Supplementary information

Large-scale genome-wide SNP analysis reveals the rugged (and ragged) landscape of global ancestry, phylogeny, and demographic history in chicken breeds

Natalia V. DEMENTIEVA[1], Yuri S. SHCHERBAKOV[1], Olga I. STANISHEVSKAYA[1], Anatoly B. VAKHRAMEEV[1], Tatiana A. LARKINA[1], Artem P. DYSIN[1], Olga A. NIKOLAEVA[1], Anna E. RYABOVA[1], Anastasiia I. AZOVTSEVA[1], Olga V. MITROFANOVA[1], Grigoriy K. PEGLIVANYAN[1], Natalia R. REINBACH[1], Darren K. GRIFFIN[2], Michael N. ROMANOV[2,3]

[1]*Russian Research Institute of Farm Animal Genetics and Breeding – Branch of the L. K. Ernst Federal Research Centre for Animal Husbandry, Pushkin, St. Petersburg, 196601, Russia*
[2]*School of Biosciences, University of Kent, Canterbury, CT2 7NJ, UK*
[3]*L. K. Ernst Federal Research Center for Animal Husbandry, Dubrovitsy, Podolsk, Moscow Oblast, 142132, Russia*

Data S2: Genetic diversity in small populations of chicken gene pool breeds and their crossbreds

Like their wild ancestors, farm animal species also evolve, although their evolutionary process, as a result of human intervention, has been greatly accelerated. During this process, the genetic diversity inherent in the totality of individuals that make up a domesticated biological species was largely redistributed among many artificially created and propagated breeds. Depending on the purpose of use and conditions of keeping, breeds have a part of the common genetic material of a species, which ensures their best adaptability to the environment (e.g., Fedorova et al., 2022) and is expressed in certain phenotypic (productive) traits. In addition to the involvement of previously formed gene complexes in the genomes of poultry species, new alleles that appear due to mutations are likely to be implicated in the evolutionary process. A number of countries have adopted national programs for the conservation of animal genetic resources. International organizations such as the European Commission, Food and Agriculture Organization (FAO) and many private foundations are also active in this matter. Maintaining small gene pool populations of poultry in a viable state requires a deeper understanding of the genetic processes in such populations. Providing the necessary level of diversity contributes to the evolutionary stability of the system, which can be ascertained by taking into account simultaneously many genetic loci in the genome. The chicken genome sequencing effort (International Chicken Genome Sequencing Consortium, 2004) and subsequent discovery of numerous SNPs enabled to look at the fundamental problem of genetic diversity and breeding outcomes in poultry in more detail, i.e., at the molecular and genomic level.

Innovative genomic and post-genomic technologies facilitate the opportunities for creating specialized lines and new poultry crosses, the broad exploitation of which will ultimately help expand the assortment of, and gradually improve the quality of, poultry products offered. At the moment, there are no specific methodological recommendations regarding which breeds and populations should be preserved, what should be the system and methods of conservation, especially considering the fact that all these breeds have a lower performance. The task of the present population analysis was to correctly pick the selection scheme to achieve the maximum effect, which implies obtaining commodity or aesthetic products from the selection subjects. For this purpose, we analyzed the genotyping data and calculated a number of population genetic parameters (e.g., F_{ST} statistics, heterozygosity, inbreeding, interbreed differences, etc.) and estimated an overall breed phylogeny. The second area of research being developed during the implementation of this research project was the creation of interbreed hybrids from crosses between several gene pool and highly productive breeds of chickens including the original old and local breeds. A selection of groups of meat-type and dual purpose chicken breeds and populations was carried out. At the age of 63 days, blood samples were taken from the crossbreds for further DNA isolation and genotyping using SNP chips.

At this stage of the work, which was connected with the development of crosses based on gene pool breeds, the following crosses were produced, and two-way crossbreds were obtained (with amount of collected samples shown):

- Tsarskoye Selo (Ts)×Sussex Light (SL): 93 birds (53 males and 40 females);
- Brahma Light (BL)×SL: 94 birds (50 males and 44 females);
- SL × Amrock (Ar): 91 birds (51 males and 40 females);
- Uzbek Game (UG)×Ar: 83 birds (38 males and 45 females).

Subsequently, the following three-way crossbreds were obtained, in two of which the meat-type White Cornish (WC; Fig. S2-1) breed was used (the number of samples collected is indicated in the parentheses):

- WC×(BL×SL) (7 males and 7 females);
- WC×(SL×Ar) (7 males and 7 females);
- Ts×(SL×Ar) (7 males and 7 females).

DNA samples were analyzed using specialized Illumina SNP chips.

Fig. S2-1. A White Cornish chicken.

Next, based on the genotyping of DNA samples using Illumina SNP chips, the genetic variability of the breeds was analyzed. Preliminary filtering and exclusion of SNPs from the database was carried out using the PLINK 1.9 program according to the scheme: –maf 0.01; –geno 0.1; –hwe 0.0001; --qual-threshold 0.8. As a result, 41,445 SNPs were used in the analysis. Calculations of inbreeding (F_{IS} and F_{ROH}), heterozygosity, LD level, number and length of runs of homozygosity (ROH) were performed, and data were obtained on genetic diversity within populations and groups of crossbred birds. All analyzed indicators (Tables S2-1 and S2-2) characterize the genetic diversity within breeds and collectively provide an adequate assessment of the state of small populations. In most cases, groups with higher inbreeding had lower heterozygosity, higher F_{ROH}, and higher ROH metrics. This was characteristic of an inbred strain of the Russian White breed (RWS; DNA samples archived in 2001), Brahma Buff (BB), and BL. In particular, as shown in Fig. S2-2a, the RWS subpopulation had higher F_{IS} (0.53), LD (0.52), and F_{ROH} (0.32). Higher homozygosity was noted in the BB and BL chicken populations. All interbreed crossbreds (Table S2-2, Figs. S2-2a and S2-2b) had very low inbreeding rates, for both F_{IS} and F_{ROH}, and number of ROH was significantly lower (1 to 5) as compared to the parent breeds. The LD level in two- and three-way crossbreds did not change as compared to purebred birds, which may be due to the limited number of roosters (2 to 3) used to produce crossbreds. The level of F_{ROH} when breeding pure breeds ranged from 0.1 to 0.2. With a small introductory crossing with related breeds and breeding *inter se*, a decrease in F_{ROH} below 0.1 was observed. But some breeds, such as UG, had higher heterozygosity, lower levels of inbreeding and lower homozygosity. Perhaps, this is due to the origin of this breed from local chickens. When maintaining a chicken gene pool collection, the exchange between subpopulations is rather necessary and obligatory for the preservation of breeds, provided that they themselves are kept pure.

