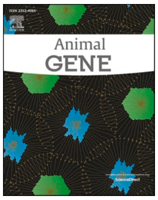
Animal Gene 27 (2023) 200139

Contents lists available at ScienceDirect

Animal Gene

journal homepage: www.sciencedirect.com/journal/animal-gene



Single Nucleotide Polymorphisms in Insulin-like Growth Factor 2 (IGF2) gene and their associations with body weight and growth rate traits in indigenous guinea fowls (*Numida meleagris*) of northern Ghana

Kurukulasuriya Mariesta Jayaroshini Ahiagbe [a], Christopher Adenyo [b], Miho Inoue-Murayama [c], Esinam Nancy Amuzu-Aweh [d], Patrick Bonney [a], Boniface Baboreka Kayang [d,*]

[a] *Animal Research Institute, Council for Scientific and Industrial Research, P.O. Box AH 20, Achimota, Accra, Ghana*
[b] *Livestock and Poultry Research Centre, University of Ghana, Legon, Ghana*
[c] *Wildlife Research Center, Kyoto University, Kyoto, Japan*
[d] *Department of Animal Science, University of Ghana, Legon, Ghana*

ARTICLE INFO

Edited by Dr. M Mousel

Keywords:
Growth rate
Guinea fowl
Insulin-like Growth Factor 2 gene
Northern Ghana
Polymorphism

ABSTRACT

Insulin-like Growth Factor 2 (IGF2) plays important roles in stimulating cell proliferation, differentiation and migration culminating its effects as a modulator of juvenile growth in animals. Although the gene that codes for IGF2 (*IGF2*) has been investigated as a candidate gene in several livestock species, there is no information on polymorphisms in *IGF2* in guinea fowls (*gIGF2*), an important indigenous poultry species from West Africa. Therefore, this study sought to identify the Single Nucleotide Polymorphisms (SNPs) in three populations of indigenous guinea fowls in northern Ghana. Target genomic regions in *gIGF2* were amplified and sequenced from 84 indigenous guinea fowls from Upper East Region ($n = 17$), former Northern Region ($n = 22$) and Upper West Region ($n = 45$) of Ghana together with domesticated French variety ($n = 3$). The sequences were aligned with the reference genomic sequence (domesticated French variety) of chromosome 6 (GenBank accession no., NC_034414.1) to identify SNPs. Statistical associations among the genotypes arising from the SNPs and juvenile growth traits were estimated using linear models. Two novel SNPs were identified in *gIGF2* among the indigenous guinea fowls. An insertion of a Guanine (G) within a poly G motif of the intron following the third exon at the 13,955,730 bp location was identified in the majority (84.5%) of indigenous guinea fowls, while the wild type allele was observed in the minority of indigenous guinea fowls and in the domesticated French variety sampled. Also, a biallelic transition arising from the substitution of G by Adenine (A) at position 13,956,496 bp (13,956,496 G > A) located on the fourth exon, which codes for most of the extension peptide of prepro IGF2, was observed in the minority (11.9%) of indigenous guinea fowls. No significant associations among the genotypes arising from the two SNPs, with body weights and weekly growth rates, were identified.

Abbreviations: A, Adenine; BLAST, Basic Local Alignment Search Tool; BLASTN, Nucleotide BLAST; BW, Body weight; BW1, Body weight at week 1; BW11, Body weight at week 11; BW2, Body weight at week 2; BW3, Body weight at week 3; BW4, Body weight at week 4; BW6, Body weight at week 6; BW7, Body weight at week 7; BW9, Body weight at week 9; C, Cytosine; *cIGF2*, Chicken Insulin-like Growth Factor 2 gene; CP, Crude Protein; DNA, Deoxyribonucleic acid; EDTA, Ethylenediaminetetraacetic acid; FNR, Former Northern Region; G, Guanine; GF, Guinea fowl; *gIGF2*, Guinea fowl Insulin-like Growth Factor 2 gene; GR, Growth rate; GR1, Growth rate from week 1 to week 2; GR2, Growth rate from week 2 to week 3; GR3, Growth rate from week 3 to week 4; GR4, Growth rate from week 4 to week 6; GR5, Growth rate from week 6 to week 7; GR6, Growth rate from week 7 to week 9; GR7, Growth rate from week 9 to week 11; GRO, Overall growth rate; HR, Homologous region; IGF, Insulin-like Growth Factor; IGF2, Insulin-like Growth Factor 2; *IGF2*, Insulin-like Growth Factor 2 gene; MAS, Marker Assisted Selection; ME, Metabolizable Energy; MJ, Megajoules; NCBI, National Center for Biotechnology Information; PCR, Polymerase Chain Reaction; SNP, Single Nucleotide Polymorphism; T, Thymine; TPNG, Three Populations of Northern Ghana; UER, Upper East Region; UTR, Untranslated Region; UWR, Upper West Region.

* Corresponding author.
E-mail address: bbkayang@ug.edu.gh (B.B. Kayang).

https://doi.org/10.1016/j.angen.2022.200139
Received 29 January 2022; Received in revised form 3 November 2022; Accepted 14 November 2022

1. Introduction

Growth is an important quantitative trait in poultry breeding and genetics. Growth in farm animals is influenced by multiple genetic, environmental factors and their interactions (Lawrence and Fowler, 2012). The somatotropic axis plays a central role in regulation of growth and development in animals. Growth hormone secreted by the hypothalamus triggers the secretion of Insulin-like Growth Factors (IGFs) that bind to their receptors in target tissues and initiate signal transduction at cellular level that ultimately results in cell proliferation, differentiation and migration culminating in overall growth of animals (Kim, 2010).

Due to the pivotal roles played by the proteins of the somatotropic axis that include growth hormone, IGFs, their high affinity receptors and binding proteins, the genes that code for them are considered good candidate genes to study growth and development particularly during the early stages of growth in poultry. These genes are also ideal targets for development of molecular markers for Marker Assisted Selection (MAS) for faster growth (Nie et al., 2005; Xu et al., 2013).

Insulin-like Growth Factor 2 (IGF2) is one of the key components of the complex protein network of the somatotropic axis. Therefore, IGF2 gene (*IGF2*) has been considered as a major candidate gene for growth in poultry (Yan et al., 2017). Several workers have studied the polymorphisms within *IGF2* in several livestock species using Single Nucleotide Polymorphisms (SNPs), due to their association with functional regions of genes and related phenotypic traits. Some associations between these SNPs and growth have also been reported.

Zwierzchowski et al. (2010) studied the polymorphisms in *IGF2* and their associations with some economic traits including growth rate in Polish Holstein-Friesian cattle. Associations between SNPs, their haplotypes of *IGF2* with growth have been established in Quinchuan cattle of China (Huang et al., 2014). Similarly, Hou et al. (2010) identified SNPs within porcine *IGF2* in a native Chinese pig breed and established their associations with growth.

