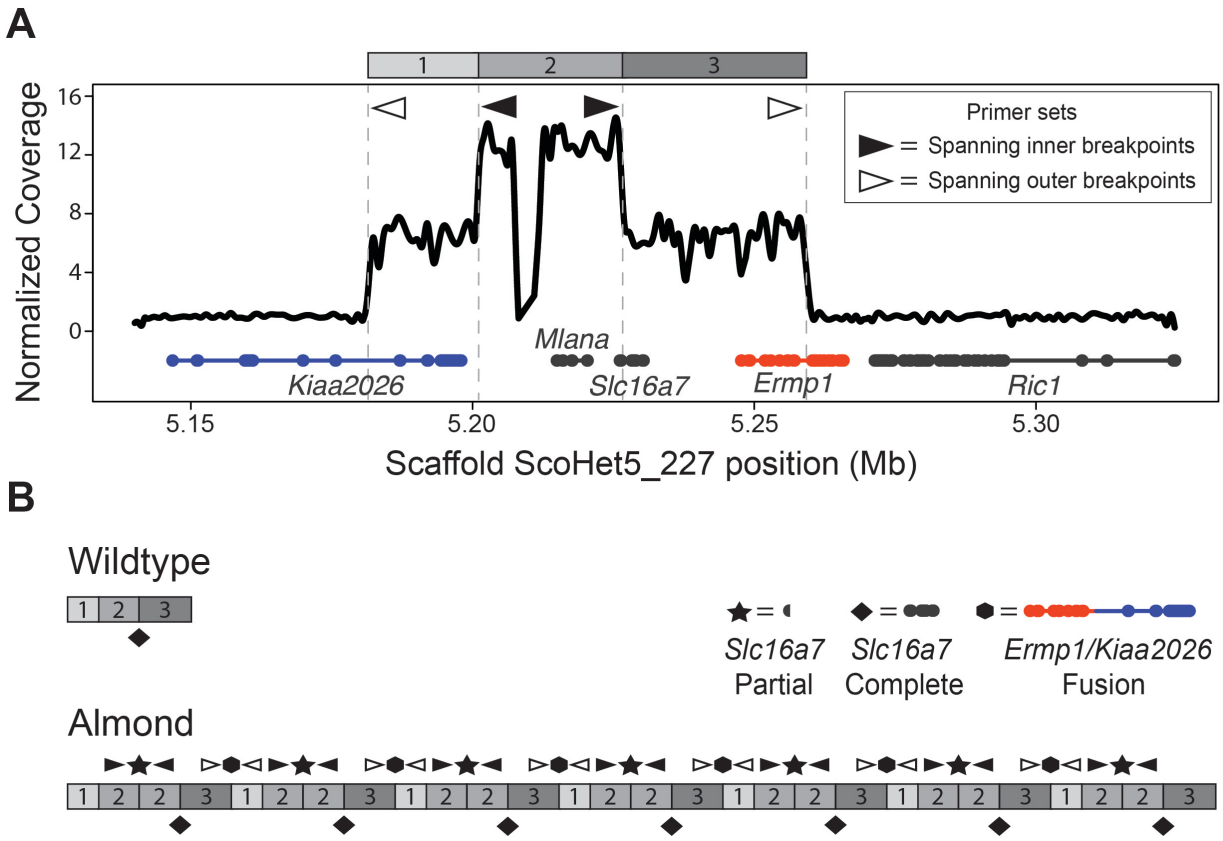


Figure 3. The Almond-associated CNV has a complex structure that results in duplicated, truncated, and fused genes. (A) Coverage diagram showing different regions of the CNV normalized to a non-CNV region on the same scaffold. Two outer regions (1 and 3, above plot) have an approximately 7-fold coverage increase, while one inner region (2) has an approximately 14-fold coverage increase. Gene models are depicted below the coverage plot in grey, orange and blue (thick lines, exons; thin lines, introns). (B) Schematic of the non-Almond (top) and inferred Almond (bottom) structures of the CNV. Gene structural changes resulting from the Almond CNV include a fusion of *Ermp1* and *Kiaa2026* at the segment 3/1 junction (hexagon), and a truncated version of *Slc16a7* at the segment 2/2 junction (star). A complete copy of *Slc16a7* occurs at each 2/3 junction (diamond).

