Computers in Biology and Medicine 100 (2018) 17–26

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Highly deleterious variations in *COX1, CYTB, SCG5, FK2, PRL* and *PGF* genes are the potential adaptation of the immigrated African ostrich population

Mohammed Baqur S. Al-Shuhaib[a], Fadhil R. Al-Kafajy[a], Milad Ali Badi[a], Sayed AbdulAzeez[b], Kasi Marimuthu[c], Hussein Ali Imran Al-Juhaishi[d], J. Francis Borgio[b,*]

[a] Department of Animal Production, College of Agriculture, Al-Qasim Green University, Al-Qasim, 51013, Babil, Iraq
[b] Department of Genetic Research, Institute for Research and Medical Consultation (IRMC), Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia
[c] Department of Biotechnology, Faculty of Applied Sciences, AIMST University, Bedong, Kedah Darul Aman, Malaysia
[d] Ojaimi Animal Protectorate, Al-Muhaweel, 51008, Babil, Iraq

ARTICLE INFO

Keywords:
In silico
In vitro
Missense mutations
Ostrich
SNPs

ABSTRACT

Because of variable inconvenient living conditions in some places around the world, it is difficult to collect reliable physiological data for ostriches. Therefore, this study aims to provide a comprehensive *in silico* insight for the nature of polymorphism of important genetic loci that are related to physiological and reproductive traits. Sixty-nine mature ostriches ranging over half of Iraq were screened. Six exonic genetic loci, including cytochrome *c* oxidase I (*COX1*), cytochrome *b* (*CYTB*), secretogranin V (*SCG5*), feather keratin 2-like (*FK2*), prolactin (*PRL*) and placenta growth factor (*PGF*) were genotyped by PCR-single stranded conformation polymorphism (SSCP). Thirty-six novel SNPs, including seventeen nonsynonymous (ns) SNPs, were observed. Several computational software programs were utilized to assess the extent of the nsSNPs on their corresponding proteins structure, function and stability. The results showed several deleterious functional and stability changes in almost all the proteins studied. The total severity of each missense mutation was evaluated and compared with other nsSNPs accumulatively. It is evident from the extensive cumulative *in silico* computation that both p.E34D and p.E60K in *PGF* have the highest deleterious effect. The cumulative predictions from the present study are an impressive guide for the genotypes of African ostriches, which bypassed the expensive protocols for wet laboratory screening, to identify the effects of variants. To the best of our knowledge, this is the first investigation **of its** kind on the **analyses** and prediction outcome of missense mutations in African ostrich populations. The highly deleterious nsSNPs in the placenta growth factor are possible adaptive mutations which might be associated with adaptation in extreme and new environments. The flow and protocol of the computational predictions can be extended for various wild animals to identify the molecular nature of adaptations.

1. Introduction

Ostriches, *Struthio camelus,* are a primitive avian species with variety of ancient characters. From antiquity, ostriches have been providing human beings with food, clothing and adornment; they have also been persecuted for their meat, skin, feathers and eggs. The consumption of meat and other ostrich-derived products has been expanded in many parts of the world [1]. Ostriches bear several unique physiological and developmental characteristics. Because of their adaptation to harsh conditions, African ostriches were able to thrive in multiple Asian, European and American areas. The different patterns of distribution varied from location to location, which might have imposed some altered functional pathways in this species. These differences were studied classically by collecting variable physiological data and comparing them with the original African living breed [2]. Additionally, several genetic studies were performed on ostriches around the world, but they mainly focused on the genetic sexing [3–5], as well as genetic identification [6–8]. Nuclear microsatellites [9,10], mitochondrial genes [11] or rRNA sequences [1] have been extensively used as molecular markers for ostrich populations. Most of these sequences are considered to be selectively neutral and are not suitable for evaluation of genetic polymorphisms for functional traits. Recently, single nucleotide polymorphisms (SNPs) have come to be considered the most common molecular markers in many applications, which are

* Corresponding author.
E-mail addresses: mohammed79@agre.uoqasim.edu.iq (M.B.S. Al-Shuhaib), fadilmarah@yahoo.com (F.R. Al-Kafajy), miladali901@gmail.com (M.A. Badi), asayed@iau.edu.sa (S. AbdulAzeez), aquamuthu2k@gmail.com (K. Marimuthu), baquryhilly@gmail.com (H.A.I. Al-Juhaishi), fbalexander@iau.edu.sa, borgiomicro@gmail.com (J.F. Borgio).

https://doi.org/10.1016/j.compbiomed.2018.06.019
Received 12 May 2018; Received in revised form 13 June 2018; Accepted 22 June 2018

Table 1
The oligonucleotide primer sets designed for the amplification of African ostrich's genes.