Chicken *IGF2 (cIGF2)* is present on the fifth chromosome and is by far the most studied avian counterpart of IGF2 genes. Chicken *IGF2* has been isolated and characterized (Darling and Brickell, 1996) and the nucleotide sequence for the complete *cIGF2* is available (GenBank gene ID, 395097; GenBank accession no., NC_052536; REGION: 13375612.0.13394087). Several studies have reported SNPs within *cIGF2* in several breeds of chicken (Amills et al., 2003; Tang et al., 2010; Wang et al., 2005). Further, Yan et al. (2017) observed associations between juvenile bodyweights and a SNP within *IGF2* in Langshan chicken.

However, neither the IGF2 protein nor the gene in guinea fowl (*Numida meleagris*) (*gIGF2*) has been isolated or characterized. Vignal et al. (2017) recently published the annotated genomic sequence of *gIGF2* as part of the whole genome sequence of guinea fowls. To the best of our knowledge there is currently no information on the SNPs within *gIGF2* gene. This study therefore sought to identify SNPs within *gIGF2* in three indigenous guinea fowl populations from northern Ghana. Further, we carried out preliminary investigations to identify associations of SNPs in *gIGF2* with body weight and growth rate during early growth in birds originating from outbred populations to lay the foundations to evaluate SNP-trait associations for populations with pedigree data in the future.

2. Materials and methods

2.1. Source of experimental animals

Eggs were collected from three main populations of guinea fowls (*Numida meleagris)* raised in Upper East Region (UER), Former Northern Region (FNR) and Upper West Region (UWR) of northern Ghana. In these regions, guinea fowls are widely raised under semi-intensive production system where birds were left to scavenge for most of the day but were provided with supplementary feeding, water, and some form of housing. Hence, there were no phenotypic records kept by the farmers to calculate pedigree of the offspring and populations can best be described as randomly mating outbred populations. Eggs collected across 32 sample locations (Fig. 1) were assembled at a central point and airlifted to the Hatchery unit of the Animal Research Institute, Accra, Ghana where they were hatched in a single batch. Eggs were handled according to standard procedures of handling breeding eggs from point of collection until incubation.

In total 1200 eggs were collected from the Three Populations of Northern Ghana (TPNG). Upon arrival, all the eggs collected from the TPNG were stored for 24 h in a storage room maintained at 18 °C and 70–80% relative humidity. Eggs that were too small, cracked or grossly misshapen were discarded ($n = 351$). The eggs were surface-disinfected and incubated at 37.5 °C and 60% relative humidity for 25 days on separate setter trays in an artificial incubator (Hyderabad Incubators, Hyderabad, India). A total of 509 fertile eggs were transferred to the hatcher on the 25th day and maintained at a temperature of 36.5 °C and a relative humidity of 70%. A total of 280 guinea keets that hatched were individually tagged and moved to the brooder house.

2.2. Management of experimental animals and collection of phenotypic data

Keets hatched from eggs were brooded at the brooder facility of the Guinea Fowl Resource Centre of Animal Research Institute, Accra, Ghana up to the eighth week according to guidelines detailed by Ahiagbe et al. (2016) with artificial light and heat provided by gas brooders while strictly adhering to guidelines to ensure biosafety. Keets (0–8 weeks) were fed with a formulated starter diet containing 24% Crude Protein (CP) and 12.5 MJ Metabolizable Energy (ME)/kg as recommended by Amoah et al. (2018) and Ahiagbe et al. (2021) for local guinea fowls. Growers were then raised in a deep litter house up to 11 weeks on a guinea fowl grower diet containing 16.4% CP and 11.23 MJ Metabolizable Energy/kg. Feed and water were available *ad libitum.* A vaccination schedule was also followed according to Ahiagbe et al. (2016).

Body weights (BW) were recorded at weeks 1 (BW1), 2 (BW2), 3 (BW3), 4 (BW4), 6 (BW6), 7 (BW7), 9 (BW9) and 11(BW11) from individual birds using an electronic balance. Growth rates (GR) from week 1 to 2 (GR1), week 2 to 3 (GR2), week 3 to 4 (GR3), week 4 to 6 (GR4), week 6 to 7 (GR5), week 7 to 9 (GR6), week 9 to 11 (GR7) and overall growth rate (GRO) were determined as weekly weight gains using the formula 1.

$$\text{Growth Rate} = \frac{\text{Final BW–Initial BW}}{\text{Time interval (weeks)}} \tag{1}$$

2.3. Sample collection and sex determination

In the 12th week, 5 ml whole blood was collected aseptically from the wing vein of the birds into Ethylenediaminetetraacetic acid (EDTA) coated tubes and DNA was extracted using DNeasy Blood and Tissue Kit (Qiagen Inc., Valencia, CA, USA). All the 84 remaining birds were used for SNP identification within *gIGF2*. In addition, three DNA samples were obtained from exotic domesticated French variety for comparisons.

Due to low level of accuracy of sex determination based on phenotypic methods or vent sexing in guinea fowls, sexes of all surviving birds were determined by PCR using methods described by Ahiagbe et al. (2017).

2.4. Amplification of IGF2 genomic sequences in guinea fowls by PCR and sequencing

Due to the unavailability of genomic sequence of *gIGF2* in public sequence databases during the laboratory phase of this study, the primers (Table 1) originally designed by Nie et al. (2005) to amplify