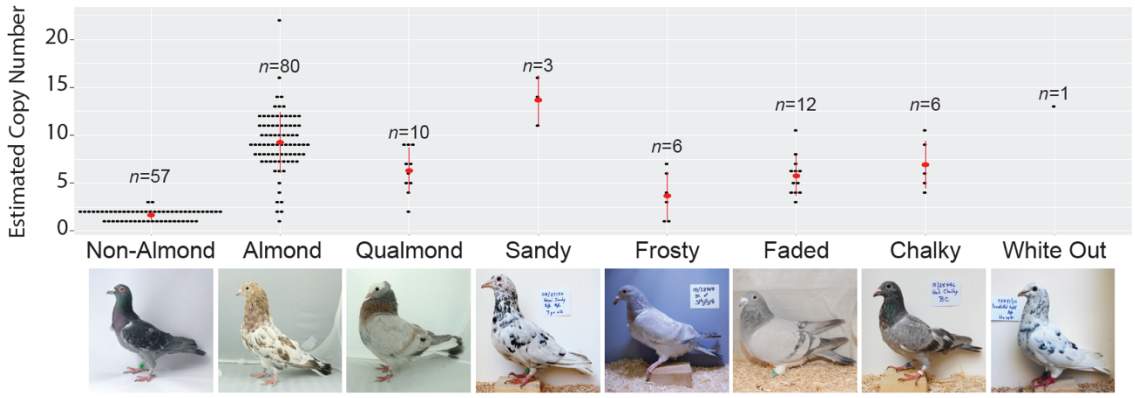
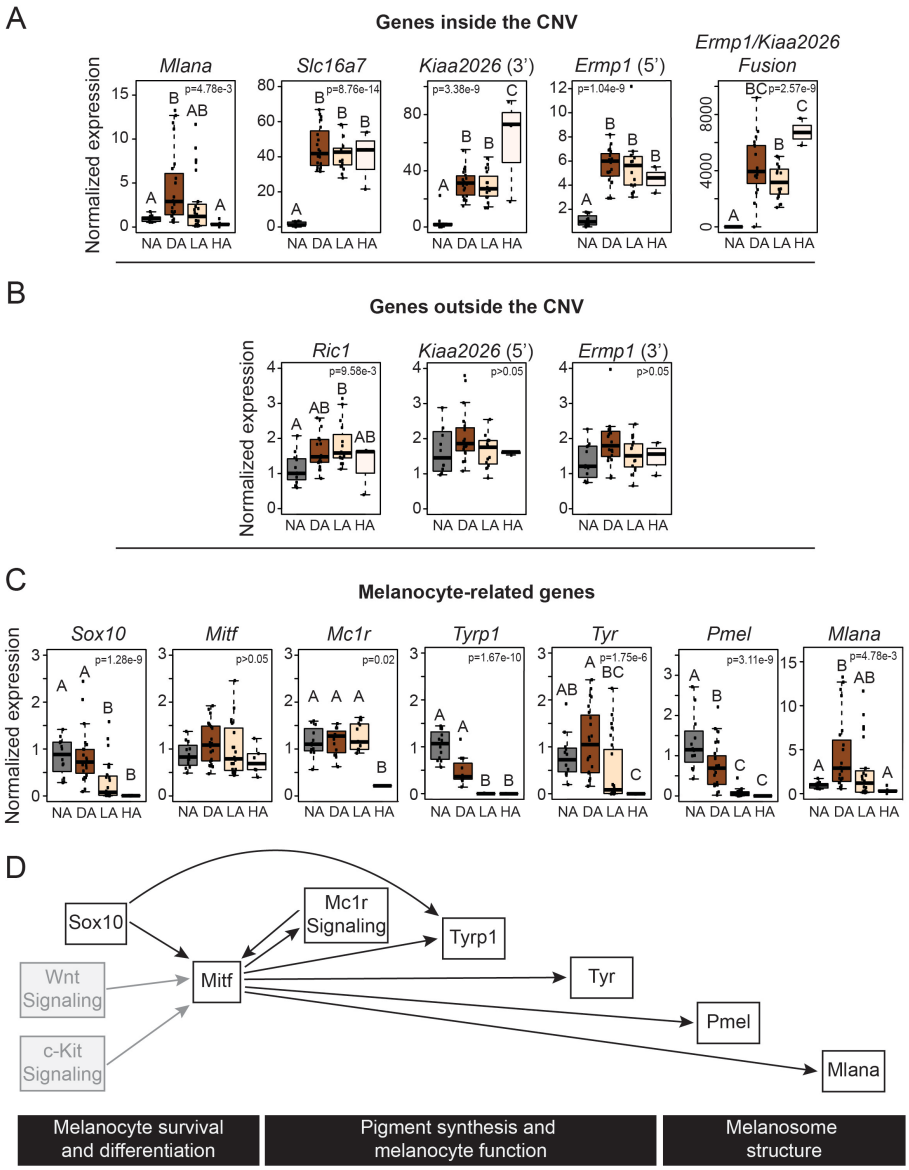