Set	Name of the Primer	Gene	Primer sequence (5′ → 3′)	NCBI Accession number of the amplified region	Position in the NCBI reference sequence		Length ~ (bp)
					(Start)	(End)	
1	COX1-F	*COX1*	TGGCATCATCAACCGTCGAA	NC_002785.1	5699	6387	689
	COX1-R		CAGAGCTCATAGGATGGGCG				
2	CYTB-F	*CYTB*	ATCATTCTGGGGCGCTACTG	NC_002785.1	14038	14677	640
	CYTB-R		GCCGATGATGATGAAGGGGT				
3	SCG5-F	*SCG5*	TCACAGGTGACAGGAGGCTA	NW_009270412.1	402862	403088	227
	SCG5-R		GTGGCCCAACCAGGTTCATA				
4	FK2-F	*FK2*	GTAGCCCAAAGACTGCCTGT	NW_009276224.1	635	929	295
	FK2-R		CTACAATGAGCCATGCGTGC				
5	PRL-F	*PRL*	TCCTTCTGGTGTCCAACACG	AB362880.1	1768	2344	577
	PRL-R		TGCAGATCAGCAAGGACAGG				
6	PGF-F	*PGF*	GGGAGAAAGAGCCCATCGAG	NW_009270863.1	218693	219227	535
	PGF-R		CACTGCTACCACGTCCTGAG				

more relevant than neutral markers due to the high prevalence of functional genomic loci. However, conventional SNP analysis techniques yield results that are difficult to interpret without a prior knowledge of phenotypes and the effect of missense mutations on proteins of interest. This problem persisted until the revolutionary breakthrough achievements of bioinformatics tools, in which no such records are necessary, while it requires input data as wildtype and variant amino acids. These freely available *in silico* tools have the advantages of being fast, easily reproducible and less expensive than conventional biological experiments [12–14]. This computational approach has been used successfully for the detection of several mutational consequences in multiple genetic loci in several organisms around the world. Synonymous or silent SNPs do not substitute amino acid sequences, while such mutations may have an effect on gene regulation [15]. Synonymous SNPs are capable of altering protein conformation, expression and function causing variable consequences [16]. Alternatively, a missense mutation or non-synonymous SNP (nsSNP), present within the coding region of a gene, results in the incorporation of an alternative amino acid in the protein composition. Thus, our attention is directed on nsSNPs. Tolerant nsSNPs can cause amino acid changes that are not deleterious, whereas damaging nsSNPs have a significant effect on protein structure, function and its interaction [13,17].

Here, we investigate several genetic loci, namely *COX1*, *CYTB*, *SCG5*, *FK2*, *PRL* and *PGF* that are known to be involved in various physiological and functional activities in the African ostrich population. *COX1, Cytochrome c oxidase subunit I*, encodes a transmembrane protein that is made of 516 amino acids in ostriches. It is the catalytic subunit of the three mitochondrial DNA encoded subunits of the respiratory chain of mitochondrial oxidative phosphorylation [18]. CYTB, the *cytochrome b* of ostriches, consists of 372 amino acids, and it functions as part of the electron respiratory chain, and is the central subunit of several transmembrane cytochrome *b* complexes [19]. Meanwhile, *SCG5* gene encodes secretogranin V (or neuroendocrine protein 7B2) that, in ostriches, consists of 208 amino acids. This acidic protein resides in the secretory granules of neuroendocrine cells and is involved in the secretory pathway in which zymogen is matured and activated [20]. The *FK2* gene encodes a feather keratin 2-like protein of 160 amino acids, which may be involved in the generation of feathers. The distribution and appearance of feathers are determining factors for sex, age and physiological performance of ostriches [21]. The *PRL* gene, encodes a 230 amino acids hormone, which is a pituitary peptide with a broad range of physiological functions [22]. The *PGF* gene encodes a placenta growth factor that is only 138 amino acids in ostrich. It is involved in the regulation of angiogenesis and vast reproductive activities [23]. **We designed specific PCR primers to partially cover the coding portions the investigated genetic loci**. However, to assess the observed SNPs in these loci, it was important to separate deleterious nsSNPs from their tolerant counterparts, to track a possible basis of these nsSNPs and the pattern of physiological performance in this population. Thus, employing exonic regions within candidate genes could provide us with markers for the prediction of the nature of genotypes within the African ostrich. To fulfil this aim, several different algorithm-based tools were implemented for each variant to serve this purpose. The present study provides the first comprehensive description of missense mutation polymorphisms in the African ostrich population.

2. Materials and methods

2.1. Ostriches

The total scanned area for the African ostriches was about 500 km, which included the majority of African ostrich populations in the middle and southern Iraq, where no prior physiological and reproductive records were available for ostrich (Suppl. Fig. 1). A total of 69 (31 male and 38 female) African ostriches were screened in this study, with an approximate age of 4–5 years. A fresh feather quill was plucked from each studied ostrich and stored at 4 °C until processed for DNA extraction.

2.2. DNA extraction

Only one mature and healthy feather was utilized to isolate the gDNA from each studied subject. About 1 cm at the bottom of the quill was cut and minced with a sterile blade. The minced tissue was processed using a universal mammalian/bird DNA extraction method [24]. The gDNA was decontaminated using RNase and its quantity and quality were assessed with a Nanodrop spectrophotometer. Additionally, the gDNA integrity was evaluated by 1% agarose gel electrophoresis.

2.3. Primer design and PCR amplification

Six total PCR fragments were selected for amplification of African ostrich DNA. Two of were mitochondrial DNA positions including *CYTB* and *COX1,* while four of them were in variable nuclear DNA positions including *SCG5, FK2, PRL* and *PGF* loci. Primers were designed using Primer BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast) (Table 1). The primers were purchased from Bioneer (Daejeon, South Korea). The PCR amplification was performed using Bioneer PCR premix. The optimum amplification conditions for the designed amplicons were empirically determined using a gradient PCR thermocycler. The PCR amplification of the designed amplicons was done as per the described conditions in Table 2. PCR amplicons were verified by electrophoresis on 1.5% agarose gel. All PCR amplicons were verified to be distinct and free from non-specific amplicons (Suppl. Fig. 2).