Doctoral thesis

Genetic diversity studies of grasscutter (*Thryonomys swinderianus*) in Ghana by microsatellite and mitochondrial markers

Christopher ADENYO

Wildlife Research Center

Kyoto University

Kyoto, Japan

2014

TABLE OF CONTENTS

ABSTRACT……..... 4

1. GENERAL INTRODUCTION……. 6

2. DEVELOPMENT OF MICROSATELLITE MARKERS FOR GRASSCUTTER (*Thryonomys swinderianus*, RODENTIA) USING NEXT-GENERATION SEQUENCING TECHNOLOGY……. 11

2.1 Introduction……... 11

2.2 Materials and Methods…….. 13

2.3 Results and Discussion…….. 14

3. GENETIC DIVERSITY OF GRASSCUTTER IN GHANA……... 30

3.1 Mitochondrial D-loop Diversity of Grasscutter (*Thryonomys swinderianus, RODENTIA, HYSTRICOMORPHA*) in Ghana……. 30

3.1.1 Introduction…… 30

3.1.2 Materials and Methods……... 31

3.1.3 Results……. 34

3.1.4 Discussion……... 35

3.2 Genetic diversity of grasscutter (*Thryonomys swinderianus*) in Ghana based on microsatellite markers……... 46

3.2.1 Introduction…… 46

3.2.2 Materials and Methods…… 47

3.2.3 Results……..... 50

3.2.4 Discussion……... 51

4. GENERAL DISCUSSION AND FINAL REMARKS……..... 61

5. REFERENCES……. 67

6. APPENDICES…… 83

7. ACKNOWLEDGMENTS…….... 87

8. PUBLICATIONS ASSOCIATED WITH THIS THESIS…… 89

9. LIST OF FIGURES…… 90

10. LIST OF TABLES…… 91

11. LIST OF APPENDICES…….. 92

ABSTRACT

The grasscutter (*Thryonomys swinderianus*) is a wild and fairly large rodent that inhabits sub-Saharan Africa. In Ghana and most parts of Western and Central Africa, the animal is hunted for its meat because it is a delicacy and relatively abundant compared to other mammals used as bushmeat. A number of hunting techniques have been devised including the use of fire and sometimes poisonous baits which are detrimental to the environment, wildlife and consumers of such bushmeat. This necessitated efforts to domesticate the grasscutter over the past few decades. Scientific studies of this species have been conducted mainly on reproduction, growth and nutrition. However, genetic information available on this species is scanty. The overall aim of this study was to develop molecular markers and apply them to determine the genetic structure of grasscutter in Ghana for the promotion of its domestication.

Firstly, I developed 116 novel microsatellite markers that can be used to map the genome of the grasscutter, do parentage analysis and to study genetic diversity and structure of the grasscutter populations in order to answer some ecological questions regarding the species. Secondly, I applied mitochondrial DNA marker (D-loop) and newly-developed microsatellite markers to investigate the genetic diversity of grasscutter populations in Ghana.

A total of 84 samples from three agro-ecological zones (Guinea Savanna, Forest and Coastal Savanna) were compared using mitochondrial D-loop. A total of 26 haplotypes were found forming two haplogroups of grasscutters in Ghana. Haplotype diversity ranged from 0.853 for Forest to 0.978 for Guinea Savanna. Analysis of molecular variance (AMOVA) revealed that 85.79% of the total variation is within populations whilst the remaining variation is between populations. Population differentiation showed that the Forest population is differentiated from both Guinea Savanna and Coastal Savanna populations (F_{ST} = 0.14, p < 0.05). A test of population neutrality indicated that the Guinea Savanna population had undergone a past population expansion event (Fu's F_S = -7.132, p < 0.05) whilst others did not.

Microsatellite analysis showed significant differentiation among all populations (p < 0.05). In terms of diversity, Guinea Savanna recorded the highest mean number of alleles (MN_A), indicating that it is the most diverse population. Both mitochondrial and microsatellite markers revealed that the Guinea Savanna population is more diverse compared to other populations. These results demonstrated that Ghanaian population of grasscutters are highly diverse but less distinctive. This study has shed light for the first time on the genetic structure

of grasscutter in Ghana and has provided molecular markers which will serve as useful tools for genetic improvement of the domesticated grasscutter as well as to study wild grasscutter populations to answer questions regarding phylogeography and to address future conservation concerns.

1. GENERAL INTRODUCTION

The grasscutter (*Thryonomys swinderianus*), also known as the cane rat (or *aulacode* in Francophone West Africa), is a fairly large hystricomorph rodent which mainly inhabits the Sudan and the Guinea savanna (Jori et al., 1995; van der Merwe, 2007). It is an herbivorous species found in grassland areas and wooded savanna and is particularly widely distributed at places where its most preferred grasses for feeding are available (Opara, 2010a). In farming areas, it is often regarded as a major crop pest, destroying cassava, maize, sugarcane, and in plantations feeding on young cocoa, coconut and oil palm (Annor et al., 2009; Opara, 2010a). Wild grasscutters are distributed throughout Ghana, except in rain forests, but they occur in secondary forests and cropped areas leading to continuous expansion of their habitat due to forest clearance for agricultural activities (Annor et al., 2009). Observations showed that grasscutters are nocturnal, gregarious and polygynous (one male with multiple females and their offspring) thereby forming extended family groups (Ewer, 1969; Cox et al., 1988).

Grasscutters breed throughout the year irrespective of seasons, often with peaks around rainy season when there is abundance of feed resources (Asibey, 1974). Adult males weigh 3 – 6 Kg whilst females weigh 2 – 4 Kg with a litter size of up to 15 (Adu et al., 1999). The females reach sexual maturity at about 6 months of age and when successfully bred, gestation lasts for about 152 days (Stier 1991 cited by Adu and Yeboah 2000). The meat of the grasscutter is a delicacy in West and Central Africa and therefore commands a premium price in the region. It was reported that the grasscutter meat is the most expensive compared to other bushmeat and meat from conventional livestock (Asibey and Child 1990). In Ghana, it has been observed that it is one of the most hunted species among various wild animals harvested for food (personal observation). A study done in an urban bushmeat market in the south-western part of Ghana revealed that the grasscutter is the most abundant species traded (Cowlishaw et al., 2005). The grasscutter bushmeat trade alone contributes estimated amount of $57.4 – 59.7 million annually to the Ghanaian economy (Ntiamoa-Baidu, 1998). The high prolificacy, relative ease of keeping under domestic conditions and the immense economic contribution, made the grasscutter an ideal species for domestication. Domestication efforts started in the 1970s in order to increase protein supply in West Africa (Asibey, 1981). Domestication has however been slow due to myriads of challenges that are encountered, including high mortality due to aggression even though farmers have the desire to produce the grasscutter commercially (Adu et al., 1999).

Many studies are being conducted on the grasscutter with respect to its physiology, nutrition, husbandry and economics of production. However, very limited studies are being done on its genetics with the ultimate goal of genetic improvement. In Ghana, studies have been conducted to estimate the heritability of docility and various morphological traits in order to inform selection decisions (Annor et al., 2011a; 2011b). Docility particularly is a very important trait in the grasscutter because it was observed that even after many generations of rearing them in captivity they are still aggressive (Mensah and Okeyo, 2006). This underscores the need for selection to improve temperament to make the grasscutters easier to handle. In Annor et al. (2011a), it was reported that the heritability of docility was high (h^2 = 0.58) which supposes that the trait is highly heritable therefore significant improvement could be made by conscious selection. They also found that docility correlated favourably with litter size therefore selecting for docility could improve litter size or vice versa. In other morphological traits such as body length, head length, tail length, heart girth and height-at-withers, they found high heritabilties (0.52 - 0.60) also indicating that selection for these traits could result in significant improvements. Medium to high genetic correlations (0.45 – 0.85) were found among these traits which means that selecting for any of these traits could favourably improve the other traits. This information could be used in the design of breeding programs to improve the grasscutter under captive conditions to better the lot of farmers although it is inadequate. It is therefore imperative to increase our knowledge on the genetics of the grasscutter especially at the molecular level to complement quantitative genetic information in order to inform breeding decisions. However, there has not been any report on the development of genetic resources such as microsatellite markers in the grasscutter. Furthermore, there is no genetic diversity study of the grasscutter in any of the countries that the animal inhabits.

In the case of mice and rats, many selected lines and inbred as well as outbred lines are available. However, in the case of grasscutter, establishment of reference population begun only recently and there is the need to identify genetically diverse individuals to be used in the establishment of reference population for subsequent linkage analysis which is our future goal. Quantifying genetic variation within and between populations is essential to identify populations that require conservation. We need to study the genetic diversity to be able to understand the role of selection (natural or artificial), the population history and its threat status. The grasscutter is not an endangered species but increasing our knowledge of its

genetic diversity is necessary to inform future conservation decisions since the hunting pressure on this species is very high. For our purpose, identifying genetically diverse populations would enable us to choose individuals for the establishment of reference families. By using mitochondrial and microsatellite markers, more insight could be gained on the diversity and dynamics of the populations. For the specific case of the grasscutter domestication, maintaining genetic diversity is necessary for a number of reasons. Firstly, in order to maintain future breeding options which could make it possible to develop different lines if need be, genetic variation should be maintained. This would make it possible to develop product specific lines just like most livestock species that are available today. Secondly, genetic variation would be needed to obtain greater response to artificial selection under domestic conditions. Traits that are desirable in the domestic setting, such as body weight, litter size and docility will be greatly enhanced after several generations of conscious selection if genetic variation is maintained. Thirdly, it is important that grasscutter populations adapt to uncertain or rapidly changing environmental conditions. Genetic variation would obviously confer such adaptation potential on individuals or populations. Fourthly, animal populations both wild and domestic are often plagued with infectious diseases and parasites which may wipe out whole populations if there is no genetic variation for resistance to those infections. Even though there is no identifiable disease of the grasscutter to date, we need to maintain genetic diversity in order to safeguard the domestic populations. There are three main agro-ecological zones in Ghana that differ in terms of climate and vegetation type. These include Guinea Savanna, Forest and Coastal Savanna. The Guinea Savanna zone is dry in most part of the year with mean annual rainfall of 1000 mm occurring in a single rainy period of between May and September (Akramov and Asante, 2008). The vegetation type is mainly grass with sparse trees. The rainfall pattern in both Forest and Coastal Savanna is bimodal with the major season occurring between March and July and the minor season occurring between September and October. The mean annual rainfall in the Forest zone is 1600 mm whilst that of the Coastal Savanna is 800 mm (Akramov and Asante, 2008). The vegetation of the Coastal Savanna is similar to that of the Guinea Savanna. Grasscutters have adapted to these areas thereby making us to reason that grasscutter populations differ according to the different zones. To be able to decipher these differences, genetic markers such as microsatellites are needed.

Microsatellite markers are a class of short nucleotides (2 – 6 bp) that are repeated in tandem. They are usually ubiquitous in the genome of both eukaryotes and prokaryotes and have been used extensively to study genetic diversity and structure of both wild and domestic populations. They are used to determine parentage and to infer relatedness among kin. They are also used to map traits that are economically significant as in many quantitative trait loci (QTL) analysis both in model and livestock species. Moreover, some qualitative traits such as disease loci and deleterious mutations have also been mapped through such studies. QTL mapping aims at uncovering genetic blueprints that underpin complex traits by identifying specific chromosomal regions that influence the expression of the traits at the phenotypic level (Shalom and Darvasi, 2002). The ultimate goal is to narrow down to genes or regulatory elements that underlie the traits (Shalom and Darvasi, 2002; Flint et al., 2005; DiPetrillo et al., 2005). In rodents such as mouse and rat, thousands of QTLs have been identified for various traits in the past few decades (Flint et al., 2005; Dwinell et al., 2009; Eppig et al., 2012) and more are being mapped due to the advancement in mapping techniques. More importantly, it is known that most genes underlying disease QTL in rodents are often associated with the corresponding disease in humans because some chromosomal regions are conserved in rodents and humans (DiPetrillo et al., 2005). It may take up to 10 years for QTL experiments to fine map and narrow down on the underlying genes (Drinkwater and Gould, 2012) but it is worth the effort. Such experiments involve crossing of clearly defined lines or genetically diverse F0 individuals to obtain F1 individuals which are then intercrossed or backcrossed to produce F2. The traits in question can then be mapped and chromosomal regions identified by genotyping all individuals with several markers (Drinkwater and Gould, 2012). Even though docility of the grasscutter has been characterised using variance component analysis (Annor et al., 2011a), identification of chromosomal segment or genes underlying such phenotype would tremendously enhance the selection process towards more docile animals. To arrive at such stage in the grasscutter, we first need to develop several markers which can be used to genotype individuals from at least three generations. This forms the basis for this study to be carried out.

In this study, I aimed to develop microsatellite markers that can complement the grasscutter improvement efforts and to characterise the genetic diversity of the grasscutter in order to identify unique populations and genotypes and to understand the dynamics of the population in the wild. As far as I know, this study is the very first attempt to develop

molecular markers and to provide first-hand genetic information on the grasscutter populations inhabiting Ghana. This study is organised into two parts. In the first part, I developed novel microsatellite markers in the grasscutter and in the second part I determined the genetic diversity and structure of the grasscutter populations in Ghana using non-coding maternally inherited marker (mitochondrial D-loop) and neutral bi-parentally inherited autosomal microsatellite markers selected from the first part of this study based on their level of polymorphism. In all, I selected 12 highly polymorphic loci from the 116 total markers developed. These markers were polymorphic enough to show any genetic differences among the populations which were classified according to the agro-ecological zones in Ghana. I hypothesised that populations inhabiting different agro-ecological zones are highly diverse and differentiated.

2. DEVELOPMENT OF MICROSATELLITE MARKERS FOR GRASSCUTTER (*Thryonomys swinderianus*, RODENTIA) USING NEXT-GENERATION SEQUENCING TECHNOLOGY

2.1 Introduction

In most parts of the region that the grasscutter inhabits, especially western and central Africa, the meat is a delicacy (Adu and Yeboah, 2000; Addo et al., 2002). In most of these places, the grasscutter is hunted from the wild using crude methods such as fire and poisonous baits, which are obviously detrimental to the habitats and survival of other wild animals, as well as pose imminent danger to the health of consumers of such bush meat (Adu et al., 1999; Asogbadjo et al., 2005). Efforts are therefore being made to domesticate the grasscutter to make the meat readily available to forestall the negative consequences of hunting the species. Successful domestication and commercialization of the grasscutter will make it the first animal to be domesticated in West Africa and the fourth in Africa, in addition to donkey, guinea fowl and ostrich (Mignon-Grasteau et al., 2005). We describe here the development of novel microsatellite markers of the grasscutter, which will be useful in the domestication process as well as to assess and monitor its conservation status.

Microsatellite markers or simple sequence repeats (SSR) are DNA sequences consisting of tandem repeats of up to six nucleotides which are distributed ubiquitously in the genome of organisms, especially eukaryotes (Tautz and Renz, 1984; Chambers and MacAvoy, 2000). They are formed as a result of polymerase slippage of short DNA motifs during replication which causes dissociation and reassociation of misaligned strands (Kruglyak et al., 1998). Some hypotheses have been proposed concerning their functional roles in the genome. They may play a role to facilitate recombination (Pardue et al., 1987; Stallings et al., 1991), involved in packaging and condensing DNA into chromosomes (Stallings et al., 1991; Gailard and Strauss, 1994) or enhance transcription of genes (Hamada et al., 1982, 1984). These hypotheses however could not be confirmed and therefore remain largely speculative (Neff and Gross, 2001). Nevertheless, microsatellite markers are very useful in forensics, mapping of the genome of species, parentage analysis, population genetic studies for population management and conservation as well as molecular typing of microorganisms (Jarne and Lagoda, 1996; Hennequin et al., 2001; Luikart et al., 2003; Lim et al., 2004). They are highly mutable and therefore show high level of intraspecific variation (Buschiazzo and Gemmel, 2006).They continue to be one of the most widely used types of genetic markers

because of their co-dominant nature, abundance in eukaryotic genomes and ease of scoring (Jarne and Lagoda, 1996; Toth et al., 2000; Katti et al., 2001; Selkoe and Toonen, 2006). They are mostly neutral and are present in both coding and non-coding regions in the genome (Toth et al., 2000; Zane et al., 2002). In many cases, microsatellite markers have to be developed *de novo* but cross-species applicability is possible in some closely related taxa even though they may not be very efficient in some instances (Inoue-Murayama et al., 2001).

A number of techniques have been developed over the years to isolate and characterise microsatellites in many species. Traditionally, there are two methods that are widely used; the use of inter-simple sequence repeat primers and enrichment using hybridization probes (Santana et al., 2009). The focus has however shifted recently with the advent of next generation sequencing technology capable of generating several megabases of partial or whole genome data from which reads containing repeats can be selected (Santana et al., 2009; Jennings et al., 2011). This approach will be especially useful for species that have low frequency of microsatellites in their genome (such as birds and some insects) of which development of microsatellite markers have proven difficult. For instance, previous attempts to develop microsatellite markers in guinea fowl have proven futile until recently when the Roche 454 sequencing was used (Kayang et al., 2010; Botchway et al., 2013). The advancement in this technology has also made it possible to sequence multiple species in one run by the use of tags which serve as barcodes that can ensure that samples are separated bioinformatically after sequencing (e.g. multiplex identifier sequences (MIDs) by Roche) (Davey et al., 2011). This is more cost-effective and a faster way of isolating microsatellites compared to the traditional methods which involve laborious and time consuming cloning step.

In order to construct genetic map for species, large number of markers is often needed to cover the entire genome as much as possible. Grasscutter is becoming relevant as a protein resource in Sub-Saharan Africa and will require detailed genetic studies for selection and genetic improvement. It is therefore imperative to develop enough markers to map its genome. The aim of this study was to develop microsatellite markers of grasscutter using next generation sequencing technology.

2.2 Materials and methods

2.2.1 DNA extraction and sequencing

Genomic DNA was extracted from a lung tissue of a semi-domesticated male grasscutter using QIAGEN Blood and Tissue Kit (QIAGEN) and quantified using Nanodrop Spectrophotometer (Thermo Scientific). The quality of the DNA was assessed by electrophoresis on 0.5% agarose gel. To prepare DNA library, 500 ng of the DNA was fragmented by nebulisation at 0.24 Mpa for 1 minute, and purified using the MinElute PCR Purification kit (QIAGEN). The fragments were end-repaired, A-tailed and ligated to the Rapid Library Adapter with RL Ligase (Roche). Short fragments were removed using AMPure XP beads, and the quality and quantity of the library were assessed using Agilent 2100 Bioanalyser (Agilent). Library fragments were mixed with capture beads and clonally amplified through emulsion PCR using the GS Junior Titanium emPCR kit (Roche). Captured fragments were enriched and annealed with sequencing primers and sequenced using GS-Junior bench-top sequencer (Roche), generating a total of 156,966 reads. To scale up the data, two other sequencing techniques were employed. First, the hybridization capture method of Glenn and Schable (2005) was followed until the point of enrichment of the DNA fragments containing repeats, where the fragments were enriched for (CA)n. The resulting enriched fragments were then sequenced directly using the GS Junior bench-top sequencer (Roche). This generated 73,823 reads of which reads containing GT (≥7 repeats) formed 42% of the total reads. Second, the DNA was sequenced with GS-FLX Titanium (Roche) using half (1/2) of the run on an eight lane plate after thorough preparation of the DNA fragments following the manufacturer's protocol. This approach generated 473,677 reads.

2.2.2 Microsatellite selection and screening

The program Msatcommander (Faircloth, 2008) was used to screen all sequences for microsatellites. Repeats including di-, tri-, tetra- and penta- nucleotides were searched for with the following settings: more than seven perfect di-repeats and more than five perfect repeats for the other repeat types. Sequences with 8 repeats or more for all repeat types were selected and a total of 7,702 primer pairs were designed by the Primer 3 software that is embedded in Msatcommander. All primers selected for screening were checked to ensure there were no duplications since some sequence reads appeared more than once or overlapped.

Initial screening of the microsatellites involved PCR with 4 DNA samples obtained from wild grasscutters (two males and two females). PCR was carried out in 10 µl reactions containing 0.75 U of LA-Taq DNA polymerase (TaKaRa), PCR buffer, 400 µM of each dNTP, 0.4 µM of forward and reverse primers, 0.1 µg of T4 Gene 32 Protein (Nippon Gene) and 20 ng of template DNA. After an initial incubation of 95˚C for 2 minutes, PCR amplification was performed for 35 cycles consisting of denaturation at 95˚C for 30 seconds, annealing at 55˚C or 60˚C for 30 seconds, extension at 74˚C for 1 minute and a final extension at 74˚C for 10 minutes. However, for ten of the tri-repeats (Tsw24, Tsw25, Tsw26, Tsw27, Tsw28, Tsw29, Tsw30, Tsw31, Tsw32 and Tsw33), the amplification was performed for 40 cycles using an annealing temperature of 60˚C. Primers that could not amplify all 4 samples with clear bands were discarded. Forward primers were then tagged with FAM, HEX and NED and tested using a semi-domesticated grasscutter family comprising sire, dam and four offspring, to ascertain Mendelian inheritance of the markers. In order to reduce cost of labelling primers, some of the forward primers were tagged with M13 instead of FAM, HEX and NED (Schuelke, 2000). PCR for this set of primers were conducted separately in a two step fashion; initial denaturation at 95˚C for 5 minutes followed by the first step consisting of 30 cycles of 94˚C for 30 seconds, 57˚C for 45 seconds, 72˚C for 45 seconds and the second step of 8 cycles of 94˚C for 30 seconds, 53˚C for 45 seconds, 72˚C for 45 seconds and final extension at 72˚C for 15 minutes. PCR mixture was as described above with addition of 0.4 µM FAM labelled M13 tag and reduction of the concentration of the forward primer to 0.2 µM.

PCR product size was measured using the ABI PRISM 3130xl Genetic Analyzer. Markers that could not amplify or show clear inheritance pattern were removed from the marker set. A final set of 116 markers were used to genotype 16 unrelated individual grasscutters sampled from Mankessim and Jukwa in the Central Region of Ghana. Number of alleles, expected and observed heterozygosities and Hardy-Weinberg Equilibrium were determined using GenAlEx ver 6.41 (Peakall and Smouse, 2006).

2.3 Results and discussion

2.3.1 Sequence characteristics and repeat structure

A total of 704,466 sequence reads were generated from the three approaches with an average read length of 377 bp (Table 2.1). For the two direct sequencing approaches, repeat classes

consisting of 2 - 6 tandem repeats constituted about 4% of the total sequence data. Among the di-repeats, AC/GT was the most frequent whereas GC was the least frequent. AAC/GTT, AAAG/CTTT, AATGG/CCATT, ACAGAG/CTCTGT were most abundant for tri-, tetra-, penta- and hexa-repeats, respectively. Across many vertebrate taxa, including primates and rodents among others, AC was found to be the most frequent di-repeat (Toth et al., 2000). Also, Meglécz et al. (2012) who analysed next-generation sequence data found that AC was the most frequent whilst CG was the rarest among chordates. This is also true for the grasscutter according to our results but different from Japanese giant salamander where AT was found to be the most frequent (Ito et al., 2013). Except for di-repeats, these results are in contrast with Toth et al. (2000) who found AGG, AGAT, AAAAC and ACAGGC to be the most abundant repeat types in rodents compared to other vertebrates. It is likely that the results of Toth et al. (2000) may be biased towards model organisms or influenced by coding sequences. For tri-repeats, AAT was found to be most frequent in vertebrates (Toth et al., 2000; Meglécz et al., 2012) whilst AAC was found to be most frequent in the grasscutter. Toth et al. (2000) found dinucleotides to be most abundant in rodents but in Ito et al. (2013) dinucleotide repeats were more frequent in Grevy's zebra than grasscutter. These differences in the repeat distribution may be due to unique microsatellite composition and distribution in the grasscutter genome. Variations in the evolutionary dynamics of microsatellites have been linked to mechanisms such as DNA repair system, duplication, transposition and recombination events (Martin et al., 2010). A combination of these mechanisms might lead to unique proportions and distributions of microsatellites in the genome of different or even closely related species. It could also be that there was repeat selection bias because our selection criteria was 8 repeats or more for dinucleotides and 5 repeats or more for other repeat types whereas Toth et al. (2000) analysed perfect repeats longer than 12 bp (i.e. at least 6 repeats for dinucleotides) and found that AT instead of AC dinucleotide shows striking dominance among longer repeats (> 24 bp). This obviously suggests that the distribution of microsatellites reported may be influenced by the selection criteria set by the researchers. The frequencies of the repeat types in the data set except for the data from the enrichment approach are presented in Fig 2.1.

2.3.2 Marker Polymorphism

Out of a total of 7,702 primer pairs designed, 404 were tested and 116 were found to be polymorphic (Table 2.2). The number of alleles per locus ranged from 2 to 13 (mean 6.5)

while the observed (H_O) and expected (H_E) heterozygosities ranged from 0.063 to 1.000 (mean 0.575) and 0.063 to 0.919 (mean 0.723), respectively (Table 2.3). These values are higher than those obtained for Laotian rock-rat (Pino et al., 2012) but a little lower than those of kangaroo mouse (Lance et al., 2010) and Florida mouse (Hargrove et al., 2012). Eleven out of the 116 loci deviated significantly from Hardy-Weinberg Equilibrium (HWE) in the population tested ($p < 0.05$) after Bonferroni correction for multiple testing (Rice, 1989). Null allele frequency (NAF) was therefore determined for all loci using FreeNA (Chapuis and Estoup, 2007). Except for loci Tsw24, Tsw33, Tsw48, Tsw51, Tsw91 and Tsw104 NAF was less than 0.2 for all loci, indicating that null alleles may not significantly change the results when these loci are employed in analysis (Latinne et al., 2011). Cumulative probability of identity (PI) for all loci was very low (PI = 6.5×10^{-111} for unrelated individuals and 2.3×10^{-42} when siblings are involved), indicating the usefulness of these markers for individual discrimination. Furthermore, significant Linkage Disequilibrium (LD) was found between several pairs of loci as assessed using GENEPOP ver4.0.10 (Rousset, 2008). Pairs of loci that show tendency of linkage are provided in Appendix A.

Hopefully, these results will be very useful for our future purpose which is to map the genome of the grasscutter. We hope to create the first generation linkage map of the grasscutter which will serve as a reference point to map genes that underlie quantitative and qualitative traits such as body weight, growth rate, litter size and docility. Mapping of genomes of species with microsatellite markers started about two decades ago with mouse (Dietrich et al., 1994) and rat (Jacob et al., 1995) being the first model organisms to be mapped. Mapping approaches were then successfully applied to livestock species such as pig, cattle, sheep, goat and even poultry species such as chicken, turkey and quail. A lot of progress has been achieved in these animals over the years. Just like these livestock and poultry species, there is the need to breed grasscutters to be amenable under domestic conditions and to make them more productive through selective breeding. At present, they are still aggressive and the domestication process is being hindered due to losses. Many grasscutter farmers may get discouraged to continue production because of the losses that they suffer even though there are good market prospects. Adu et al. (1999) advocated for genetic selection in order to overcome this challenge, thereby making this study timely. The microsatellite markers presented here will aid in marker assisted selection which is known to

accelerate the response to selection or genetic progress by increasing accuracy of selection, reducing generation interval and increasing selection differential (van Arendonk et al., 1994)

The grasscutter also has a potential to be used as a model species to study diseases just like other laboratory rodents. In a study by Addo et al. (2007), it was found that experimental infection of grasscutters elicited symptoms that were similar to symptoms developed by humans infected with *Mycobacterium ulcerans* (micro-organism that causes buruli ulcer). The genetic mechanisms of predisposition to such an enigmatic disease could be studied in the future employing genetic markers. The results of our study may therefore become handy in such a study.