Figure 4. *St*-linked pigmentation phenotypes show quantitative variation in the Almond CNV. Black dots represent results of a TaqMan copy number assay. Mean copy numbers for each phenotype are shown as red dots. Most individuals without *St*-linked phenotypes have the expected 1 or 2 copies (because *St* is a sex-linked locus, females have a minimum of 1 copy and males have a minimum of 2). All other *St*-linked phenotypes are associated with an expansion of the CNV in the Almond candidate region on scaffold ScoHet5_227, indicating an allelic series at *St*. Numbers above each phenotype indicate number of individuals sampled.



820

821 **Figure 5.** Almond and non-Almond feather buds have distinct gene expression profiles. (A) Exons
822 assayed within the CNV show expression differences in Almond feather buds compared to non-
823 Almond. Boxplots show the results of qRT-PCR assays designed to assess gene expression of
824 exons located in the CNV region. Fusion gene expression results are from qPCR primers spanning
825 exon 7 of *Ermp1* into exon 5 of *Kiaa2026*. (B) Exons assayed outside the CNV show no expression
826 differences in Almond feather buds compared to non-Almond. This indicates expression

differences are specific to exons inside the CNV. (C) Expression of melanocyte-related genes. qRT-PCR results indicate a decrease in expression of several genes involved in melanin production in Almond feather buds. (D) Model of interactions among genes and signaling pathways involved in different aspects of pigment synthesis. Gray boxes indicate pathways discussed in the text but not directly represented in our expression analyses. NA, feather buds from non-Almond individuals with wild type alleles at *St*; DA, dark Almond feather buds from hemizygous and heterozygous Almond individuals; LA, light Almond feather buds from hemizygous and heterozygous Almond individuals; HA, feather buds from a homozygous Almond individual. Bar in each box represents the median**,** box ends indicate upper and lower quartiles, whiskers indicate the highest and lowest value excluding outliers. Different letters indicate groups with statistically significant differences in gene expression determined by ANOVA and post-hoc Tukey test ($p<0.05$).

SUPPORTING INFORMATION LEGEND

Supporting_Information.xlsx. S1-S5 Tables. Individual tables are separate worksheets within the file and contain the following information: S1, NCBI SRA submission numbers and breed information for birds used for genomic analysis in this study. S2, Sample sizes and Identifiers of birds included in each phenotypic category for qPCR analysis. S3, Raw qRT-PCR results for Figure 4. S4, Copy number results from Taqman assay of *Mlana* region. S5, Primer Sequences Used in This Study. S2: