



Insights into the genetic history of Green-legged Partridge-like fowl: mtDNA and genome-wide SNP analysis

M. Siwek*, D. Wragg[†], A. Sławińska*[‡], M. Malek[‡], O. Hanotte[†] and J. M. Mwacharo[†]

*Department of Animal Biotechnology, University of Technology and Life Sciences, Mazowiecka 28, 85-084, Bydgoszcz, Poland. [†]Centre for Genetics and Genomics, School of Biology, University Park, The University of Nottingham, Nottingham, NG7 2RD, UK. [‡]FAO/IAEA Agriculture and Biotechnology Laboratory, International Atomic Energy Agency, Wagramerstrasse 5, 1400, Vienna, Austria.

Summary

The Green-legged Partridge-like (GP) fowl, an old native Polish breed, is characterised by reseda green-coloured shanks rather than yellow, white, slate or black commonly observed across most domestic breeds of chicken. Here, we investigate the origin, genetic relationships and structure of the GP fowl using mtDNA D-loop sequencing and genome-wide SNP analysis. Genome-wide association analysis between breeds enables us to verify the genetic control of the reseda green shank phenotype, a defining trait for the breed. Two mtDNA D-loop haplogroups and three autosomal genetic backgrounds are revealed. Significant associations of SNPs on chromosomes GGA24 and GGZ indicate that the reseda green leg phenotype is associated with recessive alleles linked to the *W* and *Id* loci. Our results provide new insights into the genetic history of European chicken, indicating an admixed origin of East European traditional breeds of chicken on the continent, as supported by the presence of the reseda green phenotype and the knowledge that the GP fowl as a breed was developed before the advent of commercial stocks.

Keywords autosomal markers, European chicken, *Gallus gallus*, mtDNA control region, Z chromosome markers.

Introduction

The Red junglefowl *Gallus gallus* is widely recognised as the main matriarchal ancestor of all domestic chicken (West & Zhou 1988), although it has recently been shown that other *Gallus* species also have contributed to the genetic make-up of some modern day domestic stocks (Nishibori *et al.* 2005; Eriksson *et al.* 2008). Initial evidence indicated that all domestic chicken originated from mainland South-East Asia (present day Thailand) from a single wild subspecies *G. g. gallus* (Fumihito *et al.* 1994, 1996). However, the recent observation of nine distinct mtDNA D-loop haplogroups in Eurasian domestic chicken has provided some

support for multiple domestications across South-West and South-East Asia (Liu *et al.* 2006; Miao *et al.* 2013).

The genetic diversity of European native chicken at the mtDNA level remains poorly studied. A recent study of Hungarian chicken (Revay *et al.* 2010) revealed two mtDNA D-loop lineages, with the authors suggesting that these could have been derived from the Indian subcontinent and South-East Asia/China respectively. In a similar study, Dana *et al.* (2010) showed that East Asian chickens contributed to the development of Dutch traditional and Western commercial breeds of chicken.

The majority of European breeds have yellow-coloured shanks, whilst in others, they are typically white, bluish grey, slate or black. The Green-legged Partridge-like (GP; *Zielononóżka kuropatwiana*) fowl is an old native Polish breed of chicken with reseda green feet/shanks (Fig. S1) and grey partridge-like *Perdix perdix* plumage. The origin and development of the breed remain unclear. Most information about the breed originates from anecdotal reports (see: <http://www.fao.org/ag/againfo/programmes/documents/genetics/story/story40.html>). Polish chickens that phenotypically resembled GP fowls were described in 1866 by Pietruski and later on by Graff (1887). Other reports indicate that GP-like chicken were first introduced to Poland from Italy much earlier, possibly in the 16th

Address for correspondence

M. Siwek, Department of Animal Biotechnology, University of Technology and Life Sciences, Mazowiecka 28, 85-084 Bydgoszcz, Poland.

E-mail: siwek@utp.edu.pl

and

J. M. Mwacharo, Centre for Genetics and Genomics, School of Biology, University Park, The University of Nottingham, Nottingham, NG7 2RD UK.

E-mail: joram.mwacharo@nottingham.ac.uk

Accepted for publication 26 February 2013

century (Witkowski *et al.* 2009). Whatever the case, the name Green-legged Partridge-like was used for the first time in 1879 to denote the breed in the journal *Poultry Breeder* by Obfidowicz (Witkowski *et al.* 2009). In 1894, the breed was presented for the first time at a country exhibition in Lvov/Lviv, Ukraine. Other information from grey literature indicates that GP chickens were first bred in Poland in 1904. The breed was defined as a formal strain of the indigenous fowl of Poland in 1923, and from amongst several colour varieties, the Partridge-like feathers, reseda green legs/shanks and beaks were then chosen as the defining breed standards. In 1932, a herd book and nucleus breeding flock were established. The breed was registered in 1988 with the International Registry of Poultry Genetic Stocks (Somes 1988). Although for many years, the breed was widespread across Poland, by 1973, its population size had declined to almost 2% of its original size, and by 1988, only 624 birds were known to exist (Witkowski *et al.* 2009). Currently, the breed is maintained in two separate conservation flocks, each with about 600 hens and 70 cocks that are managed by the University of Life Sciences in Lublin and the Polish Institute of Animal Genetics and Breeding respectively. Fancy breeders also maintain a small number of unregistered flocks. The breed requires minimal veterinary health care, tolerates very low temperatures, is an excellent forager and is reputed to lay eggs low in cholesterol (Krawczyk *et al.* 2005; Krawczyk & Szefer 2012). In recent times, it has been used to study the genetic background of general immune responsiveness to KLH, LTA and LPS antigens (Siwek *et al.* 2010; Sławińska *et al.* 2011; Siwek *et al.* 2012).

The two most common skin pigments in birds are melanins and carotenoids. Melanins are divided into eumelanin, which is the main pigment in black or blue feathers, eyes, skin and connective tissue, and pheomelanin, which is deposited mainly in reddish brown and buff-coloured feathers. Carotenoids and xanthophylls are responsible for the yellow colouration in skin, fat and egg yolk (Smyth 1990). A wide spectrum of leg/shank colour polymorphisms (black, white, yellow, green, etc.) can be observed in domestic and wild chicken. In wild *Gallus* species, shank colours include shades of plumbeous brown to bluish grey (Red junglefowl, *G. gallus*), shades of yellow to salmon red (Grey junglefowl, *G. sonneratii*), brownish yellow (Ceylon junglefowl, *G. lafayettii*) and white to pink (Green junglefowl, *G. varius*) (Delacour 1977). In domestic chicken, yellow, white, grey, slate and/or black shanks are commonly observed. Some breeds such as Ayam Cemani, Black H'Mong, Korean Ogol, Kadaknath and Silkie have black skin and shanks. There are also various shades of either blue or green or both depending on the interaction of genes affecting the pigmentation of epidermis and dermis (Smyth 1990).

The major loci controlling skin colour phenotypes and their respective genotypes have previously been described (Bateson 1902; Dunn 1925; Dunn & Jull 1927; Hutt 1949).

The genetic control of skin phenotypes relies on the cumulative and interactive effects of several major pigmentation loci. These include alleles for dominant white (wild type) and recessive yellow skin at the autosomal *W* locus, sex-linked dermal melanin (wild type) recessive to its inhibitor at the *Id* locus, epidermal melanin at the *E* locus, fibromelanosis at the *Fm* locus and several unidentified modifiers that enhance or restrict eumelanin (Smyth 1990). The green shank colour has previously been shown to result from the interaction of the *w/w* and *id⁺/id⁺* alleles in the Cornell random-bred White Leghorn population (McGibbon 1974, 1979).

Eriksson *et al.* (2008) identified the *BCDO2* (also known as *BCO2*) as the gene underlying the yellow skin colour. The gene spans a region between 6.26 and 6.29 Mb in size on chromosome 24 (GGA24). Beta-carotene dioxygenase-2, the enzyme encoded by the *BCDO2* gene, acts by diffusing colourful carotenoids into colourless apocarotenoids, and this changes the skin colour from yellow to white.

The biology and genetic control of dermal hyperpigmentation (black skin/shanks) in the Silkie also have been studied (Dorshorst *et al.* 2010). Significant associations were mapped for two loci: dermal melanin (*Id*) to the Z chromosome (GGAZ) at position 72.3 Mb and fibromelanosis (*Fm*) to chromosome 20 (GGA20) at position 10.3–13.1 Mb. A further study by Dorshorst *et al.* (2011) identified the causal mutation of *Fm* as an inverted duplication and junction of two genomic intervals – one of which contains the *EDN3* gene, which plays a role in melanoblast proliferation.

In this study, we investigate the genetic relationship and structure of the GP fowl using mitochondrial DNA D-loop and full genome-wide SNP scan analysis. We reveal that the breed has an admixed genetic background from two maternal and three autosomal genetic contributions. We confirm that one of its defining traits, the reseda green shank phenotype, is associated with recessive alleles at both the *W* and *Id* loci. In the light of the known history of the breed, our results support the fact that admixed populations such as the GP fowl were likely present in Europe prior to the development of modern commercial breeds.

Materials and methods

Sample collection and DNA extraction

Eighty-four DNA samples [The Roslin Institute, $n = 43$; AVIANDIV (<http://aviandiv.tzv.fal.de/>), $n = 8$; private breeders, $n = 33$] representing 35 traditional European breeds of chicken were used for genotyping (Table 1). Junglefowl samples were obtained from various sources. *G. g. murghi* DNA samples were from Lehr Brisbin's flock in Alabama, USA. Blood spotted on FTA[®] classic cards (Whatman) from two individuals each of *G. g. gallus*, *G. lafayettii*, *G. varius* and *G. sonneratii* came from a private breeder (Fazanterie de Rooie Hoeve, The Netherlands). For

Table 1 Details of the breeds that were used for genotyping using the 60K SNP chip.

Breed	Shank colour ¹	Sample source and size			Genotypes ²			QC fail (IBS)
		RI	PB	AD	W ⁺	<i>Id</i>	<i>Fm</i>	
Domestic chicken								
Appenzellor	Blue	2			W ⁺	<i>id⁺/id⁺</i>	<i>fm⁺/fm⁺</i>	1
Araucana	Willow to olive or slate	6	1		w/w	<i>id⁺/id⁺</i>	<i>fm⁺/fm⁺</i>	0
Brahma	Orange yellow or yellow	1			w/w	<i>Id</i>	<i>fm⁺/fm⁺</i>	0
Buff Orpington	White	1			W ⁺	<i>Id</i>	<i>fm⁺/fm⁺</i>	0
Cochin	Yellow	1			w/w	<i>Id</i>	<i>fm⁺/fm⁺</i>	0
Cream Legbar	Yellow		2		w/w	<i>Id</i>	<i>fm⁺/fm⁺</i>	0
Crevecoeur	Black or slate blue		3		W ⁺	<i>id⁺/id⁺</i>	<i>fm⁺/fm⁺</i>	0
Croad Langshan	Bluish black	1			W ⁺	<i>id⁺/id⁺</i>	<i>fm⁺/fm⁺</i>	0
Derbyshire Redcap	Slate blue	1			W ⁺	<i>id⁺/id⁺</i>	<i>fm⁺/fm⁺</i>	0
Dorking	White	1	2		W ⁺	<i>Id</i>	<i>fm⁺/fm⁺</i>	0
Green-legged Partridge-like	Reseda green		5		w/w	<i>id⁺/id⁺</i>	<i>fm⁺/fm⁺</i>	0
Hamburgh	Lead blue	2			W ⁺	<i>id⁺/id⁺</i>	<i>fm⁺/fm⁺</i>	1
Indian Game	Rich orange or yellow	1			w/w	<i>Id</i>	<i>fm⁺/fm⁺</i>	0
Ixworth	White	1			W ⁺	<i>Id</i>	<i>fm⁺/fm⁺</i>	0
Leghorn	Yellow or orange	1			w/w	<i>Id</i>	<i>fm⁺/fm⁺</i>	0
Lincolnshire Buff	White	1			W ⁺	<i>Id</i>	<i>fm⁺/fm⁺</i>	0
Malay	Rich yellow	1			w/w	<i>Id</i>	<i>fm⁺/fm⁺</i>	0
Marans	White	4	4	2	W ⁺	<i>Id</i>	<i>fm⁺/fm⁺</i>	0
Marsh Daisy	Pale willow green	1			w/w	<i>id⁺/id⁺</i>	<i>fm⁺/fm⁺</i>	0
Modern Langshan	Transparent White	1			W ⁺	<i>Id</i>	<i>fm⁺/fm⁺</i>	0
Norfolk Grey	Black	1			W ⁺	<i>id⁺/id⁺</i>	<i>fm⁺/fm⁺</i>	0
Old English Pheasant Fowl	Slate blue	1			W ⁺	<i>id⁺/id⁺</i>	<i>fm⁺/fm⁺</i>	0
Polish	White or bluish white	2			W ⁺	<i>Id</i>	<i>fm⁺/fm⁺</i>	0
Rhode Island Red	Yellow or red horn	1			w/w	<i>Id</i>	<i>fm⁺/fm⁺</i>	0
Scots Dumpy	Black or slate on black	1			W ⁺	<i>id⁺/id⁺</i>	<i>fm⁺/fm⁺</i>	0
Scots Grey	White or white with black spots	1			W ⁺	<i>Id</i>	<i>fm⁺/fm⁺</i>	0
Silkie	Lead	4			W ⁺	<i>id⁺/id⁺</i>	<i>Fm</i>	0
Spanish	Pale slate blue	1			W ⁺	<i>id⁺/id⁺</i>	<i>fm⁺/fm⁺</i>	0
Sultan	White or pale blue		1		W ⁺	<i>Id</i>	<i>fm⁺/fm⁺</i>	0
Sussex	White	4			W ⁺	<i>Id</i>	<i>fm⁺/fm⁺</i>	0
Totenko	Olive green		7		w/w	<i>id⁺/id⁺</i>	<i>fm⁺/fm⁺</i>	5
Villafranguina	Slate blue			2	W ⁺	<i>id⁺/id⁺</i>	<i>fm⁺/fm⁺</i>	0
Welssummer	Yellow		5		w/w	<i>Id</i>	<i>fm⁺/fm⁺</i>	0
White Star	Yellow		2		w/w	<i>Id</i>	<i>fm⁺/fm⁺</i>	0
Yurlov	Yellow or black			4	?	?	<i>fm⁺/fm⁺</i>	0
Junglefowls			PB	SI				
<i>Gallus gallus gallus</i>	Plumbeous brown to bluish grey	2	1		W ⁺	?		0
<i>Gallus gallus murghi</i>	Plumbeous brown to bluish grey	3			W ⁺	?		2
<i>Gallus gallus jabouillei</i>	Plumbeous brown to bluish grey			1	W ⁺	?		0
<i>Gallus lafayetti</i>	Brownish yellow	2	1		w/w	?		3
<i>Gallus varius</i>	White to pink	2			?	?		1
<i>Gallus sonneratii</i>	Shades of yellow to salmon red	2	1		w/w	?		2

Sources: RI, Roslin Institute; PB, Private Breeder; AD, AvianDiv; ILRI, International Livestock Research Institute; SI, Smurfit Institute); Genotype (? , unknown phenotype);

¹Delacour (1977), Roberts 1997, Smyth (1990), Scrivener (2006, 2009).

²Allele combinations inferred from shank colour and Smyth (1990).

these samples, DNA was extracted using University of Nottingham's in-house protocols. SNP data for a single individual each of *G. g. gallus*, *G. g. jabouillie*, *G. lafayetti* and *G. sonneratii* were provided by the Smurfit Institute, Dublin. For GP chicken, DNA was extracted from whole blood using the QuickExtract™ DNA extraction kit (Epicentre Biotechnologies) following the manufacturer's protocol. Phenotypes for the loci of interest were inferred from

literature and breed standards (Delacour 1977; Roberts 1997; Smyth 1990; Scrivener 2006, 2009), whereas those that were not fixed in any breed were recorded as unknown.

mtDNA studies

For mtDNA D-loop analysis, DNA was extracted from whole blood from 31 GP birds from a conservation flock of the

experimental farm of the University of Life Sciences, Lublin, Poland. This flock has been maintained as a closed population since 1960 and can be regarded as a true and pure representative of GP fowls. DNA was extracted using the DNeasy[®] Blood and Tissue kit (Qiagen). The first 800 bp of the mtDNA D-loop region was amplified via PCR and sequenced using primer pair L16750 (5'-AGGACTACG GCTTGAAAAGC-3') and CR1b (5'-CCATACACGCAAACC GTCTC-3') following Mwacharo *et al.* (2011).

Data analysis

mtDNA D-loop sequences

For each sample, two fragments were generated. These were edited manually using BIOEDIT 7.0 (Hall 1999) and then aligned against a reference sequence (GenBank accession number NC_007235; Nishibori *et al.* 2005) using CLUSTALX 2.0.12 (Thompson *et al.* 1997). Subsequent analyses were restricted to the first 397 bp of the D-loop spanning the first hypervariable region (HVR I). DNASP 4.0 (Rozas *et al.* 2003) was used to define the haplotypes. ARLEQUIN v3.5.1.2 (Excoffier & Lischer 2010) was used to estimate haplotype and nucleotide diversities and the mean number of pairwise differences and to infer historical population dynamics from mismatch distribution patterns. Departure of the observed mismatch distribution from the simulated model of demographic expansion was tested with chi-squared test of goodness of fit and Harpending's (1994) raggedness index 'r' following 10 000 coalescent simulations. These two tests were augmented with the Fu (1997) 'F_s' and Tajima (1996) 'D' coalescent-based estimators of neutrality and whose significance were also tested with 10 000 simulations. To assess the number of haplogroups in the GP fowl, a median-joining (MJ) network (Bandelt *et al.* 1999) was constructed using NETWORK 4.5 (fluxus-engineering.com). To portray the affinity of GP to other Eurasian chicken, the construction of the MJ network included 40 haplotypes downloaded from GenBank (Table S1). Nine were from Liu *et al.* (2006), 11 from Revay *et al.* (2010) and 20 from Dana *et al.* (2010). For this purpose, the central and most common haplotypes for the different clades/haplogroups observed in these three studies were selected to infer the maternal origins of GP within the geographic range of the wild ancestor, the Red junglefowl and their possible shared maternal ancestries with other European traditional breeds.

SNP analysis of genetic relationships, structure and association mapping

All samples were genotyped by a private company (DNA Landmarks, Inc.). SNP genotyping was performed using the 60K SNP Illumina iSelect[™] chicken microarray (Groenen *et al.* 2011; <http://www.illumina.com/>). Further details may be found in Wragg *et al.* (2012). STRUCTURE v2.3.3

(Pritchard *et al.* 2000) was used to evaluate the genetic structure of the study populations using a subset of the data that included a single individual selected at random from each breed and 40 000 randomly selected SNPs. A model assuming admixture between populations with unlinked loci was used for the analysis with a burn-in of 50 000 and Markov chain Monte Carlo repetitions of 100 000 assuming $1 \leq K \leq 5$ genetic clusters.

The approach used for association mapping has been described in detail in Wragg *et al.* (2012). Individuals with unknown phenotypes were excluded. SNP genotyping data were pruned using the GENABEL package (Aulchenko *et al.* 2007) for R (R Core Team 2009). Initially, the mean and standard deviation of identity-by-state (IBS) scores were calculated across all samples using the `ibs` function. The `check.marker` function was then used to prune the data set using the following attributes: `maf = 0.05`, `call = 0.9`, `per.id = 0.9`, `p.lev = 0`, `ibs.mrk = 'ALL'`, `ibs.threshold = ibscut`, `ibs.exclude = 'lower'`. The `ibscut` score was set to 0.7849959, a value that corresponds to the mean plus three times the standard deviation of IBS scores. All birds that failed this threshold are presented in Table 1 under the column heading QC fail (IBS). The `qtscore` function of GENABEL was used to perform tests of allelic associations, with reported *P* values corrected to one degree of freedom – equivalent to the Armitage test. Genotype frequencies were calculated across all samples, including those that initially failed QC, using the `summary.snp.data` function for each significant SNP mapping to previously published regions for *W* (Eriksson *et al.* 2008), *Id* and *Fm* (Dorshorst *et al.* 2010) loci. Frequencies were calculated for all birds and then independently for Silkie, junglefowl and GP.

Results

mtDNA D-loop analysis

Sequences spanning the first 397 bp of the mtDNA D-loop and which included the HVR I were analysed. From 31 sequences, three haplotypes defined by 10 polymorphic sites were generated (Fig. 1). For this study, haplotypes observed in GP are abbreviated 'GP' followed by a numeral, for example GP1 (Figs 1 and 2). Haplotype GP1 (with 23 sequences) was the most common, whereas GP2 (with five sequences) and GP3 (with three sequences) were the least frequently observed. Haplotype and nucleotide diversities for the GP fowl were 0.362 ± 0.210 and 0.0080 ± 0.0021 respectively, whereas the average number of nucleotide differences between sequences was 3.626 ± 1.888 .

The mismatch distribution pattern for the GP deviated significantly (calculated SSD = 0.217, *P* = 0.024) from a unimodal distribution pattern typical of populations that have experienced recent or past demographic expansions. The GP's distribution was characterised by three distinct

	222222333	
	1144567115	
	2736610055	N
NC007235	ATTTTCTCTT	1
GP1	23
GP2	GCCCCT·TCC	5
GP3	GCCCCTCTC·	3

Figure 1 Sequence variation of three haplotypes derived from 31 Green-legged Partridge-like chickens observed in the mtDNA D-loop region. Mutations are scored relative to the reference sequence (GenBank accession number NC007235; Nishibori *et al.* 2005). Dots (·) denote nucleotide identity with the reference sequence. Numbers in the first three lines read vertically represent the positions of the variable sites with respect to the reference sequence. The number of individuals sharing the same haplotypes is indicated in the right column by 'N'. These haplotypes have been deposited with the NCBI under accession numbers KC560148–KC560150.

peaks at 0, 2 and 9 mismatches respectively (Fig. S2). The failure to reject the null hypothesis was further supported by a non-significant Harpending's raggedness index ($r = 0.642$, calculated $P = 0.294$). Both Fu's F_S and Tajima's D statistics were positive ($F_S = 7.229$; $D = 1.410$), indicating a decline in population size in the recent past and/or balancing selection.

Together with the 40 haplotypes downloaded from GenBank, the MJ network revealed a grouping of the three GP haplotypes (GP1, GP2 and GP3) into two haplogroups (Fig. 2). The first haplogroup is made up of haplotype GP1 that is associated with clade B of Liu *et al.* (2006) (Table S2a). This haplotype is identical to haplotypes B1 (Liu *et al.* 2006), HIC11 (Revay *et al.* 2010) and B1 (Dana *et al.* 2010) observed in Eurasian, Hungarian and Dutch traditional and commercial breeds respectively (Table S2b). The haplogroup involving these haplotypes also was observed in commercial broilers and traditional chickens from North-West Europe by Muchadeyi *et al.* (2008) and in a single individual from Ethiopia by Mwacharo *et al.* (2011). Liu *et al.* (2006) proposed Yunnan province in China and/or adjacent areas as the possible centre of origin of this haplogroup. The second haplogroup is made up of haplotypes GP2 and GP3 (Fig. 2). These two are closely related to haplotype E1 from clade E of Liu *et al.* (2006), haplotypes HIC1–HIC9 observed in Hungarian chicken (Revay *et al.* 2010) and haplotypes E1–E12 observed in Dutch traditional and commercial breeds (Dana *et al.* 2010) (Table S2 (a) and S2(b)). Haplotypes from this haplogroup predominate amongst African village chicken from Zimbabwe, Malawi, Sudan (Muchadeyi *et al.* 2008), Nigeria (Adebambo *et al.* 2010) and Kenya, Uganda, Ethiopia and Sudan (Mwacharo *et al.* 2011). Liu *et al.* (2006) observed this haplogroup in Europe, the Middle East and India and suggested that it could have its roots in the Indian subcontinent. Only two (B and E) of the nine (A to I) clades observed in Eurasian chicken by Liu *et al.* (2006) are

found in the GP. These results suggest that, like other European chicken from Hungary and the Netherlands as well as commercial breeds, the GP has two maternal genetic backgrounds, one derived likely from the Indian subcontinent and the other possibly from southern China or northern South-East Asia.

Genetic structure and association mapping

We partitioned genetic variation amongst the study breeds using the parametric genetic admixture algorithm implemented in STRUCTURE 2.3.3 (Pritchard *et al.* 2000). Between one and five clusters (K) were tested assuming admixture between populations and linkage equilibrium between loci. A graphic representation of the estimated membership coefficient for each individual to each cluster at $1 \leq K \leq 5$ is shown in Fig. 3. An estimate of the posterior probability of the data ($\ln[P(D)]$) indicated that $K = 3$ was the most optimal. With the exception of Hamburg, Appenzeller, Spanish, Brahma and Cochon, all the other breeds appear to be admixed among the three clusters with the proportion of each cluster varying across breeds. We interpret these three clusters to represent European, East Asian and South Asian genetic backgrounds consistent with known breed histories (Scrivener 2006, 2009).

For each locus (W , Id , Fm), binomial association tests were performed on the pruned data set. After pruning, 29 birds with yellow shanks (w/w) and 50 with white shanks (W^+) were retained for analysis. For the inhibitor of dermal melanin phenotype, two tests were performed as the phenotypes were uncertain based on inference from breed standards. In the first test, the phenotypes were assigned as per Table 1, resulting in 39 birds exhibiting the presence of dermal melanin (id^+/id^+) and 40 exhibiting its absence (Id). In the second test, only the Silkie birds and the GP were assigned the id^+/id^+ phenotype, whilst all the other birds were assigned the Id phenotype, resulting in nine birds exhibiting the presence of dermal melanin (id^+/id^+) and 74 exhibiting its absence (Id). Finally, dermal hyperpigmentation (Fm) is only present in the Silkie resulting in four birds with the Fm phenotype and 79 of the wild type (fm^+/fm^+).

Results of the top five significant outcomes from each binomial test are presented in Table 2. These are in agreement with previously published mapping data for the W and Fm loci (Eriksson *et al.* 2008; Dorshorst *et al.* 2010). In the table, those not highlighted are considered false positives for the purpose of this study. They are likely the result of the limited number of samples involved in each test. As already indicated, the W and Fm results are consistent with previous findings, although the mapping of the Id locus appears inconclusive. By considering only the Silkie and GP birds to possess id^+/id^+ and all the other birds to possess Id , the results agree with those of Dorshorst *et al.* (2010). However, applying the allelic interactions that are

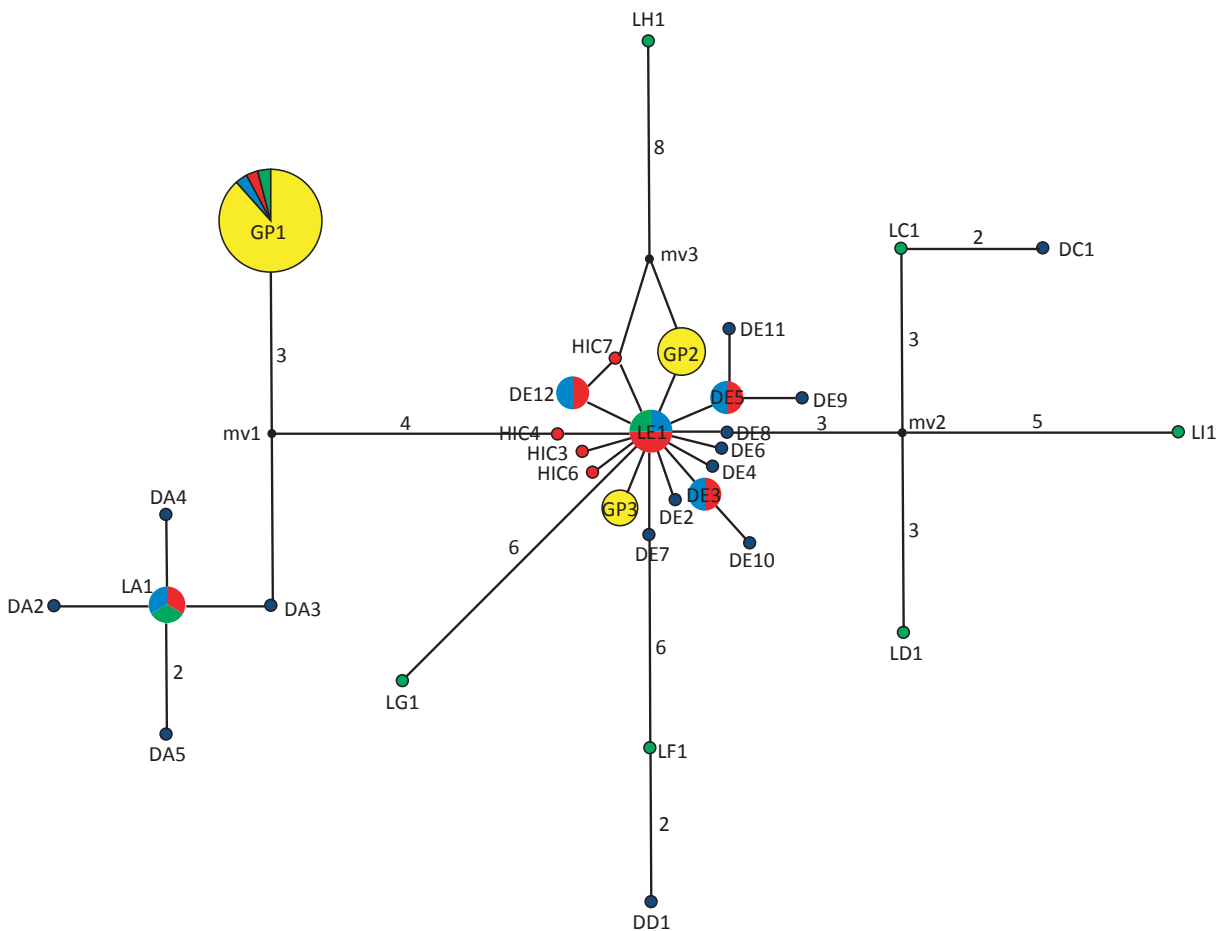


Figure 2 Median-Joining network of the three haplotypes of Green-legged Partridge-like fowl (GP1, GP2, GP3) and 40 haplotypes downloaded from the GenBank (Supplementary Table 1) and representing Eurasian (haplotypes starting with letter 'L'; Liu *et al.* 2006), Hungarian (haplotypes starting with 'HIC' Revay *et al.* 2010) and Dutch fancy and commercial breeds (haplotypes starting with letter 'D'; Dana *et al.* 2010) of chicken. Circled areas are proportional to haplotype frequencies. Median vectors are represented by 'mv'. Unless shown by using numbers, all interlinked haplotypes are separated by a single mutation. Key to colour codes: Blue = Haplotypes from Dutch fancy and commercial breeds, Green = Haplotypes of Eurasian chicken, red = Haplotypes of Hungarian chicken, Yellow = Haplotypes of Green-legged Partridge-like chicken. Circles in which more than two of these colours occur means that the haplotypes are shared between the respective groups of chicken represented by each colour. With the exception of haplotype GP1 which is shared between Eurasian, Dutch fancy and Hungarian indigenous chickens, for haplotypes shared between Eurasia, Dutch and Hungarian indigenous chickens, haplotype names (those starting with letter 'L' from Liu *et al.* (2006) are adopted while for haplotypes shared between Dutch fancy breeds and Hungarian indigenous chickens, haplotype names (those starting with letter 'D' from Dana *et al.* (2010) are used.

known to result in specific shank colours (Table 3) to the breed standards of each bird used for the test fails to map any significant associations to *Id*.

After association testing, genotype frequencies for significant SNPs, identified in the previously mapped regions for each locus, were estimated (Fig. 4). Genotype frequencies were calculated as per the phenotypes recorded in Table 1 for all birds excluding the GP, junglefowl and the Silkie. The same alleles linked to yellow skin phenotype at the *W* locus were found to be present in the GP (Fig. 4), supporting that the locus contributes to the green leg/shank phenotype. As the junglefowl group included both white- and yellow-skinned species, a similar number of homozygotes were observed for each allele at the same locus. The genotype frequencies at the *Id* locus mapped SNP were less conclusive

across all birds. However, a similar allele frequency at the SNP in the region previously mapped for the *Id* locus is observed in Silkie and GP. Finally, the SNP frequencies mapped for the *Fm* locus confirmed the allele as absent in the GP and in 96% of non-Silkie breeds, whilst one allele was observed in both heterozygous and homozygous forms in the Silkie, supporting a dominant mode of inheritance. None of the alleles for *Id* or *Fm* were observed in any of the junglefowl samples.

Discussion

The GP chicken is an indigenous Polish breed, which at one point was threatened with extinction but currently is maintained in two independent conservation schemes

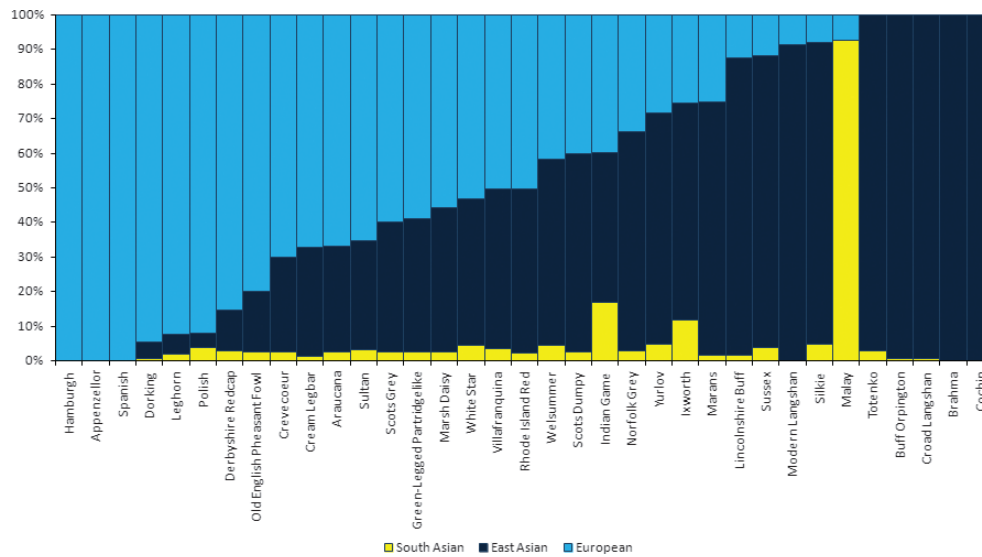


Figure 3 Genetic structure of the nuclear genome (autosomal and Z chromosome) for 35 breeds revealed from the analysis of 60K SNP microarray.

Table 2 Significant results following association tests for the loci *W*, *Id* and *Fm* using the 60K SNP chip.

Locus	Chromosome	Position	<i>P</i> -value
<i>W</i>		24	6279104
		24	6273988
		24	6261633
		12	11022567
		1	19149987
<i>Id</i>	Only GP and Silkie <i>id⁺/id⁺</i>	Z	72985598
		2	138132039
		3	41040834
		12	7186919
		Z	33105782
	Id phenotypes as recorded in Table 1	9	12500113
		13	5917154
		1	158528977
		13	5963541
		5	54486865
	<i>Fm</i>		20
		1	56353716
		28	2986479
		2	68439874
		2	72080734

Note. Chromosomal positions based on *galGal3* genome assembly. Results not highlighted are considered false positives for the purpose of this study and are likely the result of the limited number of samples involved in each test.

Table 3 Interactions of *W* and *Id* on shank colour.

Shank colour	<i>Id</i>	<i>W</i>
White	<i>Id</i>	<i>W⁺</i>
Yellow	<i>Id</i>	<i>w/w</i>
Green (willow)	<i>id⁺/id⁺</i>	<i>w/w</i>
Slate/Blue/Black	<i>id⁺/id⁺</i>	<i>W⁺</i>

Adapted from Smyth (1990).

(Witkowski *et al.* 2009). In this study, we investigate the origin of the breed and describe its genetic structure using the maternally inherited mtDNA D-loop and genome-wide SNP analysis (Groenen *et al.* 2011; Wragg *et al.* 2012). Using the latter, we demonstrate the likely genetic control of the reseda green-coloured shank phenotype, a standard trait for the breed. Our study reveals two distinct mtDNA D-loop haplogroups and three autosomal genetic clusters,

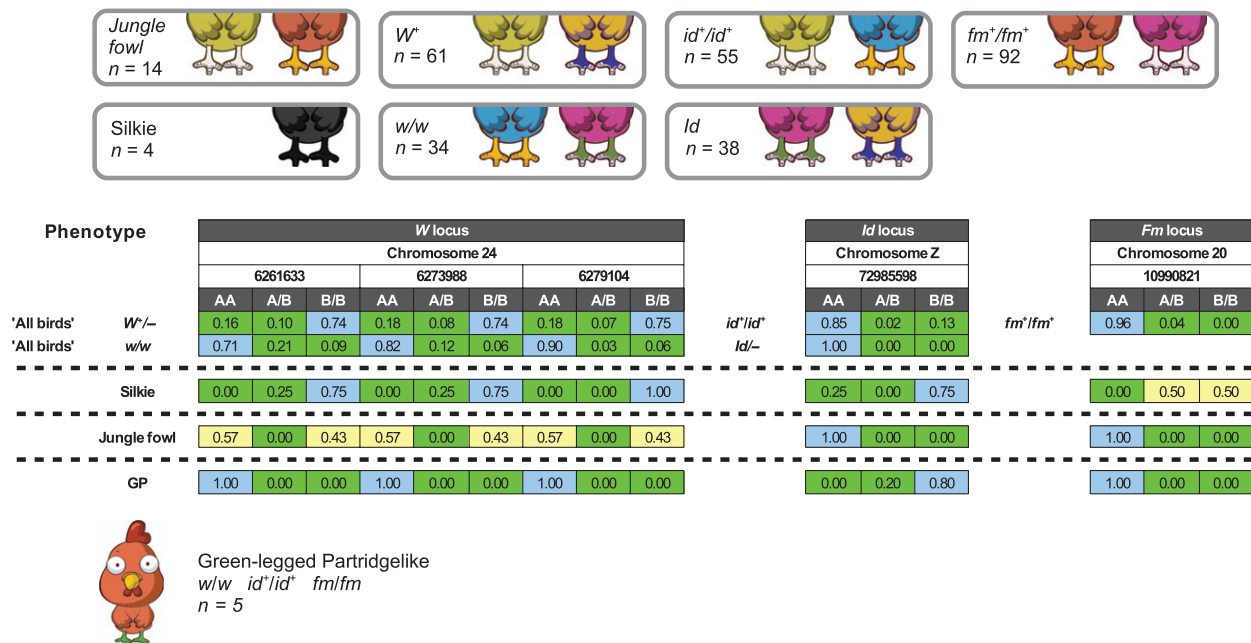


Figure 4 Genotype frequencies at *W*, *Id* and *Fm* loci-mapped SNPs controlling shank pigmentation.

indicating that the GP fowls, together with most European traditional breeds, have mixed maternal and possibly paternal ancestries. Due to its admixed nature, the expectation was that the breed would display relatively high genetic diversity. However, its haplotype and nucleotide diversities are lower than those observed in free-range scavenging village chicken from Asia (Liu *et al.* 2006; Silva *et al.* 2008) and Africa (Muchadeyi *et al.* 2008; Mwacharo *et al.* 2011). The contraction in population size (bottleneck) experienced by the breed in the 1970s, supported by positive Tajima's *D* and Fu's *F_S* statistics obtained in this study, might be responsible for the low levels of maternal diversity in the GP. The subsequent development and multiplication of the breed, from a small founder population with a narrow genetic base and the long-term maintenance of the breed as a closed flock, are also possible contributing factors.

We observed at least two mtDNA D-loop haplogroups (B and E) and three autosomal genetic clusters based on the 60K SNP chip analysis in the GP breed. The two mtDNA D-loop haplogroups are associated with similar haplogroups observed in other traditional breeds from Europe and Asia (Fig. 2). Similarly, the three autosomal genetic clusters are also present in other European traditional and fancy breeds of chicken included in the current study (Fig. 3). These results support a mixed maternal and paternal ancestry for the European breeds in general.

Moreover, our results suggest a distinct origin and history of the two haplogroups and the three autosomal genetic clusters observed in GP and European traditional breeds (Figs 2 and 3; Tables S2a and S2b). Indeed, the two mtDNA D-loop haplogroups are related to geographically distinct

Asian mtDNA haplogroups, primarily deriving from Yunnan Province in China and/or adjacent regions (haplogroup B (GP1)) and South Asia [haplogroup E (GP2 and GP3)] (Liu *et al.* 2006). On the other hand, the three autosomal genetic clusters indicate the GP fowl and most European traditional breeds analysed here to be genetically admixed (Fig. 3). The presence of several different genetic backgrounds in the GP fowl and most European traditional breeds raises interesting questions on how, when and where such admixture could have taken place.

Recent cross-breeding with commercial breeds may have contributed to the admixture. Both commercial broiler and egg layers share the same mtDNA haplogroups B and E with the GP and European traditional breeds (Fig. 2; Dana *et al.* 2010). However, we believe this to be unlikely for the GP fowl. Indeed, there is no historical record indicating that recent commercial introgression might have occurred in the GP fowl, although it might have remained undocumented. Also, such introgression does not explain the presence of the reseda green shank phenotype in the breed. Furthermore, the reseda green shank phenotype was described prior to the development of commercial breeds (Pietruski 1866; Graff 1887).