Furthermore, wild grasscutters are distributed widely over Sub-Saharan Africa. In West Africa especially, the hunting pressure is very high because of the high demand for the meat. Despite this significance, the population genetic structure, dispersal patterns and genetic diversity are not fully known in the sub-region. Hopefully, the markers obtained in this study can be used to assess and monitor the populations of grasscutters not only in Ghana but on a regional scale to inform population management decisions in the future.

In conclusion, the grasscutter is a Least Concern species in the IUCN Red List and therefore does not currently require any urgent conservation efforts. However, the species has a great economic potential with relevance to agriculture and protein supply in Sub-Saharan Africa. The microsatellite markers presented here will be useful for identification, parentage analysis and analysis of economic traits which will be necessary to enhance the domestication process as well as genetic improvement of this species. They can also be used to study the genetic diversity, genetic structure and dispersal patterns of the grasscutter populations in the wild. This information will be relevant for population monitoring for the purposes of conservation in the near future.

Table 2.1 Fragment read characteristics of the partial genome data of the grasscutter using three sequencing approaches

Technique	No. of reads	Average length (bp)	Repeat type selected	Sequences containing repeats	% of total	No. of primers designed
GS Junior	156,966	419	≥ 7 di, ≥ 5 tri, tetra, penta, hexa	6,806	4.3	2,280
GS Junior CA enrichment	73,823	331	≥ 7 di (GT)	31,100	42.1	268
GS FLX Titanium	473,677	382	≥ 7 di, ≥ 5 tri, tetra, penta, hexa	18,984	4.0	5,154
Total	704,466	377		56,890		7,702

Table 2.2 Polymorphism of different repeat loci from the partial genome data of the grasscutter

Repeat type selected	No. of primers tested	No. polymorphic	% polymorphism
Di-(> 7 repeats)	180	71	39.4
Tri-(> 5 repeats)	152	42	27.6
Tetra-(> 5 repeat)	62	3	4.8
Penta-(> 5 repeats)	10	-	-
Total	404	116	28.7

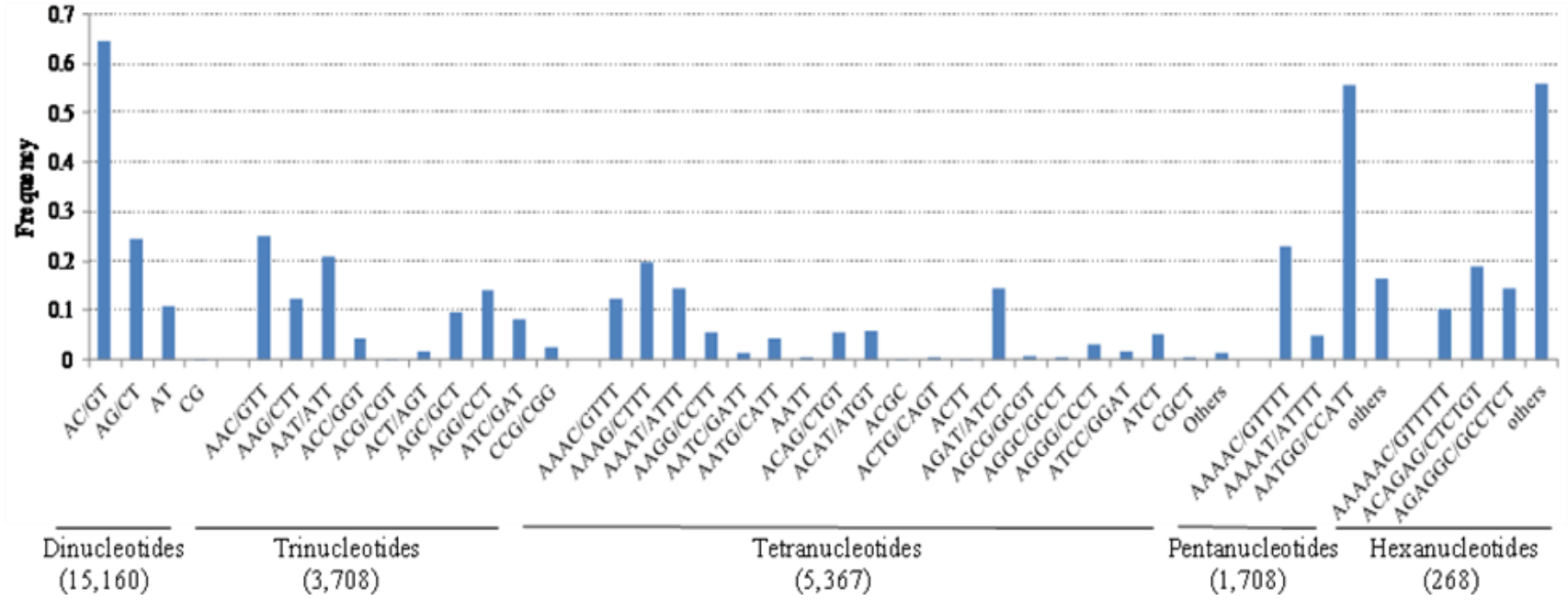

Fig 2.1 Frequency of different repeat types (Di – Hexa) of the grasscutter based on partial genome data. Frequencies were computed based on the total number of each repeat class indicated in parenthesis.

Table 2.3 Characteristics of 116 microsatellites of the grasscutter

Locus	Accession no.	Primer sequence (5' - 3')	Repeat motif	Ta	Size range (bp)	No. of alleles	H_O	H_E	NAF
Tsw01	AB725546	F: FAM-TGAGCCAGGAGAGCGTTTT	$(GGA)_8$	55	133 - 139	3	0.625	0.506	0
		R: TACAACCAGCAGCTCAGTCG							
Tsw02	AB725547	F: HEX-GATATCTTGTCCATGCAGGATT	$(GTCT)_8$	55	191 - 212	4	0.438	0.520	0.076
		R: AATGGATGCACCTTGAGACC							
Tsw03	AB725548	F: FAM-AATCATGGGCAAGATGGATG	$(GGAT)_{10}$	55	165 - 194	8	0.750	0.803	0.008
		R: TCTCCTGCCCTCATTTACCA							
Tsw04	AB725549	F: FAM-GGAAGCAACCCCATTAACAA	$(TAA)_{12}$	55	224 - 246	9	0.500	0.848	0.182
		R: TCTTCGGTTTTCTTCAAGCTG							
Tsw05	AB725550	F: FAM-ACAGCCAGCTCTAGGAACC	$(CA)_{11}$	55	304 - 308	3	0.563	0.635	0.073
		R: GCAGAAGGCCAAGGTATCG							
Tsw06	AB725551	F: FAM-TGGTTGATTTATTAGAGCCTGGA	$(CA)_{13}$	55	162 - 182	7	0.813	0.811	0
		R: TCTGTGTCCCCAAAGGAATC							
Tsw07	AB725552	F: FAM-TGCTACCTTTTCCTCCCTGT	$(CA)_9$	55	170 - 180	6	0.875	0.777	0
		R: GCAGCACCTCAACTGTGTGT							
Tsw08	AB725553	F: HEX-AAGGGGAGGGGAAAAGAGAT	$(CA)_{17}$	55	198 - 229	11	0.750	0.865	0.060
		R: CCAAGGAAAGAAGGGAAAGG							
Tsw09	AB725554	F: HEX-ACATGCACACATGCACACAC	$(CA)_{10}$	55	136 - 156	4	0.375	0.322	0
		R: GTCCACGGTCCTGATTGTCT							
Tsw10	AB725555	F: HEX-TCAGATAGGCGACGTGACTG	$(CA)_9$	55	142 - 156	5	0.500	0.736	0.119
		R: GGTGCTCCTATTCTGGTGGA							
Tsw11	AB725556	F: NED-TCATGAGCATAGCCTCAAGGT	$(CA)_{21}$	55	138 - 175	10	0.813	0.873	0.035
		R: ACCAAAACCCAATCAACCAG							
Tsw12	AB725557	F: HEX-CATCAGACGGGAGGTAAGGA	$(CA)_8$	55	134 - 158	8	0.688	0.713	0.018
		R: TCAAAATGTGGCAAGAACGA							

Tsw13	AB725558	F: NED-AGTGAGGCCCTGTCTCAAAA R: AGCTCCTCTTCCAGCCTACC	$(GAAT)_{11}$	60	194 - 220	7	0.625	0.691	0.047
Tsw14	AB725559	F: NED-GTTTCCAGCTGCTACCTTGG R: ATGGAGGCTGATACGACAGA	$(CA)_{9}$	55	204 - 208	3	0.250	0.518	0.173
Tsw15	AB725560	F: NED-TGCTGCTGCTGTTCATCTCT R: TCTCTTCACTCGTGTGTGTATGG	$(CA)_{13}$	60	238 - 242	3	0.188*	0.432	0.196
Tsw16	AB725561	F: NED-AGTCTATCATTGTTTGTTGCTCAG R: TGCCCTGTGTGTATCCATGT	$(CA)_{11}$	55	171 - 192	9	0.938	0.828	0
Tsw17	AB725562	F: NED-TGAGCCCTTGTTTTGGAGAG R: TTATCAGCCCTCCGATGAAC	$(CA)_{9}$	60	204 - 218	8	0.813	0.836	0.007
Tsw18	AB725563	F: FAM-TCAATCACTCTTCACCACACC R: TACAGGCTCCTGATGCAGTG	$(CA)_{19}$	55	182 - 208	9	0.563	0.822	0.128
Tsw19	AB725564	F: NED-AAGCCCATTCACTAAATCAAACA R: GGACGAAGTTGAAGCTCTGG	$(CA)_{15}$	55	148 - 166	7	1.000	0.783	0
Tsw20	AB725565	F: HEX-GGAGAGCCCCCAAATAACTT R: CCACGTGGTCATCTTTTACA	$(CA)_{13}$	55	214 - 220	4	0.563	0.727	0.099
Tsw21	AB725566	F: HEX-CAAGGCTGGTTTGGGCTAC R: CTGGGACAGGTAAGGGCTG	$(CA)_{10}$	60	163 - 192	8	0.750	0.813	0
Tsw22	AB725567	F: HEX-AATTACGAGGGCCTGTTCCT R: GAAGAGGCGTAGTCCGTGTC	$(CA)_{10}$	60	175 - 187	4	0.500	0.652	0.099
Tsw23	AB725568	F: FAM-AGCCCAAACTGCATACCAAG R: TCCCAAGCTCTTGTGACTCC	$(CA)_{11}$	60	140 - 160	8	0.625	0.768	0.065
Tsw24	AB725569	F: HEX-GGGTTTCAGCTCATCGTTGG R: GGTACCACCCAAAGGACCC	$(CAG)_{12}$	60	205 - 236	9	0.438*	0.863	0.236
Tsw25	AB725570	F: FAM-AGGGCACAGGTGGAGAAAC R: GAGGCCGTCTCGGTTGG	$(CAA)_{13}$	60	247 - 265	8	0.500	0.814	0.173
Tsw26	AB725571	F: NED-CCCTAGGTTGGCCTTCCAC R: ATTGCCACCTCTTGGGCTC	$(CAG)_{8}$	60	153 - 168	4	0.625	0.584	0

Tsw27	AB725572	F: HEX-AGGCAGAAGCTGTGAATGC R: GGCTTGCTCACTACGATGG	$(CAA)_9$	60	150 -156	3	0.375	0.537	0.123
Tsw28	AB725573	F: FAM-CTGCATTCCTCCCAAGCAC R: GACTTTCAGCAGCTTCGCC	$(GCT)_8$	60	180 - 189	3	0.500	0.639	0.065
Tsw29	AB725574	F: FAM-AGGCGCTGTTAAGCAAGATG R: GAGAATCGGGAGGACAGGC	$(GGA)_{10}$	60	233 - 242	4	0.438	0.682	0.145
Tsw30	AB725575	F: HEX-TGCACCTGTGAAGAACGGG R: CGTGGTGAAGACGCAATCC	$(GGC)_{10}$	60	253 - 266	5	0.563	0.691	0.064
Tsw31	AB725576	F: HEX-CCAGGGTCCTCGGTACTTC R: AGAGCGCGGTAGAACAAATC	$(TAA)_{10}$	60	234 - 264	10	0.500	0.846	0.184
Tsw32	AB725577	F: NED-AGAGCCTTCACACTGGGC R: TGATGGCTGTGAAACCCTG	$(TTA)_{13}$	60	225 - 252	9	0.688	0.750	0.036
Tsw33	AB725578	F: FAM-TCCTTGGGAAAGGATCCACC R: AATTGCAGTGCTCGCTTCG	$(TTA)_{13}$	60	126 - 162	10	0.375*	0.852	0.263
Tsw34		F: TGTTCCTCTGGCACCATCC R: GTAGTCCGTGTCCCTGGAG	$(CA)_{10}$	57	174 - 190	6	0.688	0.744	0.031
Tsw35		F: CCGCCGACTGCGTTTAAG R: TTTCCTGGTCCACTGACGC	$(CA)_{12}$	57	176 - 200	10	0.800	0.903	0.033
Tsw36		F: ACACTGAGCGTCACCTTCC R: TGTACCAAGAAGTCTAGCCCTG	$(CA)_{13}$	57	281 - 301	6	0.688	0.770	0.003
Tsw37		F: AGCTCAGTGTTCAATCAAGCC R: CCCTCCACATCCGCTTAAC	$(CA)_{14}$	57	222 - 262	12	0.813	0.919	0.036
Tsw38		F: GCCTCACAAAGGCTGGAAG R: CGAGAGCAGTGAATGTAGGTC	$(CA)_{17}$	57	224 - 240	7	0.688	0.782	0.000
Tsw39		F: GCCTGGGATTTACACCAGC R: CTGTGTTAGCTAGAGTTCTGCC	$(GTT)_8$	57	392 - 398	3	0.400	0.441	0.043
Tsw40		F: AAACATCGCAGTGCTGTCC R: CATGCTACAATGCCGGGTG	$(TTA)_{10}$	57	360 - 375	6	0.733	0.770	0.000
Tsw41		F: TTGCCTGGGCTCCATAGTG	$(CAA)_8$	57	162 - 171	4	0.600	0.598	0.000

	R: AAGAAGAGCCACCCATTGC							
Tsw42	F: GCAACGTCCCTTGGAACAG	$(TTA)_{12}$	57	305 - 349	10	0.615	0.880	0.122
	R: TGGTTCTCGCTCAGGACAG							
Tsw43	F: TCCCTGAGGACTAAAGCTGG	$(GGA)_{9}$	57	287 - 290	2	0.063	0.063	0.000
	R: AGGAGTAGTTGCCCGTGTC							
Tsw44	F: GGGAGGAGTTAGGTTCTGTCC	$(TTA)_{11}$	57	295 - 310	6	0.769	0.840	0.014
	R: CCCAGACCGGGCAATGTAG							
Tsw45	F: ACAGACTTTCGGAGCAGGG	$(CAA)_{8}$	57	237 - 246	4	0.385*	0.545	0.075
	R: CAGATCACTGGAGAAATGCGG							
Tsw46	F: GAAGCACAACTTGAGAAGCAC	$(CAA)_{11}$	57	267 - 279	5	0.400	0.687	0.160
	R: CAGTCGGTTAGAGGCAAGC							
Tsw47	F: GCCCAACTCAGTATAATGGCAC	$(GAT)_{8}$	57	239 -258	4	0.467	0.568	0.030
	R: TCCCTGTCACTGGAACTTTG							
Tsw48	F: GATCAGGAGTTTGAGGCCAG	$(CCA)_{16}$	57	278 - 334	10	0.429*	0.828	0.211
	R: GGACTTTAGCTGTTGCAGGTC							
Tsw49	F: TCTTCAATACGGCAGCCAC	$(CAA)_{11}$	57	193 - 214	3	0.188	0.433	0.176
	R: TGCAACCTCAGGACACAAATG							
Tsw50	F: GCTGGAGGAAGCCACAAAG	$(GAT)_{15}$	57	278 - 300	7	0.533	0.786	0.124
	R: TGGGTTGGTGGGACATACG							
Tsw51	F: AAACCTCGCCCTGGGATTG	$(GCT)_{23}$	57	249 - 321	10	0.333	0.855	0.265
	R: ACTTGGCTGGTTCCAAAGC							
Tsw52	F: ACACCTGTGGAGTCATTGC	$(GTT)_{10}$	57	193 - 229	7	0.750	0.718	0.000
	R: GGGCTGAGGGTATAATTTGGTG							
Tsw53	F: CAGCTGGAGGCTTTGTTGG	$(GGA)_{11}$	57	169 - 182	4	0.667	0.736	0.050
	R: TCATCTCAGAGCCGAAGGG							
Tsw54	F: AGAGGCTCATGGGAAGCAG	$(GTT)_{9}$	57	354 - 389	7	0.571	0.698	0.093
	R: AGACGTGAACTACCGTGCC							
Tsw55	F: ATACGCAAGGTACCGCTCC	$(TTA)_{17}$	57	241 - 281	12	0.733	0.910	0.099
	R: AGGACAAGAATCCATAAGATGACC							

Tsw56	F: CCTGGGACGGAAGCTCATC R: GGCCAGGTACTTAAAGCGG	$(GCC)_{10}$	57	287 - 347	10	0.750	0.889	0.093
Tsw57	F: CACCCGGCGATTCAACTTC R: GCCATGACTCCAGCAAACC	$(CAA)_{8}$	57	282 - 302	10	0.625	0.823	0.092
Tsw58	F: TGTGTGAGGCCTTGGGTTC R: AAGTGGCAGCCTATACGGG	$(TAA)_{8}$	57	212 - 353	6	0.438*	0.780	0.187
Tsw59	F: ACCTTCAGGAGACCACTTGC R: GCCCTTGGTTAAACGTCCC	$(CAT)_{14}$	57	203 - 224	6	0.667	0.772	0.000
Tsw60	F: TCCTGACTCTGAAGGAAGCC R: AGGACGTTCGTCAAATGTGG	$(TAA)_{9}$	57	328 - 337	3	0.167	0.453	0.184
Tsw61	F: TGACATAGGTTGCCAGAGATCC R: GCCATGTCGGTTTGGAAGG	$(CAT)_{9}$	57	225 - 256	4	0.231	0.563	0.177
Tsw62	F: CACAACTGCTGGGTGACAAG R: CCGTCCTTATAGTGACCTGAAATC	$(CAT)_{12}$	57	228 - 252	7	0.667	0.869	0.093
Tsw63	F: ACCCTCCGTAGCTTTGATCC R: TCACTGTATGGTGTTTGATCCC	$(CAA)_{9}$	57	183 - 198	5	0.733	0.726	0.000
Tsw64	F: TCCTGACTCTGAAGGAAGCC R: AAAGTGGAGGCACGCCAG	$(TAA)_{8}$	57	189 - 212	4	0.250	0.518	0.180
Tsw65	F: CGCCTGCATTCTGATCTGG R: CTGTGATGCCTGCGGAATG	$(CAG)_{9}$	57	230 -254	5	0.750	0.797	0.003
Tsw66	F: CTGGATCTTCTCAACATGCCC R: GCTTGCTAGGCTACATCACC	$(TAA)_{13}$	57	262 - 341	8	0.600	0.839	0.131
Tsw67	F: CAGGTTCAGCGGTTTAGCG R: GGACTTCAACCACCACAGAATG	$(CAA)_{11}$	57	309 - 333	8	0.813	0.770	0.000
Tsw68	F: CTGCAGCCATGAGCCAAAG R: CCTGAAATATGGTGGATGCCC	$(CAT)_{13}$	57	185 -210	9	0.643	0.817	0.086
Tsw69	F: AGGGCTGGGTATGTAGCTC R: CTCAACTCATCCTGAGCAAGG	$(CA)_{18}$	57	173 - 187	8	0.533	0.733	0.054
Tsw70	F: TGTTACCTGGTGCAGAGTTC	$(GT)_{11}$	57	352 - 360	3	0.400	0.543	0.120

	R: TCCAGTCTCTCCATAGGGTTC							
Tsw71	F: AGTTGCCTGGGACCTTGAG	$(CA)_{10}$	57	380 - 386	3	0.308	0.625	0.196
	R: CCACCTATAAGAGTCATTGCCC							
Tsw72	F: CTGGGCTCAAGGAATTACACC	$(CA)_{14}$	57	263 -275	6	0.875	0.782	0.000
	R: GACAGCATGAACCAGGCAC							
Tsw73	F: GCACTTTGGGAGGCTGAAG	$(CA)_{12}$	57	241 - 259	8	0.563	0.800	0.123
	R: TCACTCAGCCAACGTTCTG							
Tsw74	F: GGTCCTGTCTAGGCTGTGG	$(CT)_{8}$	57	196 - 198	2	0.563	0.417	0.000
	R: GGATAGATCTCCCTTGCATATTG							
Tsw75	F: CTCCTTTGGCACGTACAGC	$(CA)_{14}$	57	348 - 362	6	0.462	0.763	0.148
	R: ACCTGCTCTCTTCAGCCAG							
Tsw76	F: GGGCATGATGGGCTTTCAC	$(CT)_{9}$	57	206 - 210	3	0.400	0.570	0.104
	R: TGCTCTCAGTAGCTGATGGTC							
Tsw77	F: TGATCACCCTCAACACTGGG	$(CA)_{8}$	57	192 - 210	6	0.667	0.621	0.000
	R: TGTGCACCCATGATTATAGCAG							
Tsw78	F: ACAGCTGGGTGGCATTGAG	$(GT)_{15}$	57	316 - 336	8	0.917	0.870	0.000
	R: CATGGCTCCTGCCAAGTC							
Tsw79	F: GTAGGAAAGACCTTGGTGGC	$(GT)_{15}$	57	285 - 307	9	0.688	0.857	0.059
	R: AGTGCTGGACCTGAAACCC							
Tsw80	F: GGGAGGACAGATTTGGAGC	$(GT)_{9}$	57	189 - 209	7	0.333*	0.816	0.253
	R: AGAGGACAGCAAGGGTCTG							
Tsw81	F: CTGACTCTGCTCCCTGGAC	$(CA)_{10}$	57	308 - 322	5	0.714	0.709	0.000
	R: TCCTGTCTCCGAGCTTCTG							
Tsw82	F: CCCTTGCTTCTGCCTACG	$(CA)_{10}$	57	224 - 236	6	0.467	0.791	0.171
	R: CCCACGTCCTTGTAAAGCTG							
Tsw83	F: TTGGTTCCACTGGCTGAAG	$(CA)_{13}$	57	187 - 217	7	0.600	0.745	0.061
	R: ACTATGTGTCTGCCTTGCTAAC							
Tsw84	F: TGGGCATGTAGCTCAGTGG	$(CA)_{9}$	57	171 - 185	6	0.500	0.720	0.104
	R: CTTGCCGGTGCCTTTCAAC							

Tsw85	F: ATGCCACATGAGCATAACTTC R: CTCAGTATGAATGCTCTGAGTTC	$(GT)_{12}$	57	162 - 196	7	0.571	0.714	0.081
Tsw86	F: ACAGGGTTCTCCGCTGATG R: ATGACCCTTCGTCTCCACG	$(GT)_{11}$	57	270 - 282	7	0.813	0.827	0.000
Tsw87	F: ACACCTGCCTAGAAAGGGC R: GGAGTGCGAACCAATTTACCC	$(GT)_{14}$	57	263 - 291	10	0.600	0.887	0.147
Tsw88	F: ACATCTGGCAAAGGTGAGG R: ATGACTGCGTCCTTGCTAC	$(CA)_{8}$	57	175 - 179	3	0.714	0.582	0.000
Tsw89	F: GTCTAGATGCCAACCTCTGC R: CTGTCTTGCACATCAGTATCCC	$(GT)_{12}$	57	250 - 282	11	0.933	0.894	0.000
Tsw90	F: TTCTGGAAAGAGGGCTTCC R: GGAGTTGTGGGAGGCTGAC	$(CA)_{8}$	57	228 - 238	3	0.417	0.453	0.012
Tsw91	F: AAGTAGTACCGCGCCTACC R: TTTCTCCACGGATCCCAGC	$(CA)_{17}$	57	201 - 237	13	0.375*	0.875	0.251
Tsw92	F: GCCTGTCTAGCAAGCATAAGG R: ACTACTTGTCCTCTTGCCATTC	$(CA)_{19}$	57	155 - 173	10	0.857	0.892	0.000
Tsw93	F: GCACGTTTGGTGTCAGAGC R: TCCTTGTCTGCATAAAGAACTAGC	$(CA)_{12}$	57	184 - 194	5	0.600	0.740	0.065
Tsw94	F: CTGGTGAAGGAGACCAGCC R: CCTGGCATTTCAGCTCTGC	$(CA)_{11}$	57	297 - 307	5	0.667	0.754	0.026
Tsw95	F: GAACTTGAAGCTGCCGGAC R: GCGACATGCCTGTAATCCC	$(CT)_{9}$	57	244 - 258	7	0.533*	0.800	0.120
Tsw96	F: AATCCCAGCACAATGGGAG R: CAGCTTCTTCCCAGGTTAACAG	$(GT)_{10}$	57	198 - 202	3	0.500	0.534	0.006
Tsw97	F: GGCCTAAGGTCCTGAATCG R: GCAGCAGCTTACAGACGTG	$(GT)_{8}$	57	275 - 277	2	0.333	0.287	0.000
Tsw98	F: GCCCTTCCCAGAGGTCAAC R: AGCCTTTCCTCCTCTAATGGG	$(CA)_{12}$	57	303 - 336	10	0.750	0.906	0.080
Tsw99	F: CCCTACACCCATTTCCCTTTG	$(CA)_{11}$	57	310 - 318	5	0.800	0.754	0.000

	R: AGCTCAGTCTTGCAAACAACC							
Tsw100	F: CACGCTGGACAACACAAGG R: CCATTGCAGATGAGGAGCC	$(CA)_9$	57	362 - 372	5	0.833	0.652	0.000
Tsw101	F: GCCTGATTGAGCTCTTGGTG R: AGGGACTGAATGCCACCTC	$(CA)_{12}$	57	334 - 380	8	0.688	0.845	0.076
Tsw102	F: TCTGGGTTCCACCTCTAATGC R: CACAGTTTCATCAAGTGACAGC	$(CA)_9$	57	216 - 264	5	0.188	0.383	0.164
Tsw103	F: GCTGGCAGACAACTGCTTC R: CAGGATCTGAACCCACAGG	$(GT)_8$	57	312 - 322	5	0.688	0.673	0.000
Tsw104	F: GGGAGGACAGATTTGGAGC R: AGAGGACAGCAAGGGTCTG	$(GT)_9$	57	158 - 210	7	0.400	0.816	0.215
Tsw105	F: TATGGTCACCGGACTTCCC R: AGACTGTTTGGCAGAGGTG	$(CA)_{14}$	57	206 - 306	5	0.167*	0.377	0.153
Tsw106	F: AGCCATCTGTTTACTCCAGGG R: GCATCAGCTGCTTTCTGGG	$(CA)_{13}$	57	298 - 328	11	0.800	0.772	0.000
Tsw107	F: GGAAACAGGCATTCTGGCG R: GGAAACAGGCATTCTGGCG	$(CA)_{12}$	57	171 - 181	5	0.533	0.692	0.041
Tsw108	F: GTGCGTAGCATGAGAATGTCC R: GGAATCTCCAGAGGGTGCC	$(CA)_8$	57	203 - 235	6	0.400	0.616	0.135
Tsw109	F: TGCTAGAGCCATCCAGGTC R: GGGAACTTTGGGACAACCC	$(CT)_{11}$	57	338 - 364	6	0.438*	0.694	0.154
Tsw110	F: TCACTTGTCTAGGCAGGGC R: TGATGCTCCGGTTGTCCTC	$(GT)_9$	57	258 - 292	8	0.500	0.826	0.183
Tsw111	F: TGCCAAATGTCAGTCTCCG R: ACGGGTTTGGAATTGTGCTC	$(GT)_8$	57	216 - 234	8	0.800	0.789	0.028
Tsw112	F: AGGGTTCCCAGCAGAAAGG R: TGTCTGCTAGAGGCAGGTC	$(GA)_{13}$	57	150 - 228	9	0.500	0.877	0.191
Tsw113	F: CGATGTCGCTCTTCCACG R: TGTTCTCCAGGTGCAGTTTAAG	$(CA)_{13}$	57	314 - 348	12	0.667	0.910	0.120

Tsw114	F: ACCCACAATCTTCCGGGTG R: AGTCACCATCAAGGGCAGG	$(CA)_{10}$	57	250 - 272	6	0.467	0.743	0.138
Tsw115	F: TCCAGAGACCCACAGTTTCC R: ACCATCGTTCCACAACTTTCC	$(CA)_{11}$	57	362 - 392	10	0.571	0.865	0.139
Tsw116	F: GGAGCCATGACTTGTGGAAG R: GCCAGCCATTTATAGCTTGTTC	$(GT)_8$	57	190 - 224	4	0.467	0.634	0.114
				Average	6.5	0.575	0.723	

Ta: Optimal annealing temperature, H_O: observed heterozygosity, H_E: expected heterozygosity, NAF: null allele frequency
Markers Tsw 34 – Tsw 116 were tagged with M13 to reduce cost of labelling
*Significant deviation from Hardy-Weinberg equilibrium after Bonferroni correction

3. GENETIC DIVERSITY OF GRASSCUTTER IN GHANA

3.1 Mitochondrial D-loop Diversity of Grasscutter (*Thryonomys swinderianus, RODENTIA, HYSTRICOMORPHA*) in Ghana

3.1.1 Introduction

Surveys done in Ghana have revealed that grasscutter meat is the most preferred by 'chop bar' operators (cooked food vendors) and consumers (Ntiamoa-Baidu, 1998; Owusu et al., 2006). The meat is eaten by all classes of people with no religious prohibitions (Opara, 2010b), and is also exported to continental Europe and the United States, where it is sold mainly to West Africans living in those regions (Adu et al., 1999).

Grasscutter domestication started in the 1970s but efforts over the years have met with little success. Rearing attempts suffered from high mortality due to the aggressive nature of the species, referred to as 'berserk behavior' (Adu et al., 1999). In any domestication process, selection for desirable traits is of great importance to ensure ease of handling and for profitable production in the case of the grasscutter, which is being developed as a mini-livestock in Sub-Saharan Africa to alleviate poverty and to cater for the protein needs of the people. Various aspects of grasscutter biology have been studied in order to better manage the species under domestic conditions. These include reproduction (Asibey, 1981; Adu and Yeboah, 2000; Addo et al., 2002; Adu, 2003; Owusu et al., 2010; Henry, 2011), nutrition (Annor et al., 2008; Karikari and Nyameasem, 2009), parasites and diseases (Oboegbulem and Okoronkwo, 1990; Opara and Fagbemi, 2008; Kankam et al., 2009). However, much remains to be done on the genetics of the grasscutter production, especially at the molecular level.

Mitochondrial DNA is known to be maternally-inherited, and non-recombinant with a high mutation rate (Avise et al., 1987; Chen and Herbert, 1999; Larizza et al., 2002). The D-loop is hypervariable, non-coding part of the mitochondrial genome that is approximately 1 kb long and less than 7% of the total mitochondrial genome of most mammals (Larizza et al., 2002). It is also known to regulate transcription and replication (Abyankar et al., 2009). The above characteristics of the mitochondrial DNA, specifically the D-loop, make it an excellent marker for population genetics studies. Also, because of easier amplification due to the presence of multiple copies within a cell, mitochondrial DNA has been used extensively to study populations and to trace maternal lineages. Sequence analysis of the D-loop as a

dominant marker can efficiently reveal genetic structure and differentiation among populations (Hirota et al., 2004; Meyer et al., 2009). These genetic diversity measures of populations nevertheless serve as useful information for conservation.

Management of populations for conservation requires baseline information such as genetic structure and diversity. Even though the grasscutter is not expected to be endangered in the foreseeable future, increasing our understanding of the population dynamics and genetic diversity of the species is necessary for monitoring. The objective of this study was therefore to assess the genetic diversity among grasscutter populations inhabiting three agro-ecological zones in Ghana. As far as we are aware, no genetic diversity study about the grasscutter has been reported in the literature. This study therefore reports for the first time, the mitochondrial D-loop diversity of grasscutter populations in Ghana.

3.1.2 Materials and methods

3.1.2.1 Sample Collection

A total of 84 hair samples were collected from grasscutters in three agro-ecological zones in Ghana; Coastal Savanna (n = 45), Forest (n = 22) and Guinea Savanna (n = 17) (Fig 3.1.1). The different zones differ in terms of climatic conditions (e.g. rainfall pattern and temperature) and vegetation type and therefore each was considered to be habitat for a separate population of grasscutters. In the Coastal Savanna zone, samples were collected from suburbs of Accra (Greater-Accra Region), Mankessim and Jukwa (Central Region). Forest samples were obtained from Kumasi (Ashanti Region), while Guinea Savanna samples were collected from Mole and Tamale (Northern Region) in Ghana. Samples from Kumasi were collected from a centrally located bushmeat market where all hunted animals (including grasscutters) from surrounding localities were traded. The exact locations could however not be ascertained. Out of the 45 Coastal Savanna samples, 21 were obtained from semi-domesticated unrelated individuals which were kept by farmers. All wild samples were collected from hunters' kills, and DNA was extracted from the root of 15-20 hair pieces of each sample using Instagene Matrix (Bio-Rad Laboratories, USA). The DNA was then quantified using NanoDrop Spectrophotometer (Thermo Scientific, USA) and stored at -30˚C until ready for use.

3.1.2.2 PCR Amplification and Sequencing

Mitochondrial displacement loop (D-loop) was amplified in a PCR using CRmtDF (5'-CCAACTCCCAAAGCTGATGT-3') as the forward primer, and CRmtDR (5'-GGCACCAACATCATCACAAA-3') as the reverse primer. These primers were designed with a registered sequence of the grasscutter (Accession no. AJ301644) (Mouchaty et al., 2001) using Primer 3 software and could amplify 501 bp of the D-loop. The PCR mixture contained 0.75 U of *LA-Taq* DNA polymerase (TaKaRa, Shiga, Japan), PCR buffer, 400 µM of each dNTP, 0.4 µM each of forward and reverse primers, 0.1 µg of T4 Gene 32 Protein (Nippon Gene, Japan) and 20 ng of template DNA in a total volume of 15 µl. PCR cycling conditions consisted of an initial denaturation of 95° C for 2 minutes, followed by 35 cycles of 95° C for 30 seconds, 55° C for 30 seconds, 74° C for 1 minute and a final extension of 74° C for 10 minutes. Aliquots of 5 µl of the PCR products were electrophoresed on 1.5% agarose gel to check amplification. DNA bands were visualized after Ethidium Bromide staining under UV light, and expected size was determined in relation to a DNA size standard. The remaining aliquots were purified using High Pure PCR purification kit (Roche, Manheim, Germany) and sequenced using Big Dye Terminator ver. 3.1 cycle sequencing kit (Applied Biosystems, Foster City, CA, USA) according to the manufacturers protocol and electrophoresed on an ABI PRISM 3130xl sequencer (Applied Biosystems).

Both forward and reverse complements of the reverse sequences were aligned to get a consensus sequence using MEGA ver 5 (Tamura et al., 2011). Primer sequences were then deleted to obtain 501 bp of the D-loop spanning from position 15,661 to 16,161 of the mitochondrial genome (Mouchaty et al., 2001), and covering the rapidly evolving extended-termination associated sequences (ETAS) domain and part of the central domain (CD) of the D-loop (Larizza et al., 2002).

3.1.2.3 Data Analysis

Arlequin ver 3.5 (Excoffier and Lischer, 2010) was used to determine the number of haplotypes, haplotype diversity and nucleotide diversity in each agro-ecological zone. Haplotype diversity indicates genetic diversity within populations and nucleotide diversity is estimated as the function of the number of polymorphic sites and the frequency of transition or transversion and insertion or deletion within the population (Hirota et al., 2004). A haplotype network was constructed using outputs from TCS ver 1.21 (Clement et al., 2000)

and Network software ver 4.6 (www.fluxus-engineering.com). Analysis of molecular variance (AMOVA) was conducted in Arlequin, considering genetic distance between haplotypes and their frequencies to determine the variation among and within populations. It was assumed that the three populations formed one group and therefore the pooled flat genetic structure in Arlequin was chosen. In order to increase the level of accuracy, 10,000 random permutations were performed. The gamma value was set at 0 as this is proposed when mutation rates cannot be assumed to be uniform (Excoffier and Lischer, 2010). This is usually the case for mtDNA D-loop and most especially for rodents because they have short generation intervals (Bromham et al., 1996; Li et al., 1996). The significance (*p*-value) of the fixation index which is analogous to the *F*-value in the conventional analysis of variance was estimated as the probability that a random value is greater than or equal to the observed value. To determine which populations were genetically different from each other, pairwise F_{ST} values which show genetic distance between given pair of populations were computed (Reynolds et al., 1983).

According to Harpending (1994), a mismatch distribution is a measure of the distribution of pairwise differences among non-recombinant DNA sequences in a population. This distribution is multimodal, smooth and has a peak for populations that have undergone expansion but erratic or ragged for populations in equilibrium (Harpending, 1994; Schneider and Excoffier, 1999; Excoffier, 2004). Past demographic events are known to leave footprints in the DNA sequence of individuals in a population. To get a glimpse of the demographic history of the grasscutter populations, mismatch distributions (10,000 bootstrap replicates) were determined and Tajima's *D* and Fu's F_S neutrality indices were computed under the infinite site model as implemented in the Arlequin software (Excoffier and Lischer, 2010). The number of simulated samples under the infinite site model was 10,000.

To examine the relationship between geographic distance and genetic distance (equilibrium between gene flow and genetic drift), isolation-by-distance analysis was conducted as implemented in GenAlex ver 6.41 (Peakall and Smouse, 2006). Coordinates between pairs of sampling sites were converted into geographic distances, and a Mantel test was performed to assess the correlation between geographic distance and pairwise F_{ST} values.

3.1.3 Results

3.1.3.1 Haplotype Diversity

Guinea Savanna samples had 21 polymorphic sites whereas Coastal Savanna and Forest zones had 15 and 12 polymorphic sites, respectively. In all, there were 23 variable sites including five singletons and 18 parsimonious informative sites. All polymorphic sites were transitions except for one transversion at position 16,125 in a Guinea Savanna zone haploytpe. Out of a total of 26 haplotypes, 15 were from Guinea Savanna, seven from Forest and 13 from the Coastal Savanna (Table 3.1.1). The sequences were deposited in Genbank under the accession numbers AB675385 to AB675410.The Forest and Coastal Savanna zones shared three haplotypes whilst five haplotypes were shared between Coastal Savanna and Guinea Savanna. Forest and Guinea Savanna also had three haplotypes in common. In all, only two haplotypes were found to be common to the three agro-ecological zones. Haplotype diversities were 0.978, 0.853 and 0.875 for Guinea Savanna, Forest and Coastal Savanna zones respectively (Table 3.1.2). Even though there were fewer samples from the Guinea Savanna zone, the Guinea Savanna zone was found to have the highest haplotype diversity compared to the Forest and the Coastal Savanna zones.

Nucleotide diversity was 0.012 for both Guinea Savanna and Coastal Savanna zones which is almost twice as that of the Forest zone (0.007) (Table 3.1.2). AMOVA results indicated that 85.79% of the total variation was within populations whilst the remaining 14.21% was among populations (Table 3.1.3). The fixation index resulting from the AMOVA analysis which gives an indication of population differentiation, was found to be highly significant ($p < 0.001$) even though the larger part of the variation was within populations. This indicated that the populations are less structured. In terms of pairwise F_{ST}, Guinea Savanna and Coastal Savanna had the lowest F_{ST} value (0.052) indicating that they are the closest, while Forest and Coastal Savanna had the highest value (Table 3.1.4). The results indicated significant genetic differentiation ($p < 0.05$) between Forest and Coastal Savanna and also between Forest and Guinea Savanna. There was however no significant differentiation between Guinea Savanna and Coastal Savanna zone haplotypes.

Figure 3.1.2 presents a haplotype network showing two clusters, one of which is fairly simple and consists of only Savanna haplotypes (Cluster A), and the other which harbours the two common haplotypes, consists of haplotypes from all zones, and is more complex (Cluster

B). Such a network shows the relatedness of the different haplotypes based on nucleotide substitutions. It can be seen that the Savanna haplotypes are more related than Forest haplotypes. Haplotypes from both savanna zones can be seen in both clusters whereas the Forest haplotypes are found only in cluster B. This is indicative of the presence of two haplo-groups within the two savanna zones.

3.1.3.2 Mismatch Distribution and Population Neutrality

Figure 3.1.3 shows mismatch distributions of the three populations, with the observed mismatch distribution being multimodal for all populations, though not significant. In both Guinea and Coastal Savanna, two peaks were found, which reflect the history of the two savanna populations. Harpending's raggedness index, *rg*, Tajima's *D* and Fu's F_S are presented in Table 3.1.2. For both *rg* and *D,* none of the zones was found to be significant ($p > 0.05$). These statistical tests therefore indicate that the populations were under neutral selection. Fu's F_S was however significant for Guinea Savanna zone ($p = 0.002$), which is an indication of a recent past expansion event in this population (Fu, 1997). A Mantel test conducted to assess the equilibrium between genetic drift and gene flow, was found to be non-significant ($Rxy = 0.022$, $p = 0.334$).

3.1.4 Discussion

In this study, it is not unexpected that 26 different haplotypes were found coupled with high haplotype diversities (0.85 – 0.97) across the three zones, owing to the fact that the D-loop has a high mutation rate (Larizza et al., 2002). These results are comparable to a previous report on a South American hystricomorph rodent, *Microcavia australis* (Sassi et al., 2011) which indicated 0.70 – 0.93 for haplotype diversity (h) and 0.0006 – 0.0095 for nucleotide diversity (π). According to Lacy (1987), lack of gene flow diminishes genetic diversity. Also, genetic drift, dispersal and vicariance events such as habitat fragmentation may all influence genetic structure and diversity of mammals. For instance Sassi et al. (2011) found highly differentiated highland and lowland populations of *M. australis* with almost all haplotypes being unique to populations, indicating very restricted levels of gene flow between the populations. Similar results were obtained by Ojeda (2010) on another rodent, *Tympanoctomys barrerae*, in the same region.

The possible reason for the higher haplotype diversities could be that there are no or limited barriers to gene flow as grasscutters are good swimmers (Opara, 2010a) and can cross water bodies which might be perceived as barriers to dispersal. It has been observed that grasscutters in Ghana continue to expand their habitats in the Forest zone due to forest clearance for farming activities (Opara, 2010a). This possible dispersal could account for the two common haplotypes found among the three zones. We hypothesize that there is unrestricted gene flow among the populations especially between the Guinea Savanna and Coastal Savanna zones. More extensive sampling focusing on social structure will however be necessary to test this hypothesis. Another possible reason for the high haplotype diversities could be the very large population of grasscutters which existed in the recent past or presently exist in the country. This assertion could be supported with the fact that grasscutters can be found in almost every part of Ghana (Annor et al., 2009). Grasscutters are predominantly savanna species (Jori et al., 1995) which thrive on grasses and succulent stems, hence the name 'grasscutter'. It is probable that the Forest population of grasscutters arose out of a recent colonization of the Forest zone by grascutters from the Guinea Savanna zone, hence the closer genetic distance between Forest and Guinea Savanna than Forest and Coastal Savanna populations. According to Slatkin (1993), lack of equilibrium between genetic drift and gene flow or absence of isolation by distance, which is the case in this study, may be due to a recent colonization event, in turn linked to human impact in the Forest zone even though no significant population expansion was found. Agriculture-related activities such as forest clearance, bush burning and cultivation of crops result in expansion of grassland areas within the Forest zone and ultimately forest fragmentation. These farmlands and grassland areas have become suitable habitats for the grasscutter.

Non-significant population differentiation between the two savanna zones suggests that there is direct gene flow between the Guinea Savanna and the Coastal Savanna zones. This genetic distance reflects a dispersal pattern; possibly from the Guinea Savanna zone to the Coastal Savanna zone. This dispersal pattern might be a result of limited availability of feed resources due to harsh climatic conditions in the Guinea Savanna zone, characterised by relatively higher temperatures and a single rainy season between July and October (Bennett-Lartey et al., 2002). Longer periods of aridity in this zone coupled with higher temperatures, may cause grasses to dry up. This reduces the amount of feed resources available, forcing groups of individuals to migrate to places where feed resources are readily available.

The Ghanaian populations of grasscutters, except for those in the Guinea Savanna zone, have been stable as observed from the mismatch distributions. Harpending (1994) noted that populations that have undergone expansion show mismatch distributions that are unimodal, but populations that are stable have erratic distributions. None of the three zones was significant for Harpending's raggedness index, *rg* and Tajima's *D*. The results of mismatch distribution from the three populations reflect a multimodal distribution, including the Guinea Savanna. The difference is however found in the Fu's F_S test where the Guinea Savanna population displays a significant negative value (p = 0.002), suggesting a demographic expansion event. It is known that Fu's F_S has more statistical power than Tajima's *D* (Ramos-Onsins and Rozas, 2002) and it is sensitive to other demographic events such as population bottlenecks and genetic hitchhiking (Fu, 1997). Results should however be interpreted cautiously because the different tests have different sensitivities (Abyankar et al., 2009). According to Gaines et al. (1997), populations with bottlenecks show low variation in mitochondrial DNA. A population bottleneck might not be the likely reason for the significant F_S. Genetic hitchhiking is also ruled out because mitochondrial DNA is non-recombinant (Larizza et al., 2002). A past expansion event could account for this significance and this could be supported by the relatively higher haplotype diversity observed in the Guinea Savanna zone compared to other populations. Fu (1997) however cautioned that one cannot make a definite conclusion by using one gene and, that a combination of nuclear and mitochondrial markers may be necessary.

Ojeda (2010) suggested that lack of isolation-by-distance pattern could be due to recent expansion to new areas and inconsistent gene flow. We did not find any pattern of isolation-by-distance in our study, which supposes that an expansion event could influence this result. It is possible that the recent population expansion portrayed by the Guinea Savanna zone as found in the Fu's neutrality test and a recent colonization of the Forest zone as deduced from the results influenced the population structure and consequently the equilibrium between gene flow and genetic drift.

Within the Guinea Savanna and Coastal Savanna zones, double peaks were observed from the mismatch distribution (Fig 3.1.3). Even though the population history of the grasscutter in Ghana is not known, this result could be an indication of the coexistence of two haplo-groups in the two savanna zones. This tendency which is also clearly evident from the network analysis (Fig 3.1.2) is not profound in the Forest zone probably because of habitat

limitations. This could also explain why the genetic differentiation between the Guinea Savanna and Coastal Savanna populations is not significant. It can be deduced from the double peaks and the two clusters that the two savanna populations probably recently intermixed. The two populations may therefore be considered as one Savanna population.

Since the grasscutter is a 'Least Concern' species according to the IUCN Red List (IUCN, 2011), it does not require to be targeted for conservation at the present time. It is however envisaged that recent human population expansion in Ghana coupled with rampant hunting, if not checked, might cause this species to be endangered in the future. Also, bushfire associated with the hunting of the grasscutters may destroy the habitat of many wildlife species. This study has shed some light on the genetic diversity of grasscutters in Ghana, and this information may be valuable for future conservation efforts. It has also revealed the possibility that Guinea Savanna zone population of grasscutters has undergone population expansion in the past. It may be worthwhile to confirm these results with microsatellite analysis of the populations since mitochondrial and nuclear markers can sometimes present conflicting results due to different modes of inheritance and rates of evolution (Flanders et al., 2009). A more detailed study focusing on genetic structure and dispersal patterns within each agro-ecological zone would be necessary to ascertain these results. As domestication of this species is being intensified because of its good prospects as a mini-livestock (Ntiamoa-Baidu, 1998; Adu et al., 1999; Addo et al., 2002), information presented in this paper will serve as a baseline for further population genetics studies.

Table 3.1.1 Haplotypes and their frequencies in each population

Accession no.	Haplotype	Zone			Total
		Guinea Savanna	Forest	Coastal Savanna	
AB675402	G1	1	-	-	1
AB675403	G2	1	-	-	1
AB675404	G3	1	-	-	1
AB675405	G4	1	-	-	1
AB675406	G5	1	-	-	1
AB675407	G6	1	-	-	1
AB675408	G7	1	-	-	1
AB675409	G8	1	-	-	1
AB675410	G9	1	-	-	1
AB675388	F1	-	3	-	3
AB675390	F2	-	1	-	1
AB675391	F3	-	1	-	1
AB675392	C1	-	-	1	1
AB675393	C2	-	-	1	1
AB675394	C3	-	-	1	1
AB675396	C4	-	-	9	9
AB675397	C5	-	-	1	1
AB675398	C6	-	-	1	1
AB675401	C7	-	-	1	1
AB675386	GF	1	5	-	6
AB675387	FC	-	5	9	14
AB675395	GC1	3	-	9	12
AB675399	GC2	1	-	5	6
AB675400	GC3	1	-	2	3
AB675385	GFC1	1	5	3	9
AB675389	GFC2	1	2	2	5
Total		17	22	45	84

Table 3.1.2 Genetic diversity and indices of population neutrality across the three agro-ecological zones

Zone	*n*	No. of haplotypes	Haplotype diversity, *h*	Nucleotide diversity, π	Harpending's raggedness, *rg**	Tajima's *D**	Fu's F_S*
Guinea Savanna	17	15	0.978 ± 0.031	0.012 ± 0.007	0.022(0.613)	0.189(0.619)	-7.134(0.002)
Forest	22	7	0.853 ± 0.037	0.007 ± 0.004	0.034(0.886)	0.560(0.745)	0.989(0.706)
Coastal Savanna	45	13	0.875 ± 0.024	0.012 ± 0.006	0.057(0.690)	1.442(0.942)	0.839(0.673)

**p*-values are indicated in parenthesis

Table 3.1.3 Results of analysis of molecular variance

Source of variation	d.f	Sum of squares	Variance component	% variation	F_{ST}	*p*
Among populations	2	29.46	0.47	14.21	0.14	0.0002
Within populations	81	229.57	2.83	85.79		
Total	83	259.03	3.3	100		

Table 3.1.4 Matrix of pairwise F_{ST} values for the three agro-ecological zones

	Guinea Savannah	Forest
Guinea Savanna	-	
Forest	0.078 (0.018)*	-
Coastal Savanna	0.052 (0.090)	0.225 (0.000)*

*significant *p*-values

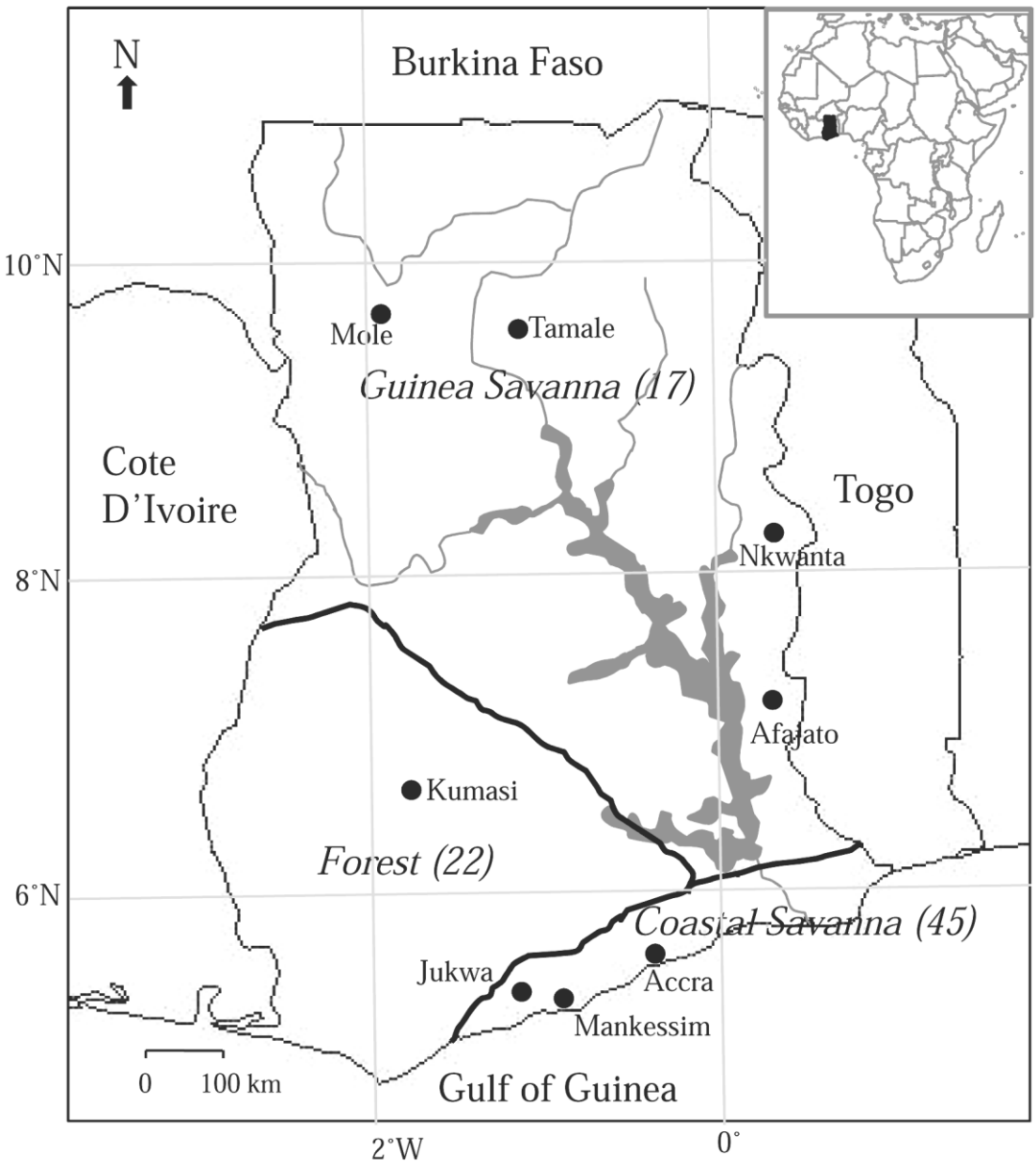

Fig 3.1.1 The map of Ghana showing three agro-ecological zones. Numbers in parenthesis indicate number of samples collected in each zone.

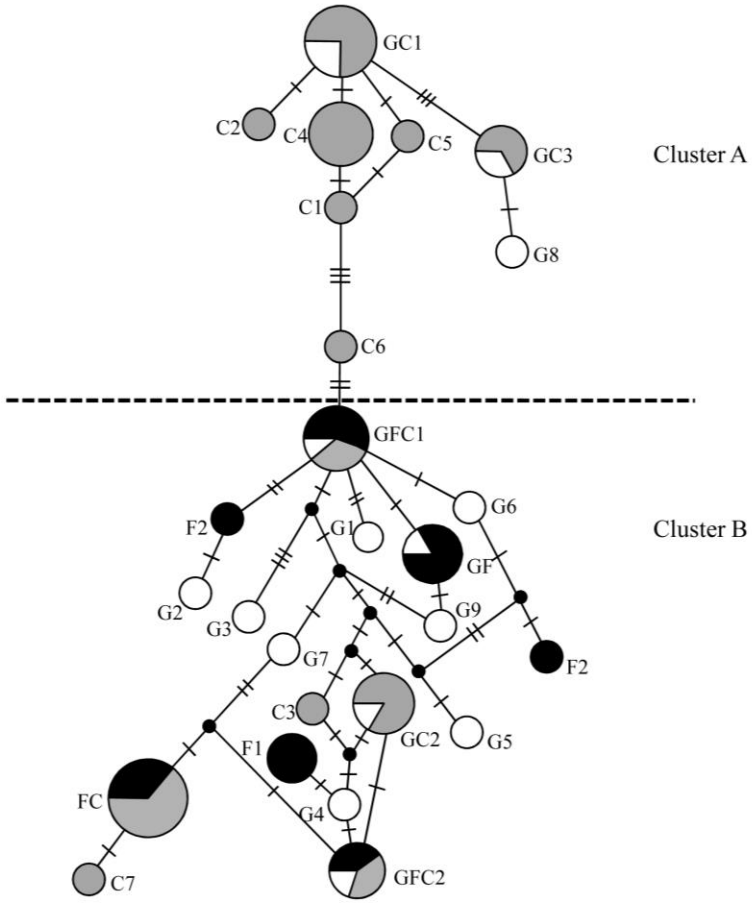

Fig 3.1.2 Network among haplotypes. G1-G9 indicate Guinea Savanna zone specific haplotypes (open circles), F1-F3 are Forest zone specific haplotypes (black circles), C1-C7 are Coastal Savanna zone specific haplotypes (gray circles), GFC denotes common haplotypes, GC denotes haplotypes shared between Guinea Savanna and Coastal Savanna zones, FC denotes haplotypes shared between Forest and Coastal Savanna zones and GF denotes haplotypes shared between Forest and Guinea Savanna zones. The proportions of individuals from each agro-ecological zone in the common and shared haplotypes are shown as slices. The size of the circle is proportional to the total number of individuals of that haplotype. The crossbars show the number of substitution between haplotypes whilst the black nodes indicate missing haplotypes.

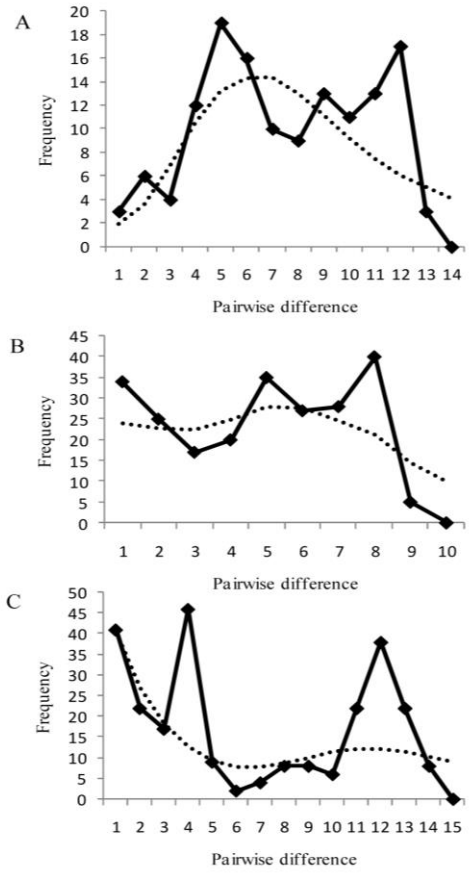

Fig 3.1.3 Harpendings mismatch distribution showing observed and simulated frequencies of pairwise differences for each population. A: Guinea Savanna zone, B: Forest zone, C: Coastal Savanna zone. Observed frequency and simulated frequency are shown by thick line and dotted line, respectively.

3.2 Genetic diversity of grasscutter (*Thryonomys swinderianus*) in Ghana based on microsatellite markers

3.2.1 Introduction

Even though there are concerns about the techniques used for hunting the grasscutter in the wild, the grasscutter bush meat trade continues to flourish, making significant contributions to the Ghanaian economy as well as the economies of other African countries where the grasscutter meat is consumed (Ntiamoa-Badu, 1997). It is therefore imperative that the grasscutter be given the attention it deserves so that it can contribute even more to those economies without jeopardizing wildlife diversity and the environment in general. Despite the enormous significance, very little is known about the ecology, demographics and population dynamics including dispersal and breeding structure of grasscutters in the wild. Determining genetic structure and variation may therefore shed light on some of these mechanisms. In this study, we used novel microsatellite markers of the grasscutter to assess the genetic diversity and population structure of their populations in Ghana.

Since their discovery, microsatellite markers have been employed by ecologists to investigate the ecology of many taxa in order to answer questions relating to population size, kinship, migration rates, bottlenecks, effect of landscape, threat status etc (Selkoe and Toonen, 2006). These questions which hitherto baffled ecologists can now be addressed due to the resolving power of microsatellites and also because there have been tremendous advancement in computing technology. Microsatellites are the most widely used markers for molecular ecological studies to date. They are selectively neutral and follow Mendelian inheritance parttern. Furthermore, they can be used to analyse ancient degraded DNA or DNA from non-invasive samples such as hair and faeces thereby making them invaluable tools for ecologists (Taberlet et al., 1999)

In Ghana, wild grasscutter populations exist throughout the country and are largely not studied. Owing to the development of novel microsatellite markers (Adenyo et al., 2012, Chapter 2), we aim to characterise the populations using these markers to understand the population dynamics in order to provide useful information for their management and for the purpose of future conservation if the need arises. There are many rivers in Ghana, including the Volta River which has now formed a lake due to a hydro electric power dam constructed in the 1960s at Akosombo. We envisaged that this lake might have a profound effect on the fauna populations (including the grasscutter) in terms of their dispersal.

Since mitochondrial and nuclear markers have different modes of inheritance and rates of evolution, they present often times conflicting results depending on the characteristics of the populations being studied (Flanders et al., 2009). Mechanisms such as sex-biased dispersal, incomplete lineage sorting, homoplasy and effective population size may have different influences on inferences made on populations, depending on whether mitochondrial or nuclear markers are used (Clutton-Brock, 1989; Colbert et al., 2001; Hewitt, 2004). In a recent study, we characterised the populations using mitochondrial DNA marker (D-loop) (Adenyo et al., 2013, Chapter 3.1). Here, we quantify the distribution of the genetic variation of grasscutters in Ghana using microsatellite markers.

3.2.2 Materials and Methods

3.2.2.1 Sampling and DNA extraction

A total of 66 hair samples were collected from grasscutters hunted from the wild by hunters and farmers in three agro-ecological zones of Ghana. In Guinea Savanna, samples were taken from around the Mole national park (n = 10) and Tamale (n = 9). Sixteen samples were collected from a bushmeat market in Kumasi which is located in the Forest zone. In the Coastal Savanna zone, samples were taken from Mankessim (n = 5), Jukwa (n = 7) and Accra (n = 4). Additionally, eight samples were taken from Nkwanta and seven from Afajato area, both located in the Volta Region which is at the eastern side of the Volta Lake (Fig 3.1.1). This was done in order to test the effect of the Volta Lake on the grasscutter populations in the Guinea Savanna zone. DNA was extracted from 1-2 mm hair root clippings of 15-20 hair pieces using Instagene Matrix (Bio-Rad Laboratories, USA). The quantity and quality of DNA was measured using NanoDrop spectrophotometer (Thermo Scientific, USA). The DNA was then stored at -30 until ready for use.

3.2.2.2 Microsatellite genotyping

Twelve highly polymorphic microsatellite markers including three tetra-repeats and nine di-repeats (*Tsw02, Tsw03, Tsw06, Tsw07, Tsw08, Tsw09, Tsw11, Tsw13, Tsw16, Tsw19, Tsw21* and *Tsw23*) were selected from a previously published panel (Adenyo et al., 2012, Chapter 2) for genotyping. Multiplex PCR was conducted with 3-5 primers per reaction using the QIAGEN Multiplex PCR Kit (QIAGEN, Valencia, CA, USA) under the following conditions; initial denaturation at 95˚C for 15 minutes, 35 cycles of 94˚C for 30 seconds,

55˚C or 60˚C for 90 seconds and a final extension at 60˚C for 30 minutes. PCR mixture contained 20 ng of DNA template, 0.2 µM of each primer of which the forward primers were fluorescently labelled (6FAM, HEX and NED), 0.1 µg of T4 Gene 32 Protein (Nippon Gene, Japan) and 2x QIAGEN Multiplex PCR Master Mix in a total volume of 10 µL. PCR was performed on GeneAmp PCR System 9700 (Applied Biosystems). Each reaction was repeated 2 – 3 times to confirm the genotype of each individual. The PCR products were thereafter diluted (1:100) and electrophoresed on an ABI 3130xl DNA Sequencer (Applied Biosystems) and the sizes of the fragments were scored with 400 HD Rox size standard using Peak Scanner Software ver 1.0 (Applied Biosystems).

3.2.2.3 Data analysis

The data were checked for the presence of null alleles and stuttering at all loci using the software program MICROCHECKER (Van Oosterhout et al., 2004). Number of alleles (*NA*), allele frequencies, observed (H_O) and expected heterozygosities (H_E) and fixation indices F_{IS}, F_{IT} and F_{ST} as well as deviations from Hardy-Weinberg equilibrium (HWE) were determined using GenAlEx ver 6.41 (Peakall and Smouse, 2006). These population parameters were calculated for each population for all loci and also across populations for each locus. Polymorphism information content (PIC) (Botstein et al., 1980) was determined for all loci using Microsatellite Toolkit ver 3.1.1 (Park, 2001). Allelic richness which is an estimation of the mean number of alleles per locus corrected by sample size, was estimated with FSTAT ver 2.9.3 software (Goudet, 2002).

Phylogenetic analysis was conducted with POPULATIONS ver 1.2.30 (Langella, 1999) using Nei's minimum genetic distance, D_m, with 1000 bootstraps over loci after which the results were viewed with TreeView (Page, 1996). To infer population structure, the STRUCTURE software ver 2.3 (Pritchard et al., 2000; Falush et al., 2003) was employed using the admixed ancestry and correlated allele frequency model. The program was run after burn-in period of 10,000 and Markov Chain Monte Carlo (MCMC) replication of 100,000. The number of putative populations, K, was set to range from 1 to 8 and number of iterations for each value of K was 50. No *a priori* population information was used in running the programme. Each individual was assigned to a cluster based on the highest likelihood of membership. Results from STRUCTURE were zipped and used as input file for the web-based programme STRUCTURE HARVESTER (Earl and vonHoldt, 2012) to calculate

posterior probability values, ln Pr(X/K). The best value of K or the most likely number of clusters was inferred from ΔK which was calculated by the method of Evanno et al. (2005). Results were then combined using the program CLUMPP ver 1.1.2 (Jacobsson and Rosenberg, 2007) which determines symmetric similarity coefficient between pairs of runs and averages individual membership proportions. Summary barplots were then generated using DISTRUCT ver 1.1 (Rosenberg, 2004).

To estimate the extent of exchange of individuals between populations, current migration was assessed using GENECLASS ver 2.0 (Piry et al., 2004). This program is able to detect migrants through a Bayesian method using multilocus genotype data (Rannala and Mountain, 1997). As a test statistics, it estimates likelihood L= L_home/L_max, where L_home is the likelihood that the individual's genotype belongs to the population from which the individual was sampled and L_max is the highest likelihood of the genotype in any population (Paetkau et al., 2004). The program was run with Monte-Carlo resampling algorithm and number of simulated individuals was pegged at 10,000 ($\alpha = 0.01$).

The rate and direction of migration among the populations were estimated using a Bayesian approach implemented in the program BAYESASS ver 3.0.3 (Wilson and Rannala, 2003). It was run with MCMC algorithm of 10,000,000 iterations sampled every 100 steps after 1,000,000 iterations were discarded as burn-in to ensure that the chain simulation reaches stationary distribution before sampling. The program was run three times with different seeds for each run to ensure concordance. The default seed of 10 was used initially followed by 100 and lastly 1000.

We tested for isolation by distance by determining the correlation between pairwise differentiation and geographic distance. Geographic distances between the locations (in km) were obtained by converting decimal degrees coordinates into distance matrix using Geographic Distance Matrix Generator ver 1.2.3 (Ersts, Internet). Pairwise geographic distances were log transformed and a Mantel test was conducted using IBDWS ver 3.14 (Jensen et al., 2005) to assess the correlation between geographic distance and pairwise genetic differentiation.

3.2.3 Results

3.2.3.1 Microsatellite characteristics

There were a total of 133 alleles ranging from 8 – 16 with a mean of 7.3 per locus across populations. Eleven specific alleles were found in Guinea Savanna, two in Forest, 10 in Coastal Savanna and nine in the Volta Region. H_O and H_E ranged from 0.460 for *Tsw09* to 0.847 for *Tsw16* and 0.452 for *Tws09* to 0.866 for *Tsw08*, respectively (Table 3.2.1). These results indicated that all markers were polymorphic. Apparently, *Tsw09* was the least polymorphic among the panel of markers used in this study. Allelic richness ranged from 5.499 for *Tsw02* to 11.435 for *Tsw08*. The range of *PIC* followed the same pattern as allelic richness (Table 3.2.1). Three loci deviated from HWE in the Volta Region population (*Tsw06, Tsw13 and Tsw19*), two loci deviated from HWE in each of Guinea Savanna (*Tsw09 and Tsw11*) and Forest (*Tsw11 and Tsw13*) populations whereas in the Coastal Savanna, only locus *Tsw08* deviated from HWE. Wright's fixation indices, F_{IS}, F_{IT} and F_{ST} for each locus over populations are listed in Table 3.2.1.

3.2.3.2 Genetic diversity and population structure

F_{IS} which indicates heterozygote deficiency and used as an indicator of the level of inbreeding, was found to be low (0.009 – 0.031) for all populations, indicating that the populations are not inbred (Table 3.2.2). Phylogenetic analysis revealed that Forest population is closer to the Coastal Savanna population than other populations whilst Volta Region population is closer to the Guinea Savanna population than other populations (Fig 3.2.1). These results are supported by fairly high bootstrap value (90%). Pairwise F_{ST} values however indicated that all populations were significantly differentiated ($p < 0.01$) after Bonferroni correction for multiple testing (Rice, 1989) (Table 3.2.3).