Table S2-1. Results of calculations of inbreeding, heterozygosity and LD levels in purebred populations (subpopulations)

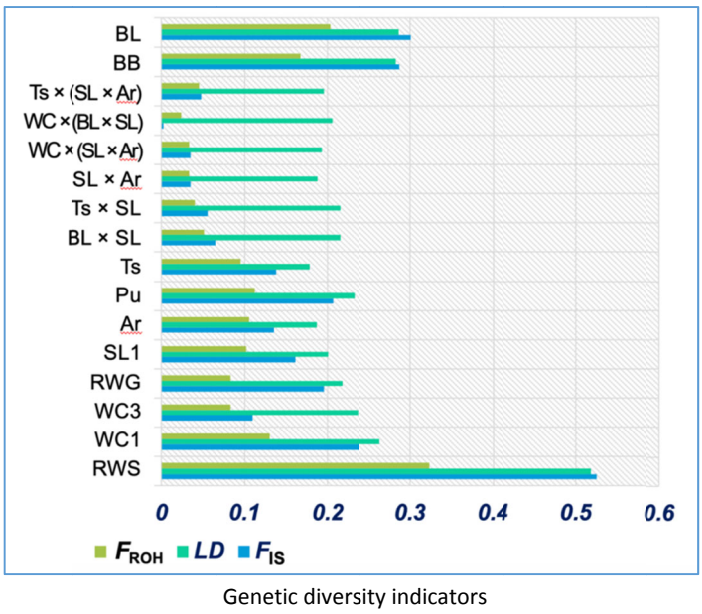
Breed (subpopulation)	Code	*n*	F_{IS}	*H*	LD
Russian White (subpopulation 2)	RWS	6	0.024±0.014	0.354±0.007	0.518±0.001
Russian White (subpopulation 3)	RWP	11	0.009±0.006	0.351±0.005	0.225±0.001
White Cornish (subpopulation 1)	WC1	22	0.004±0.002	0.349±0.002	0.263±0.001
White Cornish (subpopulation 2)	WC2	18	0.005±0.002	0.377±0.003	0.230±0.001
White Cornish (subpopulation 3)	WC3	19	0.003±0.002	0.372±0.002	0.237±0.001
Russian White (subpopulation 1)	RWG	19	0.056±0.010	0.324±0.003	0.218±0.001
Cochin Blue	CBl	18	0.017±0.007	0.354±0.005	0.212±0.001
Moscow Game	MG	20	0.026±0.006	0.359±0.0031	0.143±0.001
Sussex Light (subpopulation 1)	SL1	20	0.007±0.004	0.358±0.004	0.201±0.001
Uzbek Game	UG	19	0.010±0.007	0.373±0.008	0.155±0.001
Amrock	Ar	20	0.016±0.005	0.357±0.004	0.187±0.001
Yurlov Crower	YC	20	0.032±0.014	0.370±0.009	0.192±0.001
Pushkin	Pu	20	0.010±0.004	0.357±0.004	0.232±0.001

Tsarskoye Selo	Ts	20	0.009±0.006	0.357±0.0031	0.178±0.001
Plymouth Rock Barred	PRB	19	0.018±0.007	0.357±0.005	0.218±0.001
Brahma Buff	BB	20	0.023±0.011	0.356±0.0073	0.283±0.001
Brahma Light	BL	20	0.038±0.012	0.339±0.0110	0.286±0.001
Leningrad Mille Fleur	LMF	21	0.005±0.0023	0.362±0.005	0.189±0.001
Naked Neck	NN	19	0.002±0.002	0.362±0.002	0.211±0.001
Minorca Black	MB	19	0.003±0.003	0.379±0.005	0.187±0.001
Pavlov Spangled	PS	20	0.008±0.003	0.374±0.003	0.224±0.001
Poltava Clay	PC	17	0.024±0.012	0.359±0.005	0.163±0.001
Red White-tailed Dwarf	RWD	18	0.024±0.010	0.359±0.006	0.239±0.001
Australorp Black	AoB	9	0.006±0.003	0.378±0.004	0.208±0.001
Pavlov White	PW	15	0.025±0.013	0.345±0.007	0.259±0.001
Australorp Black Speckled	ABS	20	0.018±0.007	0.349±0.005	0.205±0.001
Sussex Light (subpopulation 2)	SL2	5	0.007±0.007	0.410±0.116	0.464±0.0003
Russian Crested	RC	20	0.022±0.006	0.359±0.003	0.160±0.0003
Pantsirevka Black	PB	17	0.012±0.005	0.371±0.005	0.185±0.0004
Zagorsk Salmon	ZS	18	0.018±0.0115	0.343±0.006	0.263±0.0005
Czech Golden	CG	16	0.113±0.030	0.306±0.013	0.252±0.0005
Faverolles Salmon	FS	20	0.062±0.014	0.321±0.006	0.239±0.0005
New Hampshire	NH	19	0.007±0.004	0.366±0.003	0.156±0.0003
Cochin Bantam	CBm	20	0.162±0.0216	0.262±0.008	0.280±0.0006
Poland White-crested Black	PWB	18	0.1166±0.0295	0.314±0.018	0.214±0.0005
Ukrainian Muffed	UM	18	0.0180±0.0093	0.365±0.005	0.154±0.0003
Aurora Blue	AB	20	0.0097±0.0038	0.357±0.003	0.188±0.0004
Orloff Mille Fleur	OMF	20	0.0147±0.0107	0.355±0.006	0.195±0.0004
Leningrad Golden-and-grey	LGG	20	0.0198±0.0083	0.348±0.006	0.246±0.0005
Bantam Mille Fleur	BMF	20	0.0284±0.0119	0.3265±0.009	0.374±0.0008
Hamburg Silver Spangled Dwarf	HSSD	20	0.0131±0.0063	0.328±0.0032	0.438±0.0010
Leghorn Light Brown	LLB	19	0.0140±0.0063	0.335±0.0073	0.2882±0.0006
Silkie White	CW	19	0.0245±0.0110	0.366±0.0103	0.289±0.0006
Frizzle	F	17	0.0038±0.0032	0.376±0.0031	0.171±0.0003
Pervomai (subpopulation 1)	Pm1	20	0.0155±0.0096	0.359±0.0056	0.209±0.0004
Rhode Island Red	RIR	24	0.019±0.0068	0.356±0.0036	0.158±0.0003
Pervomai (subpopulation 2)	Pm2	5	0.0048±0.0048	0.379±0.007	0.310±0.0006

Table S2-2. Calculation results of inbreeding (F_{IS}), heterozygosity (H_o), LD and F_{ROH} levels in crossbred populations

Cross	*n*	F_{IS}	H_o	LD	F_{ROH}
BL×SL	16	0	0.407±0.0042	0.215±0.001	0.051±0.008
TS×SL	14	0	0.406±0.0035	0.215±0.001	0.040±0.009
UG×Ar	14	0	0.405±0.0034	0.206±0.001	0.039±0.008
SL×Ar	14	0	0.401±0.0035	0.188±0.001	0.033±0.008
WC×(SL×Ar)	14	0	0.429±0.0031	0.193±0.001	0.033±0.008
WC (BL×SL)	14	0	0.421±0.0035	0.206±0.001	0.024±0.008
TS (SL×Ar)	14	0	0.414±0.0153	0.195±0.001	0.046±0.008

(a)



(b)

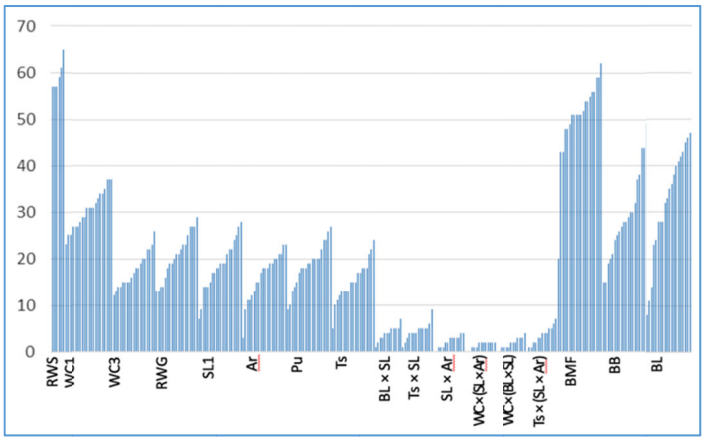
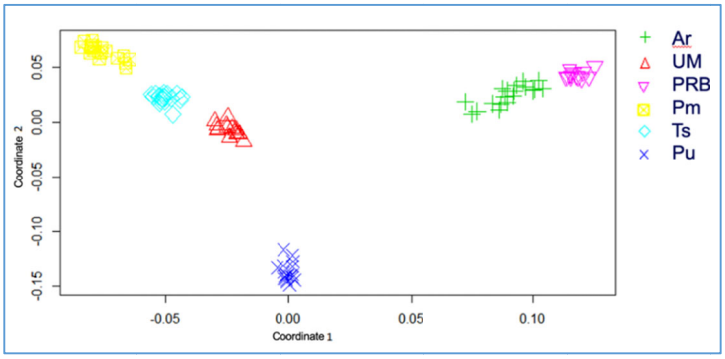


Fig. S2-2. Indicators of genetic diversity in several studied populations (subpopulations) of chickens, both purebred and crossbred. (**a**) F_{ROH}, LD and F_{IS}; (**b**) number of ROH. Breeds (subpopulations) and crossbreds: Amrock (Ar), Bantam Mille Fleur (BMF), Brahma Buff (BB), Brahma Light (BL), Pushkin (Pu), Russian White (including two subpopulations RWS and RWG), Sussex Light (subpopulation SL1), Tsarskoye Selo (Ts), White Cornish (WC, two subpopulations WC1 and WC3), BL×SL, SL×Ar, TS×SL, TS (SL×Ar), WC (BL×SL), and WC×(SL×Ar).

In addition, genetic identification of several individual breeds (subpopulations) and crossbreds was executed using multivariate scaling (MDS) as shown in Fig. S2-3.

(a)



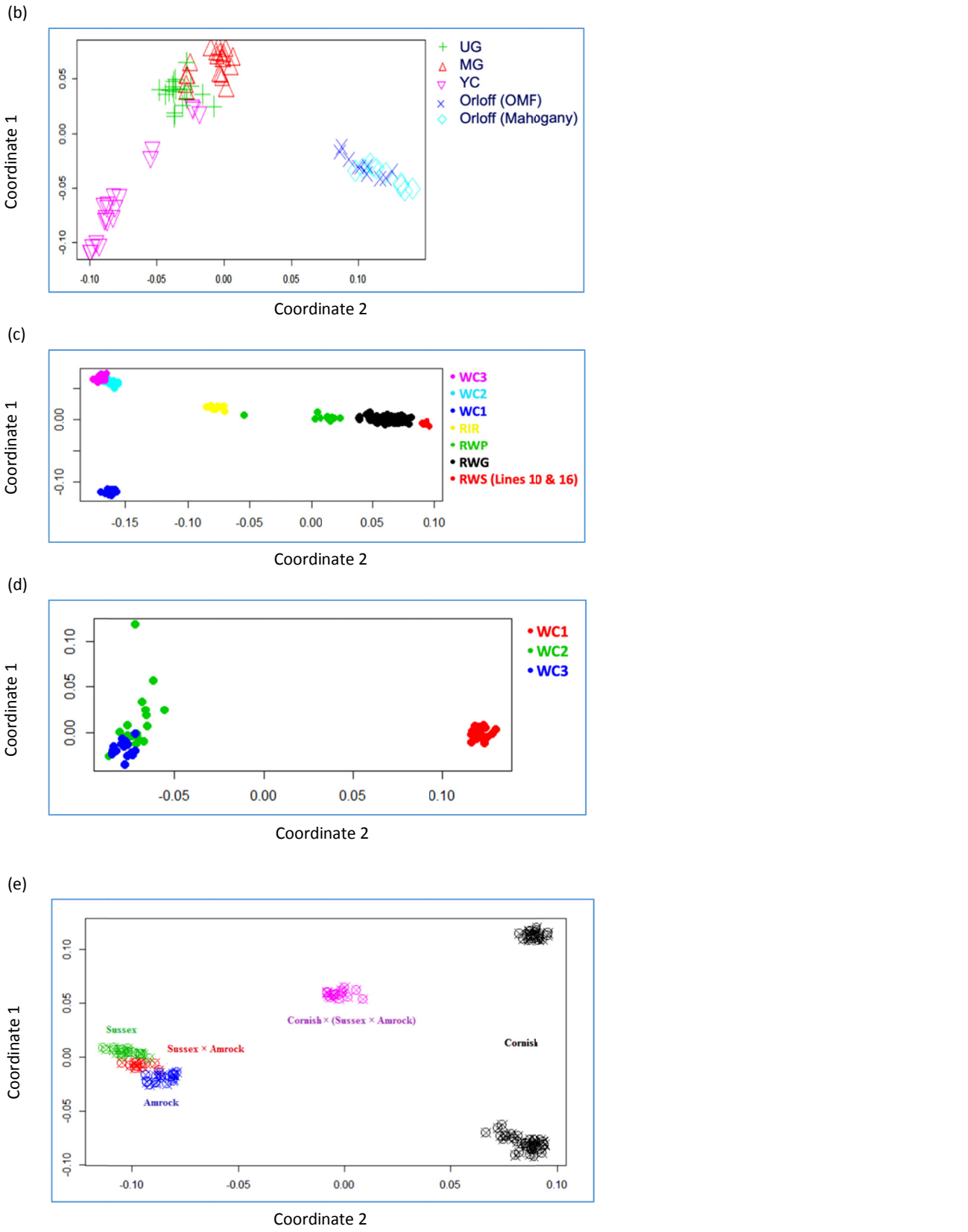


Fig. S2-3. Genetic identification by multivariate scaling (MDS): (**a**) dual purpose breeds; (**b**) game and related breeds; (**c**) meat-type, egg-type and dual purpose breeds and subpopulations; (**d**) WC subpopulations; and (**e**) hybrid populations and their parental breeds. Breeds (subpopulations): Amrock (Ar), Moscow Game (MG), Orloff (including Mille Fleur, OMF, and Mahogany subpopulations), Pervomai (subpopulation Pm1), Plymouth Rock Barred (PRB), Pushkin (Pu), Rhode Island Red (RIR), Russian White (RWS, RWP and RWG), Sussex Light (SL), Tsarskoye Selo (Ts), Ukrainian Muffed (UM), Uzbek Game (UG), White Cornish (WC, three subpopulations WC1, WC2 and WC3), and Yurlov Crower (YC). Crossbreds: SL×Ar, and WC×(SL×Ar).

A comparative analysis of the variability of the crossbreds showed that it depended on the homozygosity and the crossing combination of the original breeds. F_{ST} values calculated between breeds and interbreed crossbreds demonstrated that the best three-way combination, with the maximum genetic distance in F_1 and F_2, is WC×(BL×SL) (Table S2-3). The meat productivity of crossbreds during the anatomical processing of carcasses suggested the superiority of the same cross.

Table S2-3. Results of F_{ST} analysis of breeds and hybrids of the proposed cross

Compared populations (F_1)	**F_{ST}**	**Compared populations (F_2)**	**F_{ST}**
Tsarskoye Selo (Ts)		Ts×SL	
	0.108±0.0024		0.209±0.0055
Sussex Light (SL)		White Cornish (WC)	
SL		SL×Ar	
	0.111±0.0024		0.191±0.0054
Amrock (Ar)		WC	
SL		SL×BL	
	0.190±0.0035		0.220±0.0058
Brahma Light (BL)		WC	
Uzbek Game (UG)		UG×Ar	
	0.089±0.0022		0.202±0.0057
Ar		WC	

Thus, with the obvious superiority of individual crossbreds in terms of meat productivity, their genotyping also made it possible to calculate the level of inbreeding, linkage disequilibrium, the number and length of ROH, which is necessary for the analysis of genetic diversity both within the original breeds and in the genomes of the resulting crossbred animals and crossing combinations. The obtained intra- and interbreed data characterized the genetic diversity in combination with an adequate assessment of the state of small populations. It was found that the variability of the crossbred birds was dependent on the parent breeds. It is necessary and planned to continue research on the study of interbreed crossbred at a deeper level.

Additionally, the effective size of gene pool populations was examined to track their demographic history and changes in the genetic structure over time (Table S2-4; Fig. S2-4).

Table S2-4. Effective size of some populations

Breed (subpopulation)	**Code**	**Effective population size (N_e)**							
		Generations ago							
		25	**50**	**100**	**200**	**400**	**800**	**1600**	**3200**
Russian White (subpopulation 2)	RWS	14	29	54	94	175	312	540	962
Brahma Light	BL	29	46	69	143	319	992	1726	
Sussex Light (subpopulation 2)	SL2	15	29	58	117	234	467	934	1868
Bantam Mille Fleur	BMF	21	42	84	168	331	678	1729	
Pervomai (subpopulation 1)	Pm1	28	55	110	221	441	882	1765	3530
White Cornish (subpopulation 3)	WC3	70	121	197	358	610	1009	1790	3309
Brahma Light	BL	32	53	81	181	342	711	1812	2148
Pervomai (subpopulation 2)	Pm2	52	80	131	281	415	1073	1827	5954
Silkie White	SW	31	62	123	246	492	962	1919	3850
Russian White (subpopulation 3)	RWP	75	109	198	365	646	1100	1920	3406
Cochin Bantam	CBm	33	66	132	263	526	1053	2106	4211
Zagorsk Salmon	ZS	36	71	142	285	569	1139	2278	4555
Amrock	Ar	48	69	108	227	386	1189	2316	23497
Czech Golden	CG	38	76	152	304	608	1215	2430	4860
Leningrad Golden-and-grey	LGG	39	78	156	312	624	1248	2496	4992
Rhode Island Red	RIR	74	152	273	491	863	1479	2500	4484
Pavlov White	PW	36	72	143	288	576	1189	2525	4701
Faverolles Salmon	FS	41	81	162	324	648	1296	2592	5185
Red White-tailed Dwarf	RWD	41	81	162	324	648	1293	2617	5166

Russian White (subpopulation 1)	RWG	124	183	320	567	974	1628	2770	4949
Naked Neck	NN	48	96	191	383	766	1540	3088	6082
Pavlov Spangled	PS	44	88	175	351	702	1431	3127	5666
Australorp Black	AoB	49	98	194	389	778	1564	3143	6255
Australorp Black Speckled	ABS	50	99	198	397	793	1587	3173	6347
Leningrad Mille Fleur	LMF	54	108	216	433	865	1730	3461	6921
Minorca Black	MB	55	110	220	440	882	1754	3516	7022
Aurora Blue	AB	55	111	221	442	885	1770	3539	7078
Pantsirevka Black	PB	56	112	223	447	894	1787	3574	7149
Plymouth Rock Barred	PRB	40	62	106	192	395	1303	3625	17351
Poltava Clay	PC	65	130	260	520	1041	2104	4240	8385
Russian Crested	RC	67	133	266	532	1066	2132	4264	8527
New Hampshire	NH	68	136	272	543	1087	2174	4347	8778

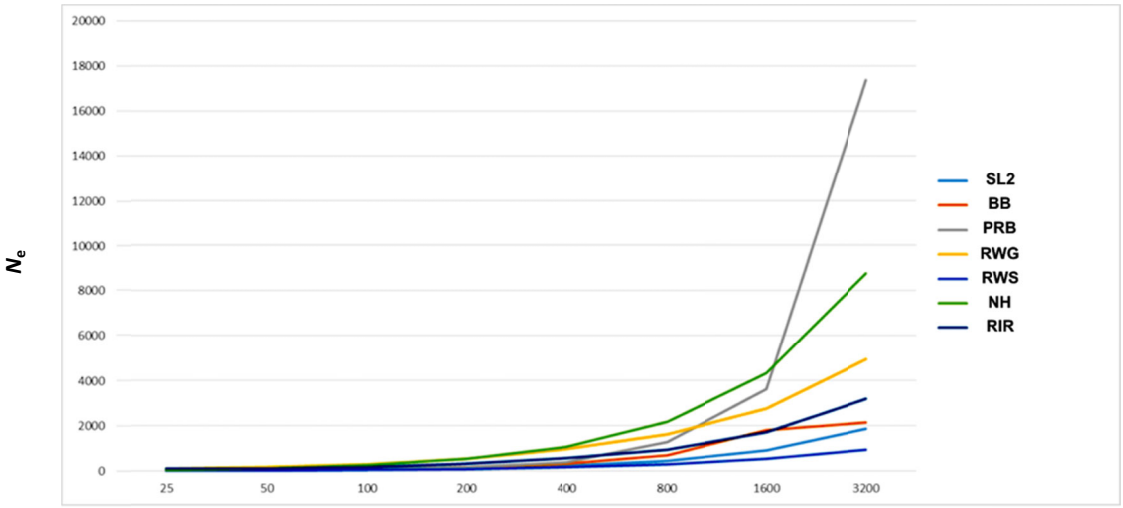
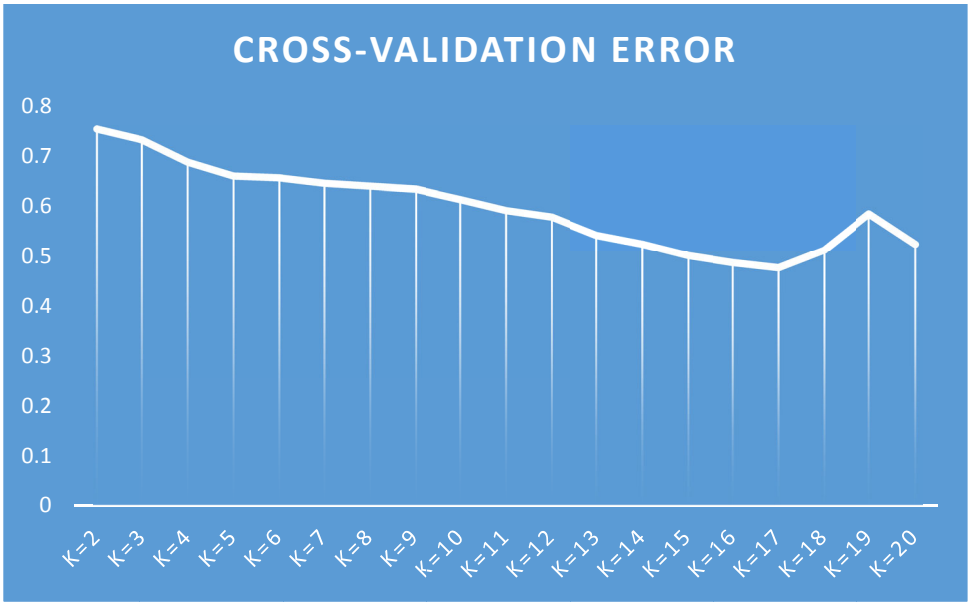


Fig. S2-4. Effective population size (N_e) in some breeds and subpopulations: Sussex Light (subpopulation SL2), Brahma Buff (BB), Plymouth Rock Barred (PRB), Russian White (subpopulations RWG and RWS), New Hampshire (NH), and Rhode Island Red (RIR).

In the global ancestry analysis (Alexander et al., 2009) using the ADMIXTURE 1.23 program (Fig. S2-5b), their optimal gradation was found in the studied breeds when the total set of breeds was divided into 17 ancestral groups (K=17; Fig. S2-5a). Although we did not specifically aim to prove that all gene pool populations at our disposal were either purebred or crossbred, the main objective in any such study is to reflect the diverse ancestry, genetic divergence and demographic history of populations. The latter may involve crossings (purposeful or random) with other breeds, as was confirmed and detailed here by studying the ADMIXTURE-inferred landscape and breed phylogeny (Fig. 3 in main body).

(a)



(b)

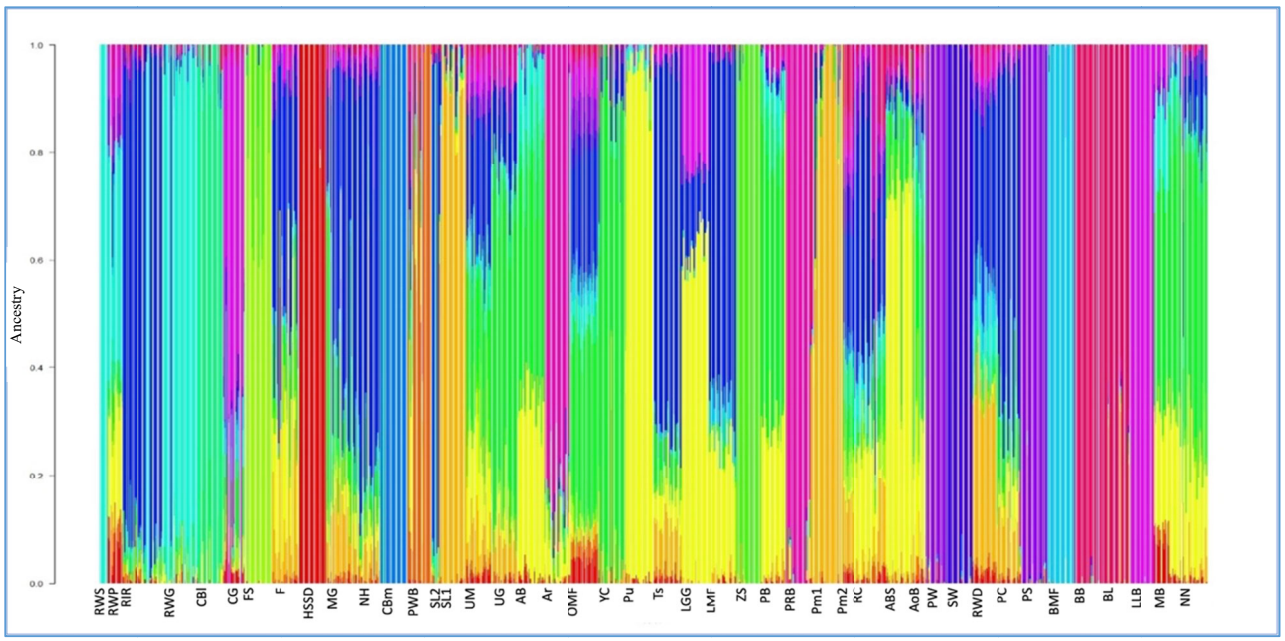


Fig. S2-5. The rugged (and ragged) landscape of admixture, with the optimal number of ancestral populations being *K*=17 (**a**), as shown using the respective bar plots for the gene pool breeds/subpopulations studied (**b**): Amrock (Ar), Aurora Blue (AB), Australorp Black (AoB), Australorp Black Speckled (ABS), Bantam Mille Fleur (BMF), Brahma Buff (BB), Brahma Light (BL), Cochin Bantam (CBm), Cochin Blue (CBl), Czech Golden (CG), Faverolles Salmon (FS), Frizzle (F), Hamburg Silver Spangled Dwarf (HSSD), Leghorn Light Brown (LLB), Leningrad Golden-and-gray (LGG), Leningrad Mille Fleur (LMF), Minorca Black (MB), Moscow Game (MG), Naked Neck (NN), New Hampshire (NH), Orloff Mille Fleur (OMF), Pantsirevka Black (PB), Pavlov Spangled (PS), Pavlov White (PW), Pervomai (two subpopulations Pm1 and Pm2), Plymouth Rock Barred (PRB), Poland White-crested Black (PWB), Poltava Clay (PC), Pushkin (Pu), Red White-tailed Dwarf (RWD), Rhode Island Red (RIR), Russian Crested (RC), Russian White (three subpopulations RWS, RWP and RWG), Silkie White (SW), Sussex Light (two subpopulations SL1 and SL2), Tsarskoye Selo (Ts), Ukrainian Muffed (UM), Uzbek Game (UG), Yurlov Crower (YC), and Zagorsk Salmon (ZS).

Visualization of phylogenetic relationships between all 49 chicken breeds and (sub)populations included in this study that was based on SNP genotypes was conducted in the form of a final phylogenetic tree shown in Fig. 3 (in main body). The respective Newick tree format can be represented as follows:

(RWG:-0.01605,(RWP:0.04529,(HSSD:0.23908,((PWB:0.08575,(PW:0.02320,PS:0.01380):0.06575):0.00704,((LGG:0.07929,(CG:0.03415,LLB:0.08285):0.05871):0.01111,((Pu:0.06684,ABS:0.04516):0.01521,(BMF:0.20088,((UM:0.01429,OM:0.06071):0.00834,((((((NH:0.00183,RIR:0.02417):0.01051,((PC:0.00387,MG:

-0.00287):0.00598,Ts:0.03152):0.00374):0.00532,F:0.01881):0.00133,(RC:0.01562,LMF:0.03737):0.00961):0.00354,(RWD:0.08460,(SL1:0.02473,Pm1:0.03127):0.03790):0.00559):0.00278,(((((WC2:-0.00104,WC3:0.01204):0.05080,WC1:0.10070):0.04489,(((BB:0.07638,BL:0.07562):0.05719,(CBl:0.06304,(SW:0.13807,(CBm1:0.01684,(CBm3:0.13830,CBm2:0.05170):0.03716):0.07843):0.01246):0.00216):0.00868,(SL2:0.12864,Pm2:0.09336):0.00713):0.00251):0.00686,(ZS:0.09179,FS:0.08721):0.01794):0.00279,(((NN:0.05329,(((MB:0.02609,AB:0.02591):0.00692,PB:0.02108):0.01134,AoB:0.00466):0.00571):0.00097,(UG:0.01765,YC:0.04335):0.01399):0.00282,(Ar:0.00547,PRB:0.02953):0.04871):0.00295):0.00079):0.00175):0.00537):0.00308):0.00254):0.00548):0.01271):0.0112):0.03571,RWS:0.14905);

where the breeds (subpopulations) were designated as follows: Amrock (Ar), Aurora Blue (AB), Australorp Black (AoB), Australorp Black Speckled (ABS), Bantam Mille Fleur (BMF), Brahma Buff (BB), Brahma Light (BL), Cochin Bantam (three subpopulations CBm1, CBm2 and CBm3), Cochin Blue (CBl), Czech Golden (CG), Faverolles Salmon (FS), Frizzle (F), Hamburg Silver Spangled Dwarf (HSSD), Leghorn Light Brown (LLB), Leningrad Golden-and-grey (LGG), Leningrad Mille Fleur (LMF), Minorca Black (MB), Moscow Game (MG), Naked Neck (NN), New Hampshire (NH), Orloff Mille Fleur (OMF), Pantsirevka Black (PB), Pavlov Spangled (PS), Pavlov White (PW), Pervomai (two subpopulations Pm1 and Pm2), Plymouth Rock Barred (PRB), Poland White-crested Black (PWB), Poltava Clay (PC), Pushkin (Pu), Red White-tailed Dwarf (RWD), Rhode Island Red (RIR), Russian Crested (RC), Russian White (three subpopulations RWS, RWP and RWG), Silkie White (SW), Sussex Light (two subpopulations SL1 and SL2), Tsarskoye Selo (Ts), Ukrainian Muffed (UM), Uzbek Game (UG), White Cornish (three subpopulations WC1, WC2 and WC3), Yurlov Crower (YC), and Zagorsk Salmon (ZS).

The analysis of the generated schematic phylogenetic tree for all 49 surveyed populations (Fig. 4 in main body), however, demonstrated that since this was a rootless tree (i.e., we did not have an outgroup), the PHYLIP program assigned RWS by default as an "imaginary" outgroup, resulting in some tree distortion specifically for populations of this breed. At the same time, we also drew attention to the fact that in the work by Moiseyeva et al. (2003), on one of the trees built on the basis of 24 discrete morphological characters (or 48 phenetic traits), one can clearly see that two typical American dual purpose breeds, RIR and NH, formed an independent, basal branch on the kinship cladogram.

The applied methodological approaches of whole genome genotyping in chicken gene pool populations using SNP chip technology allowed us to identify genetic variants quickly and accurately in tens of thousands of SNP loci in our study. The scientific significance of such studies lies in the identification of valuable markers and genes, their variations and combinations (still unclaimed), which can be in demand in the process of breeding commercial lines, breeds and crosses of poultry.

Thus, within the framework of this research, we addressed the problem of assessing the genetic potential of the global and native gene pool of chicken breeds on the basis of a genome-wide analysis of their genetic variability and divergence using SNP markers. Their effective population number was also determined, and data on the genetic structure and diversity of the studied breeds were obtained, including the assessment of heterozygosity, inbreeding, linkage disequilibrium, phylogenetic relationship, and subdivision of populations.

During this project, by implementing SNP chip technology and other effective methodological approaches in the analysis of the genotyping results, reference groups were formed for the following gene pool breeds: RWG, OMF, UG, YC, Pm, Pu, ABS, PS, Ts, LGG, LFM, ZS, HSSD, PW, RWD, PRB, LLB, CG, BL, BB, CBl, Ar, CW, PWB, CBm, BMF, and F.

Investigation based on advanced genome-wide SNP analysis technology is of great importance for developing ways to effectively maintain, breed and improve chicken gene pool populations, as well as their inclusion in genomic selection programs (Fig. S1-4 in Data S1). In this regard, the study we reported here is of undoubted importance for both fundamental and applied science, since it was aimed at in-depth study of the landscape of the global and domestic gene pool of chicken breeds using modern genomic methods. The conducted research will help in creating the basis for the implementation of genomic and post-genomic methodology in the field of directed poultry breeding. The development of an innovative technology for genomic selection in poultry industry using SNP markers will significantly expand the possibilities for maintaining and creating highly productive lines, breeds and crosses and will significantly increase the economic efficiency of using the genetic potential of chicken gene pool breeds.

Conclusions

The gene pool breeds are carriers of peculiar phenotypic features, such as adaptability to local conditions, resistance to certain diseases, unique productive, aesthetic (ornamental) and other traits. Understanding the fine genetic structure is important in panmictic breeding and tracking historical changes in the molecular organization of the genome in a gene pool population of limited size. Based on the results obtained, a program for the conservation and effective breeding of small populations will be developed. Highly effective SNP loci can be used to solve various breeding problems, and the chicken SNP genotyping technique is instrumental in the schemes for using the developed selection criteria to significantly accelerate breeding progress in gene pool populations.

Moreover, our study is one of the few projects in the world in the field of whole genome genotyping involving the widest range of small chicken breeds to assess their genetic diversity using SNP markers. Thus, we demonstrated the relevance and prospects of such studies in the field of application of genomic and postgenomic technologies in poultry farming and in agriculture in general, taking into account the results obtained here. Straightforward applications can be genome-wide association studies (GWAS) and search for selection signatures (e.g., Cha et al., 2021; Dadousis et al., 2021; Liu et al., 2021; Li et al., 2021, 2022; Dou et al., 2022; Wang S et al., 2022; Wang Y et al., 2022; He et al., 2022; Lyu et al., 2023; Tan et al., 2023).

The conducted studies helped determine the strategy for further development of the methodology for the whole genome assessment of small chicken populations. This also leads to possible options for the practical use of the best performance traits of dual purpose gene pool breeds and populations that were proposed to expand the consumer market and create alternative domestic crosses.

General remarks

Phenomes and genomes of most animal species used in modern agriculture are products of the domestication process of their wild progenitors. Due to artificial divergent selection and depending on the purpose of use and keeping conditions, useful variants and combinations of genes involved in the formation and manifestation of performance traits were effectively “concentrated” in domestic breeds. With the development of lands by man, new centers of domestication of animals appeared. Thus, the distribution of domestic birds took place in different parts of the world. Due to their wide distribution range, chickens have become the most numerous domesticated livestock species (Malomane et al., 2021). On the one hand, intensive selection of poultry for certain economically useful traits has greatly contributed to an increase in the genetic potential and increased adaptability to human exploitation and the environment (Zhang et al., 2020). On the other hand, there are concerns that directional selection through inbreeding or genetic drift could lead to the loss of genetic material, which will negatively affect the development of animal husbandry (Rege and Gibson, 2003; Woelders et al., 2006; Malomane et al., 2019). Breeding progress in poultry farming has caused significant changes in recent years. Whereas the earlier focus was mainly on quantitative aspects (e.g., egg production and egg weight in layers and daily weight gain in broilers), production efficiency, product quality, resistance to diseases and adaptability to local conditions are now increasingly important. One possible solution in addressing these issues may be utilization of the gene pool breeds using modern genomic technologies, as was undertaken in this work.

In recent decades, as a result of the progressive development of poultry farming, there has been a gradual displacement of the gene pool of older breeds. This, in turn, entails the loss of valuable genetic data, since genetic diversity is the source of the original genetic material for improving the adaptive mechanisms of resistance to various conditions in commercial lines (Bosse, 2019; Roh et al., 2020). All this leads to an increase in genetic distances between breeds, which in turn affects the level of genetic diversity in traditional and domesticated populations (Restoux et al., 2022).

The domestication of chickens in developed countries has been directed to different traits, which has led to the divergence of originally close groups of birds. As a result of artificial selection for certain phenotypic traits, the process of breed formation took place on different continents and in different countries (Momen et al., 2018). Improving

genotyping methods expands research opportunities in the field of variation assessment and allows a more detailed look at the problem of genetic diversity and selection results at the molecular level (Sánchez-Martín and Keller, 2019). This is a necessary step to successfully predict the selection effect and understand the biological mechanisms of adaptive and other breed traits (Hoban et al., 2022).

Rational management of the breeding process of farm animals, taking into account the obtained data on the genetic relationships of breeds, will not only solve the problems of food security, but also increase the genetic potential of existing breeds, so the study of the architecture of the poultry genome is in demand in many countries. Modern farm bird populations require in-depth study and analysis of DNA sequences to further maintain genetic diversity and the ability to create new and recreate previously lost breeds (Wang et al., 2020).

Genetic assessment of modern breeds and populations of poultry will make it possible to detect previously unknown parts of the genome, thanks to the study of which it is possible to identify a large number of new molecular genetic markers, selection for which will positively affect not only the conservation of genetic diversity, but also be used in poultry industry.

References

Alexander DH, Novembre J, Lange K, 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*, 19:1655–1664. https://doi.org/10.1101/gr.094052.109.

Bosse M, 2019. No "doom" in chicken domestication? *PLoS Genet*, 15:e1008089. https://doi.org/10.1371/journal.pgen.1008089.

Cha J, Choo H, Srikanth K, et al., 2021. Genome-wide association study identifies 12 loci associated with body weight at age 8 weeks in Korean native chickens. *Genes*, 12:1170. https://doi.org/10.3390/genes12081170.

Dadousis C, Somavilla A, Ilska JJ, et al., 2021. A genome-wide association analysis for body weight at 35 days measured on 137,343 broiler chickens. *Genet Sel Evol*, 53:70. https://doi.org/10.1186/s12711-021-00663-w.

Dou D, Shen L, Zhou J, et al., 2022. Genome-wide association studies for growth traits in broilers. *BMC Genom Data*, 23:1. https://doi.org/10.1186/s12863-021-01017-7.

Fedorova ES, Dementieva NV, Shcherbakov YS, et al., 2022. Identification of key candidate genes in runs of homozygosity of the genome of two chicken breeds, associated with cold adaptation. *Biology*, 11:547. https://doi.org/10.3390/biology11040547.

He Y, Shi H, Li Z, et al., 2022. Identification of new genes and genetic variant loci associated with breast muscle development in the mini-Cobb F2 chicken population using a genome-wide association study. *Genes*, 13:2153. https://doi.org/10.3390/genes13112153.

Hoban S, Archer FI, Bertola LD, et al., 2022. Global genetic diversity status and trends: towards a suite of Essential Biodiversity Variables (EBVs) for genetic composition. *Biol Rev Camb Philos Soc*, 97:1511–1538. https://doi.org/10.1111/brv.12852.

International Chicken Genome Sequencing Consortium, 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature*, 432:695–716. https://doi.org/10.1038/nature03154.

Li YD, Liu X, Li ZW, et al., 2021. A combination of genome-wide association study and selection signature analysis dissects the genetic architecture underlying bone traits in chickens. *Animal*, 15:100322. https://doi.org/10.1016/j.animal.2021.100322.

Li YD, Liu X, Bai X, et al., 2022. Genetic parameters estimation and genome-wide association studies for internal organ traits in an F2 chicken population. *J Anim Breed Genet*, 139:434–446. https://doi.org/10.1111/jbg.12674.

Liu J, Zhou J, Li J, et al., 2021. Identification of candidate genes associated with slaughter traits in F2 chicken population using genome-wide association study. *Anim Genet*, 52:532–535. https://doi.org/10.1111/age.13079.

Lyu S, Arends D, Nassar MK, et al., 2023. High-density genotyping reveals candidate genomic regions for chicken body size in breeds of Asian origin. *Poult Sci*, 102:102303. https://doi.org/10.1016/j.psj.2022.102303.

Malomane DK, Simianer H, Weigend A, et al., 2019. The SYNBREED chicken diversity panel: a global resource to assess chicken diversity at high genomic resolution. *BMC Genomics*, 20:345. https://doi.org/10.1186/s12864-019-5727-9.

Malomane DK, Weigend S, Schmitt AO, et al., 2021. Genetic diversity in global chicken breeds in relation to their genetic distances to wild populations. *Genet Sel Evol*, 53:36. https://doi.org/10.1186/s12711-021-00628-z.

Moiseyeva IG, Romanov MN, Nikiforov AA, et al., 2003. Evolutionary relationships of Red Jungle Fowl and chicken breeds. *Genet Sel Evol*, 35:403–423. https://doi.org/10.1186/1297-9686-35-5-403.

Momen M, Ayatollahi Mehrgardi A, Amiri Roudbar M, et al., 2018. Including phenotypic causal networks in genome-wide association studies using mixed effects structural equation models. *Front Genet*, 9:455. https://doi.org/10.3389/fgene.2018.00455.

Rege JE, Gibson JP, 2003. Animal genetic resources and economic development: issues in relation to economic valuation. *Ecol Econ*, 45:319–330. https://doi.org/10.1016/S0921-8009(03)00087-9.

Restoux G, Rognon X, Vieaud A, et al., 2022. Managing genetic diversity in breeding programs of small populations: the case of French local chicken breeds. *Genet Sel Evol*, 54:56. https://doi.org/10.1186/s12711-022-00746-2.

Roh HJ, Kim SC, Cho CY, et al., 2020. Estimating genetic diversity and population structure of 22 chicken breeds in Asia using microsatellite markers. *Asian-Australas J Anim Sci*, 33:1896–1904. https://doi.org/10.5713/ajas.19.0958.

Sánchez-Martín J, Keller B, 2019. Contribution of recent technological advances to future resistance breeding. *Theor Appl Genet*, 132:713–732. https://doi.org/10.1007/s00122-019-03297-1.

Tan X, Liu R, Zhao D, et al., 2023. Large-scale genomic and transcriptomic analyses elucidate the genetic basis of high meat yield in chickens. *J Adv Res*, Mar 5:S2090-1232(23)00067-X. https://doi.org/10.1016/j.jare.2023.02.016.

Wang MS, Thakur M, Peng MS, et al., 2020. 863 genomes reveal the origin and domestication of chicken. *Cell Res*, 30:693–701. https://doi.org/10.1038/s41422-020-0349-y.

Wang S, Wang Y, Li Y, et al., 2022. Genome-wide association study and selective sweep analysis reveal the genetic architecture of body weights in a chicken F_2 resource population. *Front Vet Sci*, 9:875454. https://doi.org/10.3389/fvets.2022.875454.

Wang Y, Liu L, Liu X, et al., 2022. SLC16A7 Promotes triglyceride deposition by de novo lipogenesis in chicken muscle tissue. *Biology*, 11:1547. https://doi.org/10.3390/biology11111547.

Woelders H, Zuidberg CA, Hiemstra SJ, 2006. Animal genetic resources conservation in the Netherlands and Europe: poultry perspective. *Poult Sci*, 85:216–222. https://doi.org/10.1093/ps/85.2.216.

Zhang J, Nie C, Li X, et al., 2020. Genome-wide population genetic analysis of commercial, indigenous, game, and wild chickens using 600K SNP microarray data. *Front Genet*, 11:543294. https://doi.org/10.3389/fgene.2020.543294.