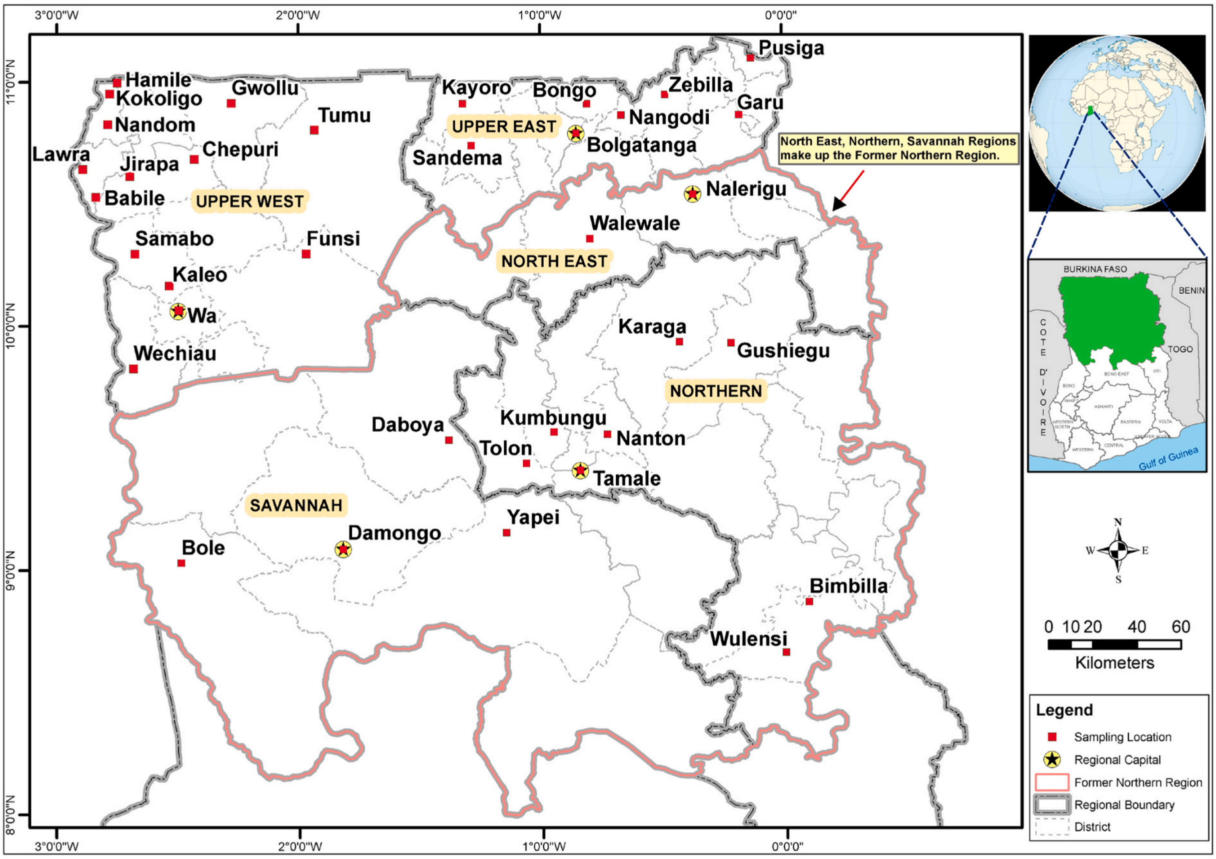


Fig. 1. Map of northern Ghana showing the sampling locations.

Table 1
Primers used for amplification of genomic targets of guinea fowl *IGF2*.

Homologous region in *cIGF2*	Primer name	Primer sequence (5′ to 3′)	Length of amplicon[a] (bp)	Annealing temperature (°C)
5’ UTR, exon 1, intron 1	IGF2E1F	CAGAGATGTGTGCTGCCAGG	339	60
	IGF2E1R	CGAAAGCAGCACTCCTCCA		
Exon 2	902F	GGTAGACCAGTGGGACGAAAT	470	60
	902R	CCTTTGGGCAACATGACATAG		
Intron 2	904F	ATCCCACTCCTATGTCATGTTGC	469	61
	904R	GGGAAGGGAGAACAACACAGTG		
3’ UTR, Exon 3	903F	GGGCGAGCAGCAATGAGTAGAGG	449	68
	903R	CCGGAGCGGCGTGATGGTG		

[a] Amplicon sizes in chicken IGF2 gene.

homologous *cIGF2* sequences were used. An additional primer pair (IGF2E1F and IGF2E1R) was designed based on *cIGF2* genomic sequence (GenBank accession no., AH005039) using Primer 3 (Untergasser et al., 2012) targeting the 5’UTR and exon 1 of *gIGF2*. The PCR reaction contained 20 ng of template DNA, 400 μM of each dNTP, 0.4 μM of each forward and reverse primers, 0.75 U TaKaRa LA-*Taq* DNA Polymerase, 1 x GC buffer I and 1.5 mM Magnesium Chloride (Takara Bio Inc., Shiga, Japan) in a final volume of 15 μl. PCR was performed with initial denaturation at 94 °C for 10 min, followed by 35 cycles of denaturation at 95 °C for 30 s, annealing at specific temperatures within 60–68 °C (Table 1) for 30 s, elongation at 74 °C for 60 s and final extension for 10 min at 74 °C.

Gel electrophoresis was performed to resolve 5 μl aliquots of the amplicons on 1.5% agarose gel in TBE buffer (1 M Tris base, 1 M Boric acid, 0.02 M EDTA) at 100 V for 20 min. Resolved amplicons were stained with gel red and visualized under a UV transilluminator relative to ΦX174 DNA-*Hae* III digest as the DNA size marker (Promega Corporation, Madison, USA).

The PCR products of fragments that amplified successfully were purified using a commercial PCR product purification kit (Roche Diagnostics, Mannheim, Germany) and sequenced in both forward and reverse directions with respective primers using BigDye Terminator version 3.1 cycle sequencing kit (Applied Biosystems, Foster City, CA, USA) and electrophoresed on an ABI PRISM 3130xl sequencer (Applied Biosystems, Foster City, CA, USA).

2.5. Sequence alignment and SNP discovery

Consensus genomic sequences of DNA targets were obtained by aligning forward and reverse sequences for each individual animal genotyped. In total genomic sequences of selected *IGF2* targets of 84 local guinea fowls and three exotic guinea fowls of French origin were obtained. The FASTA formatted consensus sequences from all 87 birds genotyped were aligned using MEGA 7 (Kumar et al., 2016) using the default parameters for SNP identification and genotyping.

Genomic sequences representing the amplified targets of *IGF2* in guinea fowls were submitted to GenBank genomic database (Table 2). Presence of novel SNPs and their locations were determined by performing Nucleotide BLAST with BLASTN version 2.8.0+ of NCBI (National Center for Biotechnology Information, United States National Library of Medicine, National Institutes of Health, Maryland, USA; Zhang et al., 2000) with respect to reference guinea fowl sequence (GenBank accession no. NC_034414.1; Vignal et al., 2017).

2.6. Determination of frequency of SNPs and genotypes

Frequency of SNP variants and genotypes arising from polymorphisms among the three populations from UER, FNR and UWR were calculated using the formulae 2 and 3 implemented in GenAlEx software ver. 6.5 (Peakall and Smouse, 2012).

$$\text{SNP Frequency} = \frac{2(\text{No.of homozygotes}) + \text{No.of heterozygotes}}{2\ (\text{Total No.of birds genotyped})} \tag{2}$$

$$\text{Genotypic frequency} = \frac{\text{No.of individuals with the given genotype AA (or AB, BB)}}{\text{Total No.of birds genotyped}} \tag{3}$$

2.7. Estimation of association between body weight traits, growth rates and individual SNPs

The effects of genotypes at each of the SNP identified in *gIGF2* on the dependent variables including body weights at selected weeks (1, 2, 3, 4, 6, 7, 9, 11), weekly growth rates measured at weekly intervals of 1 and 2, 2 and 3, 3 and 4, 6 and 7, 9 and 11 and the overall growth rate were determined using the linear model $Y_{ijkl} = \mu + P_i + S_j + GTSNP_k + e_{ijkl}$ where, Y_{ijkl} was the dependent variable per each model, μ was the overall mean for a given dependent variable, P_i the effect of i^{th} sample Population, S_j, the effect of j^{th} Sex, $GTSNP_k$, the effect of the k^{th} genotype at a given SNP and e_{ijkl}, the random error. The linear models were implemented using the statistical package R version 0.99.489 (R core team, 2016).