Alternatively, the genome admixture present in the GP might have originated prior to the standardisation of the breed or during its formation. Our association analysis confirms that an interaction between the recessive yellow skin allele and the wild recessive allele at the dermal melanin locus present in the Silkie is responsible for the green shank phenotype. The failure to map a significant association at the *Id* locus using all breeds based on expected allelic combinations raises an interesting question:

might there be multiple alleles for the *Id* locus as was also noted by Smyth (1990). The alternatives are that the phenotypes of the birds used do not meet breed standards or that none of the SNPs on the panel are in strong linkage disequilibrium with the locus in the other breeds analysed.

Eriksson *et al.* (2008) revealed that the allele responsible for the yellow skin phenotype is likely to have introgressed into domestic chicken from the Grey junglefowl *G. sonneratii*, which is endemic on the Indian subcontinent. The origin of the non-wild-type allele at the *Id* locus remains uncertain. All the junglefowl tested shared the same alleles at the SNP that was mapped for the *Id* locus in the Silkie and GP, suggesting that *Id* is possibly the result of a *de novo* mutation in domestic stocks.

The demonstration here that the reseda green shank phenotype in the GP fowl arises from the interaction of recessive alleles at the *W⁺* and *Id* loci further supports at least two genetic backgrounds in the breed. These may have already been present in the ancestral Asian domestic populations or could have arisen following admixture on the European continent of chickens from different origins and/or introductions.

Literature suggests the geographic origin of the GP fowl to be the Mediterranean region, specifically Italy, from where it diffused to Poland in the 16th century (Witkowski *et al.* 2009). This seems to suggest that its hybrid origin may have occurred in Italy. In this respect, it is worth mentioning that the Sicilian Buttercup, which originally came from Sicily, has willow green shanks and feet (Roberts 1997; Scrivener 2006). The presence of green shanks in the GP fowl might have resulted from cross-breeding with white leg chicken carrying the sex-linked dermal melanin (*id⁺/id⁺*) alleles. Such chicken might have been, for example, Silkie-type birds known to have been present across Europe since at least the 16th century (Corti & Corti 2009).

Finally, in contrast to chicken from Africa and Asia, the diversity of indigenous European chicken is found today in fancy or traditional breeds rather than in village chicken populations. Often with small effective population sizes, these fancy breeds are exposed to various threats (FAO 2009) and are at risk of extinction (FAO 2007). There is a need to ensure that they are sustainably conserved. Such an action will help to achieve two major goals: 1) provide model organisms for academic and scientific research aimed, for example, at understanding the inheritance pattern and interaction pathways of genes underlying admixed phenotypes and 2) conserve the legacy of the intellectual ability, knowledge and efforts of past breeders in creating such unique phenotypes.

Acknowledgements

We thank Ewa Gorzycka for kindly providing photographs of the GP chicken legs/shanks and to two anonymous reviewers for their insightful suggestions. We acknowledge

Dr A. Witkowski from University of Life Sciences in Lublin, Poland, for collecting the blood samples of Green-legged Partridge-like fowls, the Roslin Institute, the AVIANDIV project and fancy breeders for providing either blood or DNA samples of other breeds used in this study. Thanks also to Dan Bradley at the Smurfit Institute for providing genotyping data for some of the junglefowl.

References

- Adebambo A.O., Mobegi V.A., Mwacharo J.M. *et al.* (2010) Lack of phylogeographic structure in Nigerian village chicken revealed by mitochondrial DNA D-loop sequence analysis. *International Journal of Poultry Science* **9**, 503–7.
- Aulchenko Y.S., Ripke S., Isaacs A. & van Duijn C.M. (2007) GENABEL: an R library for genome-wide association analysis. *Bioinformatics* **23**, 1294–6.
- Bandelt H.J., Forster P. & Röhl A. (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology Evolution* **16**, 37–48.
- Bateson W. (1902) Experiments with poultry. *Reports to the Evolution Committee of the Royal Society* **1**, 87–124.
- Corti E. & Corti F. (2009) *The Chicken of Ulysses Aldrovandi*. A translation of Ulysses Aldrovandi's "Fourteenth Book of Ornithology Concerning Domestic Dust Bathing Fowls". Available at: <http://archive.org/details/TheChickenOfUlyssesAldrovandi>.
- Dana N., Megens H-J., Crooijmans R.P.M.A., Hanotte O., Mwacharo J., Groenen M.A.M. & van Arendonk J.A.M. (2010) East Asian contributions to Dutch traditional and western commercial chickens inferred from mtDNA analysis. *Animal Genetics* **42**, 125–33.
- Delacour J. (1977) *The Pheasants of the World*. 2nd edn. Spur Publications Co. Liss, Hants, England. 395pp.
- Dorshorst B., Okimoto R. & Ashwell C. (2010) Genomic regions associated with dermal hyperpigmentation, polydactyly and other morphological traits in the silkie chicken. *Journal of Heredity* **101**, 339–50.
- Dorshorst B., Molin A-M., Rubin C-J., Johansson A.M., Strömstedt L., Pham M., Chen C., Hallböök F., Ashwell C. & Andersson L. (2011) A complex rearrangement involving the *endothelin 3* locus causes dermal hyperpigmentation in the chicken. *PLoS Genetics* **7**, e1002412, doi: 10.1371/journal.pgen.1002412.
- Dunn L.C. (1925) A new lethal gene in fowls. *American Naturalist* **57**, 345–9.
- Dunn L.C. & Jull M.A. (1927) On the inheritance of some characteristic of the Silkie fowl. *Journal of Genetics* **19**, 27–63.
- Eriksson J., Larson G., Gunnarsson U. *et al.* (2008) Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken. *PLoS Genetics* **4**, e1000010, doi: 10.1371/journal.pgen.1000010.
- Excoffier L. & Lischer H.E.L. (2010) ARLEQUIN suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**, 564–7.
- FAO. (2007) *The State of the World's Animal Genetic Resources for Food and Agriculture*. Commission on Genetic Resources for Food and Agriculture. Food and Agriculture Organization of the United Nations, Rome, Italy.
- FAO. (2009) *Threats to Animal Genetic Resources – Their Relevance, Importance and Opportunities to Decrease their Impact*. Background