3. Results

3.1. Amplified guinea fowl IGF2 genomic sequences

Out of the four primer pairs originally designed from *cIGF2* genomic sequences, the primer pairs 902 F, 902 R and 903 F, 903 R yielded specific PCR products with guinea fowl DNA (Fig. 2). The remaining two primer pairs did not yield specific PCR products with guinea fowl genomic DNA amidst several optimization trials. The guinea fowl DNA sequences amplified by the primers 902 F and 902 R yielded a 486 bp long genomic sequence that was mapped to exon 3 of the guinea fowl *IGF2* and part of the adjacent intron (Table 2) according to computational annotation of *gIGF2* sequence (GenBank gene ID, 110400777; GenBank accession no., NC_034414.1). Guinea fowl genomic sequence amplified by the primers 903 F and 903 R produced a 450 bp sequence that was mapped to the protein coding segment of the exon 4 of *gIGF2* (Table 2). Relative locations of genomic sequences generated within *gIGF2* established by Nucleotide BLAST are further represented in Fig. 3.

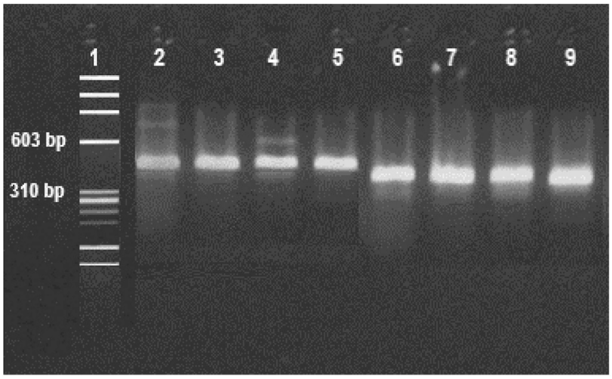


Fig. 2. Appearance of successfully amplified genomic targets of guinea fowl IGF2 gene upon gel electrophoresis. Amplified 902 FR in lanes 2 to 5 (486 bp); amplified 903FR in lanes 6 to 9 (450 bp); ΦX174 DNA-*Hae* III digest as the DNA size marker (lane 1).

Table 2
Locations of guinea fowl IGF2 genomic targets generated and comparison with the homologous chicken sequences.

Primer name	Location on **GF** chromosome 6[a]	Location in the **GF** *IGF2 gene*[b]	Observed amplicon size	Target regions in **GF** *IGF2* [c]	GenBank Accession numbers of submitted genomic sequences	**HRs** in chicken	Location of **HRs** on chicken chromosome 5[d]	% Homology between target **GF** and chicken sequences
902F **902R**	13,955,499 bp to 13,955,984 bp	12,566 bp to 13,051 bp	486 bp	Exon 3 and a portion of intron between exons 3 and 4	LC491601 LC491602	Exon 3 and a portion of intron between exons 3 and 4	13,781,432 bp to 13,781,840 bp	93%
903F **903R**	13,956,345 bp to 13,956,794 bp	13,412 bp to 13,861 bp	450 bp	Exon 4 or third protein coding Exon	LC491599 LC491600	Exon 4	13,782,253 bp to 13,782,702 bp	96%

GF, Guinea Fowl; **HRs**, Homologous Regions; [a] Location on guinea fowl chromosome 6 with reference to GenBank accession no., NC_034414.1; [b]Location within the guinea fowl IGF2 gene with reference to sequence of GenBank gene ID, 110400777; [c] Identification of functional regions of amplicons was according to annotation provided by NCBI for GenBank gene ID, 110400777; [d] Homology of amplicons with chicken chromosome 5 with reference to GenBank accession no. NC_052536 (Region: 13375612.0.13394087).

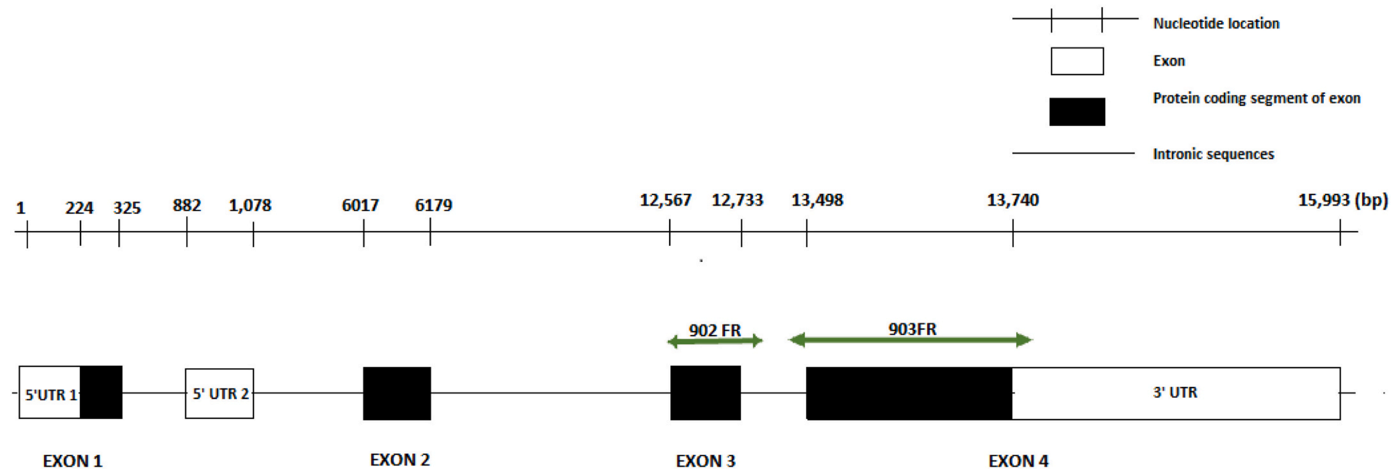


Fig. 3. Schematic representation of locations of genomic sequences generated during the current study within *gIGF2* (*gIGF2* gene map was proposed based on annotation provided by NCBI for GenBank gene ID, 110400777; not drawn to scale of bp length).

Table 3
Summary of Single Nucleotide Polymorphic markers identified within IGF2 gene in guinea fowl.

SNP location on chromosome 6[a] (bp)	**SNP** location in the **GF** IGF2 gene[b] (bp)	Nucleotide in reference sequence	Nucleotide/ Nucleotide sequence in alternative allele	Type of **SNP** variation	Amino Acid change	Genomic region	Description of **SNP**
13,955,730	Followed by 12,797	**T**	**TG**	Indel (Insertion with respect to the reference)	none	Intron between third and fourth exon	13,955,730 T > TG
13,956,496	13,563	**G**	**A**	substitution	none	fourth exon	13,956,496 G > A

GF, Guinea Fowl; **SNP**, Single Nucleotide Polymorphisms; [a] SNP location within chromosome six with reference to GF reference sequence (GenBank accession no. NC_034414); [b] SNP location within **GF** IGF2 gene with reference to GenBank gene ID 110400777.

3.2. *Single Nucleotide Polymorphisms in guinea fowl IGF2*

Two novel Single Nucleotide Polymorphic sites including a substitution and an indel were identified within *gIGF2* of guinea fowls sampled from the three populations of northern Ghana. The location of each SNP with respect to the reference genomic sequence of chromosome six and within *gIGF2* and other salient features of each SNP are summarised in Table 3 according to the convention used in Variant Call Format accepted by NCBI (Danecek et al., 2011). The indel was observed next to a Thymine (T) residue which is located at 13,955,730 bp on chromosome 6. In the reference sequence and also in the wild type allele sequence of sample population (GenBank accession no. LC491601) following the T residue, there is a polyG segment consisting of eight Guanine (G) residues while the alternative allele that was found only in local populations (GenBank accession no. LC491602) contains a nine-residue long polyG segment giving rise to an insertion with respect to the reference sequence (Fig. 4). The substitution of G (GenBank accession no. LC491599) by Adenine (A, GenBank accession no. LC491600) at 13,956,496 bp location of the guinea fowl chromosome six is a biallelic transition.

3.3. *Distribution of single nucleotide polymorphisms*

3.3.1. *Frequency of SNP variants*

Frequencies of allelic variants identified within *gIGF2* among the three guinea fowl populations from UER, FNR and UWR are given in Table 4. The insertion of a G at the 13,955,730 T > TG indel site was present in the majority of birds sampled in all the three populations. The distribution of allelic variants of the indel was significantly different among the three populations ($\chi^2 = 7.28$, $p < 0.05$). The 13,956,496 G > A substitution was only present in the minority (11.9%) of the total population while, the majority reported the allelic variant (G) in the reference genome. Distribution of 13,956,496 G > A substitution was not significantly different among guinea fowls in the three populations of northern Ghana ($\chi^2 = 2.16, p > 0.05$).

3.3.2. *Genotypic frequencies*

The 13,956,496 G > A substitution gave rise to three genotypes **GG, AA,** and **GA** (Table 5) within the populations. Majority of birds carried GG genotype in all the three populations with no significant difference in their distribution across populations ($\chi^2 = 3.64, p > 0.05$). The homozygote genotype for the alternative allele (AA) was only seen in one individual from the Upper West Region.

The indel was distributed in the three genotypes **GG** (homozygote for the alternative allele or insertion reported here), **G0** (heterozygotes carrying the allele in the reference genome, **0** and the insertion, **G**), and **00** (homozygotes for allele present only in the reference sequence). The distribution was not statistically different ($\chi^2 = 8.87, p > 0.05$) but had a higher likelihood ratio of 9.296 ($p \leq 0.05$) for varied distribution. The genotype present in the reference sequence and also in the exotic guinea fowls ($n = 3$) of French origin was only present in the minority of the birds in all the three populations of northern Ghana.

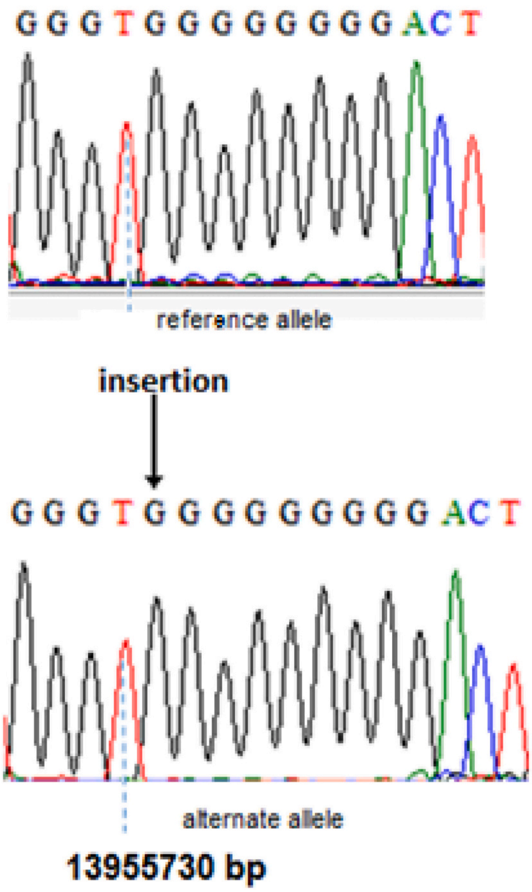


Fig. 4. Appearance of chromatograms in genotypes present in the reference allele and alternative allele (with insertion). Insertion is immediately followed by the T (Thymine) at 13,955,730 bp adding an additional Guanine residue to the eight-residue long poly Guanine in the reference allele according to NCBI Nucleotide BLAST with guinea fowl reference sequence of chromosome 6 (GenBank accession no., NC_034414.1).

Table 4
Allelic frequencies for novel SNPs identified in *IGF2* in the three guinea fowl populations of northern Ghana.

Reference sequence	Description of SNP	SNP variant	UER (n = 17)	FNR (n = 22)	UWR (n = 45)
GenBank accession no., NC_034414.1	13,955,730 T > TG	Null (**0**; allele reported in the reference sequence[1])	0.324	0.273	0.322
		Insertion (**G**; alternative allele with insertion)	0.676	0.727	0.678
GenBank accession no., NC_034414.1	13,956,496 G > A	**A**	0.176	0.136	0.089
		G	0.824	0.864	0.911

[1] Allele/SNP variant in the reference sequence, wild type allele, also observed among the domesticated French breed included in the study (*n* = 3).

3.4. Effect of SNPs on body weight and growth rate traits

None of the novel SNPs identified with *gIGF2* had significant effects on weekly body weights or weekly growth rates during the study period (Table 6).

Table 5
Genotypic frequencies for SNPs in *IGF2* in local guinea fowls from the three populations of northern Ghana.

SNP	Genotype	UER (*n* = 17)	FNR (*n* = 22)	UWR (*n* = 45)
13,955,730 T > TG	**GG**	0.235	0.636	0.600
	G0	0.529	0.182	0.289
	00	0.235	0.182	0.111
13,956,496 G > A	**GG**	0.647	0.864	0.778
	GA	0.353	0.136	0.200
	AA	0.000	0.000	0.022

GG, Homozygotes for the novel insertion observed only among the local populations; **00**, Homozygotes for the wildtype allele found in the reference sequence and the domesticated French birds included in the study; **G0**, Heterozygotes with the novel insertion and the wild type allele.

Table 6
Least Squared Means for Body Weights and Growth Rates for various genotypes at novel polymorphic sites in *gIGF2.*