- Study Paper No. 50. Commission on Genetic Resources for Food and Agriculture, Rome. Available at: <http://ftp.fao.org/docrep/fao/meeting/O17/ak572e.pdf>.
- Fu Y.-X. (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**, 915–25.
- Fumihito A., Miyake T., Sumi S., Takada M., Ohno S. & Kondo N. (1994) One subspecies of the red jungle fowl (*Gallus gallus gallus*) suffices as the matriarchic ancestor of all domestic breeds. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 12505–9.
- Fumihito A., Miyake T., Takada M., Shingu R., Endo T., Gojobori T., Kondo N. & Ohno S. (1996) Monophyletic origin and unique dispersal patterns of domestic fowls. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 6792–5.
- Graff K. (1887) Hodowla zwierząt. Wydawnictwa rolnicze pod redakcją Aleksandra Trylskiego, tom 2 In: *Animals Breeding*. (Ed. by A. Trylski), Agriculture publications, volume 2.
- Groenen M.A., Megens H.J., Zare Y., Warren W.C., Hillier L.W., Crooijmans R.P.M.A., Vereijken A., Okimoto R., Muir W.M. & Cheng H.H. (2011) The development and characterization of a 60K SNP chip for chicken. *BMC Genomics* **12**, 274. Available at: <http://www.biomedcentral.com/1471-2164/12/274>.
- Hall T.A. (1999) BIOEDIT: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**, 95–8.
- Harpending H. (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology* **66**, 591–600.
- Hutt F.B. (1949) *Genetics of the Fowl*. McGraw Hill Books Co., New York.
- Krawczyk J. & Szefer M. (2012) Quality of eggs from Greenleg Partridge hens maintained in organic backyard production system. *EAAP Book of Abstracts* No. 18, pp. 153.
- Krawczyk J., Cywa-Benko K. & Wezyk S. (2005) Effect of housing system on egg yolk cholesterol levels in native breeds of hens. International Conference “First International Scientific Poultry Days”, Nitra, Slovakia, 12–14 September, pp. 18–20 ISBN 80-8069-576-8.
- Liu Y.P., Wu G.S., Yao Y.G., Miao Y.W., Luikart G., Baig M., Beja-Pereira A., Ding Z.L., Palanichamy M.G. & Zhang Y-P. (2006) Multiple maternal origins of chickens: out of the Asian jungles. *Molecular Phylogenetics and Evolution* **38**, 12–9.
- McGibbon W.H. (1974) A shank color mutation in Cornell random bred S. C. white leghorns. *Poultry Science*, **53**, 1251–3.
- McGibbon W.H. (1979) Green shanks and adult mortality in chickens. *Journal of Heredity* **70**, 44–6.
- Miao Y-W., Peng M-S., Wu G-S. *et al.* (2012) Chicken domestication: an updated perspective based on mitochondrial genomes. *Heredity* (Edinb), **110**, 227–82.
- Muchadeyi F.C., Eding H., Simianer H., Wollny C.B.A., Groeneveld E. & Weigend S. (2008) Mitochondrial DNA D-loop sequences suggest a Southeast Asian and Indian origin of Zimbabwean village chickens. *Animal Genetics* **39**, 615–22.
- Mwacharo J.M., Bjørnstad G., Mobegi V., Nomura K., Hanada H., Amano T., Jianlin H. & Hanotte O. (2011) Mitochondrial DNA reveals multiple introductions of domestic chicken in East Africa. *Molecular Phylogenetics and Evolution* **58**, 374–82.
- Nishibori M., Shimogiri T., Hayashi T. & Yasua H. (2005) Molecular evidence for hybridization of species in the genus *Gallus* except for *Gallus varius*. *Animal Genetics* **36**, 367–75.
- Pietruski S.K. (1866) *Historia naturalna i hodowla ptaków zabawnych i użytecznych. Tom IV. Kury swojskie*. Lwów, Nakładem Karola Wilda. (Natural History and risible and usable birds breeding. Volume IV. Familiar fowls).
- Pritchard J.K., Stephens M. & Donnelly P. (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–59.
- R Core Team (2009). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: <http://www.R-project.org/>.
- Revay T., Bodzsar N., Mobegi V.E., Hanotte O. & Hidas A. (2010) Origin of Hungarian indigenous chicken breeds inferred from mitochondrial DNA D-loop sequences. *Animal Genetics* **41**, 548–50.
- Roberts V. (1997) *British Poultry Standards*, 5th edn. (Ed. by V. Roberts) Victoria Roberts, Blackwell Science, Oxford, viii, pp. 368.
- Rozas J., Sanchez-DelBarrio J.C., Messeguer X. & Rozas R. (2003) DNASP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**, 2496–7.
- Scrivener D. (2006) *Rare Poultry Breeds*. Crowood, Ramsbury, Wiltshire.
- Scrivener D. (2009) *Popular Poultry Breeds*. Crowood, Ramsbury, Wiltshire.
- Silva P., Guan X., Ho-Shing O., Jones J., Xu J., Hui D., Notter D. & Smith E. (2008) Mitochondrial DNA-based analysis of genetic variation and relatedness among Sri Lankan indigenous chickens and the Ceylon Junglefowl (*Gallus lafayetti*). *Animal Genetics* **40**, 1–9.
- Siwek M., Sławińska A., Nieuwland M., Witkowski A., Zieba G., Minozzi G., Knol E.F. & Bednarczyk M. (2010) A quantitative trait locus for a primary antibody response to keyhole limpet hemocyanin on chicken chromosome 14 - confirmation and candidate gene approach. *Poultry Science* **89**, 1850–7.
- Siwek M., Szyda J., Sławińska A. & Bednarczyk M. (2012) Detection of two QTL on chicken chromosome 14 for keyhole lymphet heamocyanin. *Journal of Applied Genetics* **53**, 115–9.
- Sławińska A., Witkowski A., Nieuwland M., Minozzi G., Bednarczyk M. & Siwek M. (2011) Quantitative trait loci associated with the humoral innate immune response in chickens were confirmed in a cross between Green-Legged Partridge-like and White Leghorn. *Poultry Science* **90**, 1909–15.
- Smyth J.R. Jr. (1990) Genetics of plumage, skin and eye pigmentation in chickens. In: *Poultry Breeding and Genetics: Developments in Animal and Veterinary Sciences*, **22** (Ed. by R.D. Crawford), pp. 109–67. Elsevier B.V, The Netherlands.
- Somes R.G. Jr. (1988) International Registry of Poultry Genetic Stocks. Storrs Agricultural Experiment Station. Paper No. 29. Available at: <http://digitalcommons.uconn.edu/saes/29>.
- Tajima F. (1996) The amount of DNA polymorphism maintained in a finite population when the neutral mutation rate varies among sites. *Genetics* **143**, 1457–65.
- Thompson J.D., Gibson T.J., Plewniak F., Jeanmougin F. & Higgins D.G. (1997) The CLUSTALX windows interface. Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **24**, 4876–82.
- West B. & Zhou B.-X. (1988) Did chickens go north? New evidence for domestication. *Journal of Archaeological Science* **15**, 515–33.

Witkowski A., Kasperczyk K., Sławińska A. & Bednarczyk M. (2009) The breed of the month: Green-legged Partridge-like. An old native Polish hen. *GlobalDiv Newsletter* 9, 12–5.

Wragg D., Mwacharo J.M., Alcalde J.A., Hocking P.M. & Hanotte O. (2012) Analysis of genome-wide structure, diversity and fine mapping of Mendelian traits in traditional and village chickens. *Heredity* 109, 6–18.

Supporting information

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

Figure S1 Example of reseda green and yellow shank phenotypes in chicken. Picture courtesy of Ewa Gorzycka.

Figure S2 Mismatch distribution pattern for the mtDNA D-loop sequence haplotypes generated from 31 Green-legged Partridge-like chickens.

Table S1 Accession numbers for the 40 haplotypes that were downloaded from the GenBank and included in the study.

Table S2 a. Clustering patterns between the three haplotypes observed in Green-legged Partridge-like chickens and those downloaded from the GenBank as observed in the MJ Network; b. Clustering patterns for the three haplotypes observed in Green-legged Partridge-like chickens and those downloaded from the GenBank on the MJ Network.