Trait	13,955,730 T > TG			13,956,496 G > A		
	GG	**G0**	**00**	**GG**	**GA**	**AA**
BW1^{NS}	38.60 ± 0.99	38.80 ± 1.25	37.39 ± 1.65	38.62 ± 0.84	37.92 ± 1.42	37.09 ± 6.02
BW2^{NS}	47.74 ± 1.44	46.37 ± 1.70	46.98 ± 2.37	47.54 ± 1.20	46.12 ± 2.04	46.02 ± 8.64
BW3^{NS}	66.05 ± 2.70	63.22 ± 3.37	63.97 ± 4.44	64.38 ± 2.26	66.25 ± 3.84	57.94 ± 16.24
BW4^{NS}	85.12 ± 4.23	80.56 ± 5.36	80.81 ± 6.93	82.27 ± 3.59	84.90 ± 6.00	79.41 ± 25.39
BW6^{NS}	135.92 ± 8.79	125.01 ± 11.29	129.12 ± 13.87	126.86 ± 7.47	143.32 ± 11.91	126.32 ± 50.41
BW7^{NS}	181.60 ± 11.56	178.30 ± 15.15	172.41 ± 18.45	172.45 ± 10.00	195.28 ± 15.53	157.01 ± 64.06
BW9^{NS}	262.91 ± 17.36	264.93 ± 21.62	246.19 ± 25.52	251.72 ± 14.41	281.07 ± 22.56	244.35 ± 92.72
BW11^{NS}	366.96 ± 22.56	367.08 ± 27.86	350.20 ± 32.89	349.04 ± 18.51	400.48 ± 28.79	330.41 ± 118.33
GR1^{NS}	9.14 ± 1.01	7.57 ± 1.26	9.59 ± 1.66	8.91 ± 0.85	8.19 ± 1.44	8.93 ± 6.11
GR2^{NS}	18.31 ± 1.74	16.85 ± 2.18	16.98 ± 2.86	16.85 ± 1.44	20.13 ± 2.45	11.92 ± 10.38
GR3^{NS}	18.98 ± 1.98	17.09 ± 2.52	16.84 ± 3.26	17.72 ± 1.69	18.66 ± 2.82	21.31 ± 11.93
GR4^{NS}	25.03 ± 2.64	22.45 ± 3.40	24.11 ± 4.17	22.16 ± 2.22	29.17 ± 3.54	23.38 ± 14.99
GR5^{NS}	43.77 ± 3.70	50.44 ± 4.84	39.99 ± 5.90	43.42 ± 3.26	48.92 ± 5.06	31.97 ± 20.86
GR6^{NS}	48.03 ± 3.45	47.68 ± 4.33	37.71 ± 5.11	45.11 ± 3.03	46.53 ± 4.59	45.51 ± 18.41
GR7^{NS}	51.91 ± 3.64	51.08 ± 4.50	52.00 ± 5.31	48.67 ± 2.94	59.70 ± 4.57	43.02 ± 18.79
GRO^{NS}	31.44 ± 2.09	29.39 ± 2.61	30.70 ± 3.44	28.96 ± 1.70	35.83 ± 2.90	29.15 ± 12.26

BWn, Body weight at week **n**, ***n*** = 1,2,3,4,6,7,9,11; **GR**, Growth Rate; **GR1**, **GR** between weeks 1 and 2; **GR2**, **GR** between weeks 2 and 3; **GR3**, **GR** between weeks 3 and 4; **GR4**, **GR** between weeks 4 and 6; **GR5**, **GR** between weeks 6 and 7; **GR6, GR** between weeks 7 and 9, **GR7**, GR between weeks 9 and 11; **GRO,** overall **GR** between weeks 1 and 11; ^{NS} Comparisons among the means of genotypes per SNP were not statistically significant at 95% confidence level.

4. Discussion

The primers used in this study were those originally designed by Nie et al. (2005) to amplify *IGF2* genomic targets in chicken. The genomic regions in guinea fowls corresponding to exon 3 and exon 4 were successfully amplified by 902 F, 902R and 903F, 903R, respectively. Perfect alignment of the nucleotide sequences of these amplicons with the annotated guinea fowl reference sequence of chromosome 6 confirms the identity of the generated genomic sequences and the specificity of primers for chicken and guinea fowl *IGF2*. Therefore, primers 902F, 902R and 903F, 903R can be utilized for studying and genotyping exon 3 and exon 4 in guinea fowls. However, the other primers were not successful in amplifying homologous regions within *gIGF2.* This may be due to complete absence of complementary primer binding sites within the homologous genomic sequences or dissimilarities of homologous sequences. This suggests that these regions exhibit greater sequence variation between the two species compared to sequences spanning exon 3 and exon 4.

The genomic sequences amplified from primer pairs 902F, 902R were mapped to a region spanning exon 3 and beyond including part of the adjacent intron in *IGF2* in both guinea fowl (Fig. 3) and chicken inferring a higher degree of homology of these sequences between the two species. However, the guinea fowl amplicon was 487 bp long compared to the homologous chicken sequence which was 410 bp long. This region covers amino acids coding exon 3 present in both Insulin-like Growth Factor 2 isoform X1 (GenBank accession no., XP_015142011.1) and Insulin-like Growth Factor 2 precursor (GenBank accession no., NP_001025513.1). A BLASTX performed with BLASTX 2.2.29+ interface of UniProt Protein database (The UniProt Consortium, 2017) revealed that this genomic sequence is homologous to chicken sequences coding for amino acid residues 53 to 107 including part of mature protein and 15 residues of E peptide. Although guinea fowl IGF2 protein sequence has not been isolated or sequenced so far, translation of mRNA predicts a similar amino acid sequence.

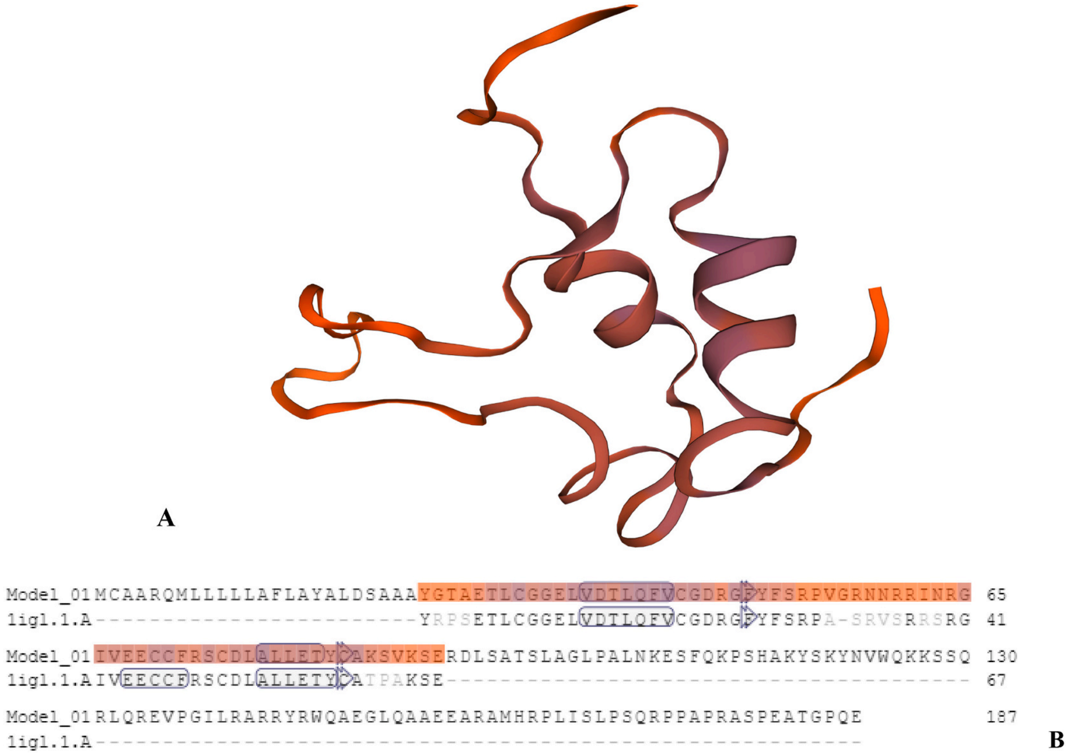
Genomic sequence amplified with primers 903F and 903R represents the amino acid coding fragment of the exon 4 in both chicken (UniProt accession number**,** P33717**)** and guinea fowl. Based on the BLASTX and homology to chicken sequences, it can be predicted that this region codes for the Extension peptide (E peptide). Extension peptides are present at the C-terminus of prepro proteins and get cleaved before release.

To the best of our knowledge, there are no reports of isolated IGF2 protein in guinea fowls or its three-dimensional structure. Therefore, we predicted a three-dimensional structure for guinea fowl IGF2 protein using SWISS-MODEL workspace version 8.05 (Swiss Institute of Bioinformatics, Lausanne, Switzerland) (Waterhouse et al., 2018). We used amino acid sequences predicted for IGF2 isoform X2 (GenBank protein ID., XP_021256640.1), as well as amino acid sequences predicted from genomic sequences generated in our current study and compared with human IGF2 as a template (Fig. 5).

Of the two SNPs found within *gIGF2,* the 13,956,496 G > A substitution could be specifically located on the Extension peptide of the predicted pro *IGF2* protein. The function of E peptides in IGF2 is not clear. However, there is evidence that they play important roles in tethering the propeptides to extracellular matrix, thereby increasing autocrine functions in growth factors in mouse models (Hede et al., 2012). The substitution at 13,956,496 bp was a transition. BLASTX results revealed that this did not result in an amino acid change and hence is a synonymous mutation.

Thus, this substitution is not likely to alter the primary structure, tertiary structure or the biological function. Consolidated report on SNPs in chicken revealed that the most substitutions within exonic regions were transitions (Schmid et al., 2005). Therefore, notwithstanding the ambiguity of function of IGF2 in avian growth regulation, it is unlikely that these can influence the bioavailability and autocrine functions of the protein. Indeed, we found no significant association between the genotypes arising from this mutation with juvenile body weights and growth rates in guinea fowls from randomly mating outbred populations in northern Ghana.

An extensive review of literature revealed no other previously reported SNPs within IGF2 gene in guinea fowl and so the SNPs reported here are novel. However, there were several other SNPs reported within IGF2 genes in chicken. Wang et al. (2005) reported a C > G transition within the second exon within the codon, coding for the 71st amino acid in a cross between a broiler line and Tauhe Silky chicken. Yan et al. (2017) also reported a synonymous mutation in Langshan chicken of China. The two SNPs reported by Amills et al. (2003) including C > T transition within exon 3 and G > A substitution within intron 2 were not



```
Model_01 MCAARQMLLLLLAFLAYALDSAAAYGTAETLCGGELVDTLQFVCGDRGFYFSRPVGRNNRRINRG 65
1igl.1.A ------------------------YRPSETLCGGELVDTLQFVCGDRGFYFSRPA-SRVSRRSRG 41

Model_01 IVEECCFRSCDLALLETYCAKSVKSERDLSATSLAGLPALNKESFQKPSHAKYSKYNVWQKKSSQ 130
1igl.1.A IVEECCFRSCDLALLETYCATPAKSE--------------------------------------- 67

Model_01 RLQREVPGILRARRYRWQAEGLQAAEEARAMHRPLISLPSQRPPAPRASPEATGPQE        187
1igl.1.A ---------------------------------------------------------
```

Fig. 5. Predicted three-dimensional structure of guinea fowl IGF2 protein: (**A**), The three-dimensional structure predicted using SWISS-MODEL workspace version 8.05 based on comparisons with human IGF2 as a template (SWISS-MODEL Template Library ID, 2 l29.1.B) (**B**), Alignment between proposed model guinea fowl IGF2 protein and human IGF2 (SWISS-MODEL Template Library ID, 2 l29.1.B).

homologous to the two SNPs identified during the current study.

Majority of local guinea fowls sampled had the insertion at 13,95,5730 T > TG polymorphic site following 13,955,730 bp (GenBank accession no., NC_034414.1), unlike the allelic variant present in the reference sequence and the exotic birds of French origin. The insertion adds a G residue to a short GC rich segment of the intron (Fig. 4). Although introns were initially described as 'Junk DNA', recent research has unveiled some functional elements within them and their influence on gene expression including regulation of transcription and splicing control (Zhu et al., 2016). Enrichment of short GC rich regions have been implicated in controlling function of the next exon and transcriptional splicing among others (Khuu et al., 2007). However, the genotypes arising from the 13,955,730 T > TG indel had no significant effect on the body weight traits and growth rates during the early growth of guinea fowls. Therefore, insertion of G within this GC rich short sequence is not likely to have any influence in RNA splicing as observed by Khuu et al. (2007) within this context.

None of the two SNPs identified within the *gIGF2* including 13,955,730 T > TG and 13,956,496 G > A located at 13,955,730 bp and 13,956,496 bp on chromosome 6 had any statistical effects on body weight traits and growth rates during early growth in local guinea fowls and, hence, are not likely to be useful as candidate SNPs for early growth in guinea fowls. Nevertheless, the influence of these SNPs on other phenotypes could be a subject for future research. Influence of these genotypes should also be explored with a larger sample size of outbred populations. So far, there have been no dedicated breeding programmes with phenotypic recording schemes to improve local guinea fowls for early growth. Therefore, we could not examine the effects of these novel SNPs with pedigreed populations. We recommend initiation of such programmes to enable exploration of effects of these SNPs in pedigreed populations.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Data availability

Data will be made available on request.

Acknowledgements

Authors acknowledge the World Bank for providing funding support for the field and on station work of this study through the West Africa Agricultural Productivity Programme, Phase 11(A). We are also grateful to the Wildlife Research Center, Kyoto University, for supporting Mrs. K. M. J. Ahiagbe to conduct laboratory analysis of this study. This work was partially financially supported by KAKENHI Grant No. 16H05801 to Professor Miho Inoue-Murayama.

References

Ahiagbe, K.M.J., Amuzu-Aweh, E.N., Bonney, P., Nyameasem, J.K., Avornyo, F.K., Adenyo, C., Amoah, K.O., Naazie, A., Kayang, B.B., 2021. Comparison of early growth and survivability in indigenous guinea fowls from Northern Ghana. Trop. Anim. Health Prod. 53 (89) https://doi.org/10.1007/s11250-020-02510-4.

Ahiagbe, K.M.J., Karbo, N., Avornyo, F., Nyameasem, J.K., Adu-Aboagye, G., Amoah, K. O., Affedzie-Obresi, S., Beckley, C.S.K., Duncan, J.L., 2016. Improving Guinea Fowl Keet Survivability in Ghana, a Manual for Management of Keet Mortalities Using an Integrated Approach. CSIR-Animal Research Institute, Accra, Ghana, pp. 1–28. ISBN, 978–9988–2-41248-5.

Ahiagbe, K.M.J., Kayang, B.B., Naazie, A., Avornyo, F.K., Botchway, P.K., Ali, S.Z., 2017. Non-invasive sex determination of guinea fowl keets (*Numida meleagris*) by Polymerase Chain Reaction. Ghana J. Sci. 57, 3–11. https://doi.org/10.4314/gjas.v52i1.

Amills, M., Jiménez, N., Villalba, D., Tor, M., Molina, E., Cubiló, D., Marcos, C., Francesch, A., Sànches, A., Estany, J., 2003. Identification of three single nucleotide polymorphisms in the chicken insulin-like growth factor 1 and 2 genes and their associations with growth and feeding traits. Poult. Sci. J. 82, 1485–1493. https://doi.org/10.1093/ps/82.10.1485.

Amoah, K.O., Nyameasem, J.K., Asiedu, P., Adu-aboagye, G.A., Wallace, P., Ahiagbe, K. M.J., Rhule, S.W.A., 2018. Protein and energy requirements for indigenous guinea keets (*Numida meleagris*) in southern Ghana. Ghana J. Agr. Sci. 52, 105–111.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., et al., 2011. The variant call format and VCF tools. J. Bioinform. 27 (15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330.

Darling, D.C., Brickell, P.M., 1996. Nucleotide sequence and genomic structure of the chicken Insulin-like Growth Factor-II (IGF-II) coding region. Gen. Comp. Endocrinol. 102, 283–287. https://doi.org/10.1006/gcen.1996.0071.

Hede, M.S., Salimova, E., Piszczek, A., Perlas, E., Winn, N., Nastasi, T., Rosenthal, N., 2012. E-peptides control bioavailability of IGF-1. PLoS One 7 (12), e51152. https://doi.org/10.1371/journal.pone.0051152.

Hou, G., Wang, D., Guan, S., Zeng, H., Huang, X., Ma, Y., 2010. Associated analysis of single nucleotide polymorphisms of IGF2 gene's exon 8 with growth traits in Wuzhishan pig. Mol. Biol. Rep. 37, 497–500. https://doi.org/10.1007/s11033-009-9681-5.

Huang, Y.Z., Zhan, Z.Y., Li, X.Y., Wu, S.R., Sun, Y.J., Xue, J., Lan, X.Y., Lei, C.Z., Zhang, C.L., Jia, Y.T., Cheng, H., 2014. SNP and haplotype analysis reveal IGF2 variants associated with growth traits in Chinese Qinchuan cattle. Mol. Biol. Rep. 41, 591–598. https://doi.org/10.1007/s11033-013-2896-5.

Khuu, P., Sandor, M., DeYoung, J., Ho, P.S., 2007. Phylogenomic analysis of the emergence of GC-rich transcription elements. Proc. Natl. Acad. Sci. U. S. A. 104 (42), 16528–16533. https://doi.org/10.1073/pnas.0707203104.

Kim, J.W., 2010. The endocrine regulation of chicken growth. Asian-Australas. J. Anim. Sci. 23 (12), 1668–1676. https://doi.org/10.5713/ajas.2010.10329.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33 (7), 1870–1874. https://doi.org/10.1093/molbev/msw054.

Lawrence, T.L.J., Fowler, V.R., 2012. Growth of Farm Animals, 2nd ed. CABI Publishing, CAB International Wallingford, Oxfordshire, England.

Nie, Q., Lei, M., Ouyang, J., Zeng, H., Yang, G., Zhang, X., 2005. Identification and characterization of single nucleotide polymorphisms in 12 chicken growth-correlated genes by denaturing high performance liquid chromatography. Genet. Sel. Evol. 37, 339–360. https://doi.org/10.1186/1297-9686-37-4-339.

Peakall, R., Smouse, P.E., 2012. GenALEx 6.5: Genetic Analysis in Excel. Population genetic software for teaching and research - an update. J. Bioinform. 28 (19), 2537–2539. https://doi.org/10.1093/bioinformatics/bts460.

R Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL. https://www.R-project.org/.

Schmid, M., Nanda, I., Hoehn, H., Schartl, M., Haaf, T., Buerstedde, J.M., et al., 2005. Second report on chicken genes and chromosomes. Cytogenet. Genome Res. 109 (4), 415–479. https://doi.org/10.1159/000084205.

Tang, S., Sun, D., Ou, J., Zhang, Y., Xu, G., Zhang, Y., 2010. Evaluation of the IGFs (IGF1 and IGF2) genes as candidates for growth, body measurement, carcass and reproduction traits in Beijing you and Silkie chickens. Anim. Biotechnol. 21 (2), 104–113. https://doi.org/10.1080/10495390903328090.

The UniProt Consortium, 2017. UniProt: the universal protein knowledgebase. Nucleic Acids Res. 45 (D1), D158–D169. https://doi.org/10.1093/nar/gkw1099.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S. G., 2012. Primer3-new capabilities and interfaces. Nucleic Acids Res. 40 (15), 1–12. https://doi.org/10.1093/nar/gks596.

Vignal, A., Thebault, N., Boitard, S., Yapi-gnaore, V., Youssao, I., Tixier-boichard, M., Warren, W., Rognon, X., 2017. Guinea fowl whole genome assembly and application for genetic diversity in African and European populations. In: Proceeding of the Xth European Symposium on Poultry Genetics, Saint Malo, France, pp. 84–88.

Wang, G., Yan, B., Deng, X., Li, C., Hu, X., Li, N., 2005. Insulin-like growth factor 2 as a candidate gene influencing growth and carcass traits and its bialleleic expression in chicken. Sci. China Ser. C.-Life Sci. 48, 187–194. https://doi.org/10.1007/BF02879672.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., Schwede, T., 2018. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 46, 296–303. https://doi.org/10.1093/nar/gky427.

Xu, Z., Nie, Q., Zhang, X., 2013. Overview of genomic insights into chicken growth traits based on genome-wide association study and microRNA regulation. Curr. Genomics 14 (2), 137–146. https://doi.org/10.2174/1389202911314020006.

Yan, L., Fang, X., Liu, Y., Elzo, M.A., Zang, C., Chen, H., 2017. Exploring the genetic variants of insulin-like growth factor II gene and their associations with two production traits in Langshan chicken. J. Appl. Anim. Res. 45 (1), 60–63. https://doi.org/10.1080/09712119.2015.1124328.

Zhang, Z., Schwartz, S., Wagner, L., Miller, W., 2000. A greedy algorithm for aligning DNA sequences. J. Comput. Biol. 7 (1), 203–214. https://doi.org/10.1089/10665270050081478.

Zhu, S., Li, W., Liu, J., et al., 2016. Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR-Cas9 library. Nat. Biotechnol. 34 (12), 1279–1286. https://doi.org/10.1038/nbt.3715.

Zwierzchowski, L., Siadkowska, E., Oprządek, J., Flisikowski, K., Dymnicki, E., 2010. An association of C/T polymorphism in exon 2 of the bovine insulin-like growth factor 2 gene with meat production traits in polish Holstein-Friesian cattle. Czech J. Ani. Sci. 55 (6), 227–233. https://doi.org/10.17221/94/2009-CJAS.