STRUCTURE clustering analysis showed a more detailed structure of the populations (Fig 3.2.2). At K = 2, the pattern of clustering was similar to the clustering observed in the Neighbour-joining tree. However, at K = 3, the Volta Region population split from the Guinea Savanna population whilst the Forest and the Coastal Savanna populations remained together. Interestingly, no clear structure was observed at K = 4. Moreover, using the method developed by Evanno et al. (2005), the highest ΔK was found at K = 3 suggesting that there were three distinct clusters.

3.2.3.3 Migration rate

Four migrants were detected in the dataset by GENECLASS including two females, one male and one individual of unknown sex. One individual in the Volta population was detected as a migrant from the Guinea Savanna, two were detected in the Coastal Savanna (one from Volta and the other from Forest) and one individual detected in the Forest originating from Coastal Savanna. The rates of migration between populations are presented in Table 3.2.4. Except for migration rate between Coastal Savanna and Forest, none of the population pairs recorded migration rate that was significantly different from zero.

3.2.4 Discussion

In this study we found significant differentiation among all pairs of populations which is not entirely in agreement with mitochondrial D-loop results that we reported previously (Adenyo et al., 2013, Chapter 3.1). In our previous results, we found that the Guinea Savanna and the Coastal Savanna populations are essentially one Savanna population because we did not find any significant differentiation between them whereas in the current study we found significant differentiation between them ($p < 0.01$). However, since mitochondrial DNA is maternally inherited, our previous results represented only meternal gene flow (uni-parental). The current results show that Coastal Savanna is rather closer to Forest than to Guinea Savanna. This probably reflects effective dispersal of females even though grasscutters like most mammals are expected to exhibit male-biased dispersal. This can be seen from the perspective of the breeding structure of the grasscutter which usually consists of one male with multiple females and their offspring. Even though the mechanisms of dispersal are not well understood and therefore need thorough investigation, we can deduce from our results that long distance female-mediated gene flow is possible in the grasscutter. This could also be aided by their high prolificacy (litter size of up to 15) (Adu et al., 1999) which makes it less likely that all females would stay in their natal groups. We found that genetic diversity within Guinea Savanna is slightly higher compared to other populations, which is in agreement with our previous mitochondrial results (Adenyo et al., 2013, Chapter 3.1).

The disparity or discordance between mitochondrial and microsatellites which is defined as significant difference in the patterns of differentiation between the two markers can be attributed to sex-biased dispersal or some demographic events (e.g. genetic drift) that could cause biased change of either of the markers (Toews and Brelsford, 2012) since both markers

have different mutation rates. The mutation rate of microsatellites is higher than that of mitochondrial which could explain why the microsatellites could detect the differentiation between Guinea Savanna and Coastal Savanna but the D-loop could not detect it in our previous results. This situation is not unique to the grasscutter but has been reported in some other rodents such as field voles, common voles, Dalton's mouse, Douglas red squirrels and American red squirrels (Braaker and Heckel, 2009; Bryja et al., 2010; Beysard et al., 2011; Chavez et al., 2011).

The fact that we detected four migrants indicates that the populations are not entirely isolated but exchange individuals, albeit at a minimal rate. Both male and female migrants were detected, which suggests that not only males but also females disperse from their natal groups. Even though the direction of migration cannot be ascertained because the migration rates were not significant (except Coastal Savanna-Forest), a cursory look at the migrant individuals detected, shows that Forest and Guinea Savanna exchange individuals in either direction whilst one direction (Guinea Savanna-Volta Region) was observed between Guinea Savanna and Volta. Interestingly, no migrant was detected in the Guinea Savanna population. This suggests a north-south migration but will require further data to elucidate. A north-south dispersal may be more likely as conditions in the north are harsh and feed resources are less compared to the south. This result indicates gene flow among the different populations which agrees with previous mitochondrial results where we detected two common haplotypes among the populations (Adenyo et al., 2013, Chapter 3.1). We considered the possibility of human-mediated gene flow, in other words, transfer of grasscutters among regions by humans. However, this may not be the case because it is less likely that captive grasscutters transferred from one region would be released into the wild.

The splitting of Volta population from the Guinea Savanna population is interesting and noteworthy, as it suggests the effect of human activity on a natural population. Water bodies such as rivers are known to form barriers that limit gene flow. For instance, it has been shown that a river served as an effective barrier to gene flow in European ground squirrels in Serbia (Ćosić et al., 2013). Also, Bryja et al. (2010) found that rivers in West Africa form effective barriers to gene flow of Dalton's mouse. Similar observations were previously made by Nicolas et al. (2008) for *Praomys sp* and Dobigny et al. (2005) for *Taterillus sp* in the West African sub-region. The extent of rivers forming barriers to dispersal depends to some extent on the swimming ability of the species in question. The effect of the Volta River was clearly

evident in the study of phylogeography of multimammate mouse by Brouat et al. (2009). The above studies indicate that rivers can be effective barriers to dispersal irrespective of the swimming ability of the species.

The Guinea Savanna and the Volta populations of grasscutter could be considered as one population in the recent past. However, in 1962, a hydro-electric power plant was constructed at Akosombo on the Volta River to serve the populace and budding industries in Ghana. As a result, a large man-made lake (largest in the world), spanning about 400 km in length and 21 km wide was created. This probably had some impact on fauna populations by hindering gene flow across the lake. Even though grasscutters are good swimmers (Opara, 2010), we do not expect that they can swim across such a wide lake. We speculate that this situation may not affect only grasscutters but also many other mammals that inhabit Ghana. This may not be significant for grasscutter populations because they are very prolific and agile. However, there could be profound effects of such separation on the population viability of large mammals and especially primates which have low reproduction rate. The effects of construction of such dams on fauna populations should therefore be considered when embarking on such efforts in the future in order to safeguard populations from extinction.

No tendency of isolation-by-distance was found among the grasscutter populations, which is in agreement with mitochondrial results. This is probably due to limited number of sampling localities. More extensive sampling may show otherwise, because presence of isolation-by-distance has been shown in other rodents on a regional scale (Nicolas et al., 2008; Bryja et al., 2010). It will be interesting to scale up this study by including samples from different places in the species geographic range to get a better view of the grasscutter's phylogeography.

In this study, we have shown one way of applicability of the novel microsatellite markers developed for the grasscutter. We have shown that the markers can be used to address ecological questions such as genetic diversity, population structure and dispersal of the grasscutters in the wild. We observed discordance between our previous results (mtDNA) and the current results (microsatellites), therefore we propose that further investigation be carried out to find the actual cause(s) of this disparity. As far as genetics of the grasscutter is concerned, this is the first time genetic markers have been used to study the populations in Ghana and of course the first of its kind in the species range. These results will form a

baseline from which inferences can be made for comparison in future phylogeographic studies on the grasscutter. As we advance the course of domestication, there will be the need to determine the effect of domestication on the genetic structure of the grasscutter by comparing the domestic populations to their wild counterparts or finding genetically diverse individuals. The markers will be invaluable tools that will be used to do such an assessment.

Table 3.2.1 Profile of 12 microsatellite loci

Locus	N_A	H_O	H_E	F_{IS}	F_{IT}	F_{ST}	**Allelic richness**	***PIC***
Tsw02	9	0.571	0.557	-0.025	0.004	0.028	5.499	0.549
Tsw03	12	0.786	0.801	0.019	0.078	0.060	7.883	0.835
Tsw06	10	0.718	0.776	0.075	0.117	0.046	7.106	0.791
Tsw07	8	0.710	0.729	0.026	0.124	0.100	6.167	0.786
Tsw08	16	0.802	0.866	0.074	0.117	0.047	11.435	0.902
Tsw09	8	0.460	0.452	-0.017	0.320	0.332	5.517	0.633
Tsw11	14	0.691	0.826	0.163	0.229	0.079	9.968	0.889
Tsw13	8	0.747	0.709	-0.055	0.020	0.071	6.290	0.733
Tsw16	12	0.847	0.819	-0.034	0.005	0.038	8.213	0.834
Tsw19	11	0.805	0.782	-0.029	0.028	0.056	7.729	0.811
Tsw21	13	0.828	0.827	-0.001	0.042	0.043	8.574	0.852
Tsw23	12	0.803	0.794	-0.012	0.039	0.050	8.046	0.817
Mean	7.3	0.731	0.745	0.015	0.094	0.079		
SE	0.3	0.023	0.020	0.018	0.028	0.024		

N_A: number of alleles, H_O: observed heterozygosity, H_E: expected heterozygosity, F_{IS}: fixation coefficient of an individual within a subpopulation, F_{IT}: fixation coefficient of an individual within the total population, F_{ST}: fixation coefficient of a subpopulation within the total population, *PIC*: polymorphic information content.

Table 3.2.2 Characteristics of grasscutter populations

Population	***n***	MN_A **± SE**	N_E **± SE**	**Specific alleles**	H_O **± SE**	H_E **± SE**	F_{IS} **± SE**
Guinea Savanna	19	8.333 ± 0.620	5.250 ± 0.444	11	0.780 ± 0.042	0.789 ± 0.024	0.009 ± 0.047
Forest	16	7.167 ± 0.613	4.735 ± 0.438	2	0.735 ± 0.033	0.760 ± 0.030	0.031 ± 0.031
Coastal Savanna	16	7.417 ± 0.596	4.796 ± 0.542	10	0.729 ± 0.054	0.740 ± 0.047	0.009 ± 0.043
Volta Region	15	6.500 ± 0.571	3.914 ± 0.451	9	0.679 ± 0.053	0.690 ± 0.050	0.007 ± 0.043

MN_A: number of alleles, N_E: effective number of alleles, H_O: observed heterozygosity, H_E: expected heterozygosity, F_{IS}: fixation coefficient of an individual within a subpopulation

Table 3.2.3 Matrix of pairwise F_{ST} (below diagonal) and Nei's genetic distance (above diagonal)

	Guinea Savanna	Forest	Coastal Savanna	Volta Region
Guinea Savanna		0.253	0.234	0.203
Forest	0.043		0.118	0.411
Coastal Savanna	0.044	0.032		0.462
Volta Region	0.047	0.075	0.087	

Table 3.2.4 Mean (95% confidence interval) migration rates among populations of grasscutter

		Migration from:			
		Guinea Savanna	Forest	Coastal Savanna	Volta Region
Migration into:	Guinea Savanna	0.7678	0.0178 (-0.015 – 0.051)	0.0377 (-0.029 – 0.105)	0.1767 (0.000 – 0.346)
	Forest	0.0145 (-0.013 – 0.042)	0.7010	**0.2650 (0.206 – 0.324)**	0.0196 (-0.015 – 0.055)
	Coastal Savanna	0.0297 (-0.021 – 0.081)	0.0409 (0.023 – 0.105)	0.9057	0.0237 (-0.018 – 0.066)
	Volta Region	0.0302 (-0.023 – 0.083)	0.0253 (-0.018 – 0.068)	0.0220 (-0.019 – 0.063)	0.9226

Columns correspond to populations from which individuals migrated whilst rows correspond to populations from which individuals were sampled. Values in parentheses indicate 95% confidence interval (CI). The value in bold shows 95% CI that is significantly different from zero. Values along the diagonal indicate the proportion of individuals derived from the source population.

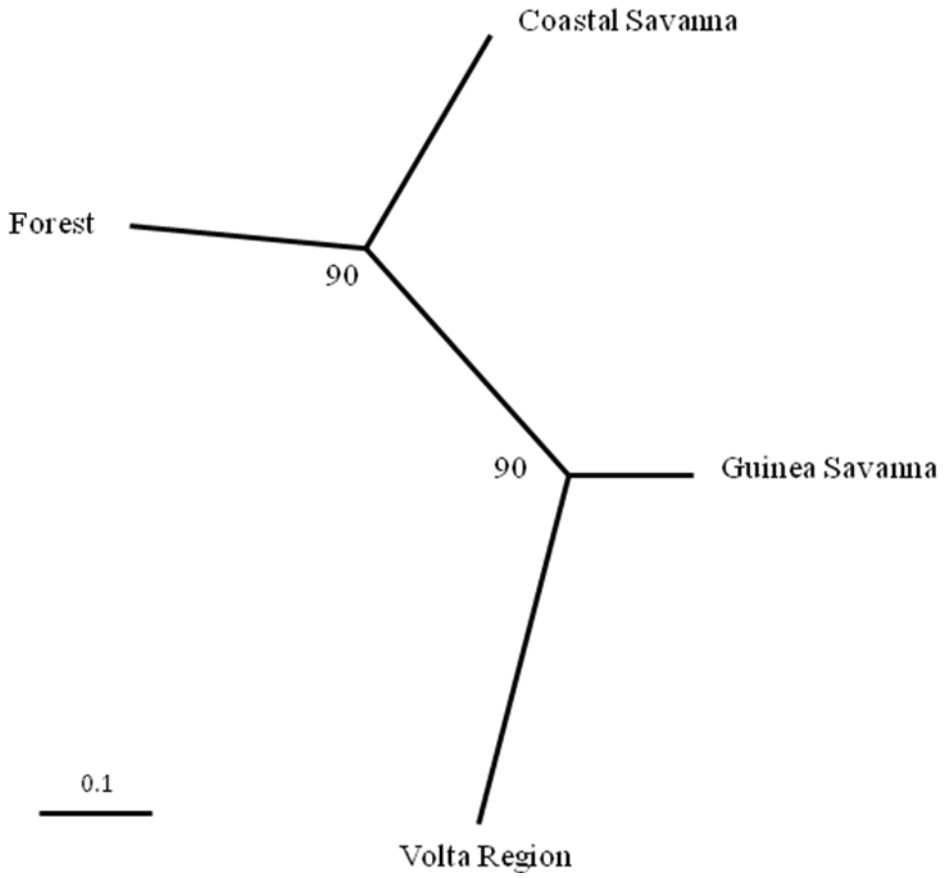

Fig 3.2.1 Neighbour-joining tree of grasscutter populations in Ghana based on Nei's minimum genetic distance, D_m. Bootstrap values are shown (1000 bootstraps were conducted over loci).

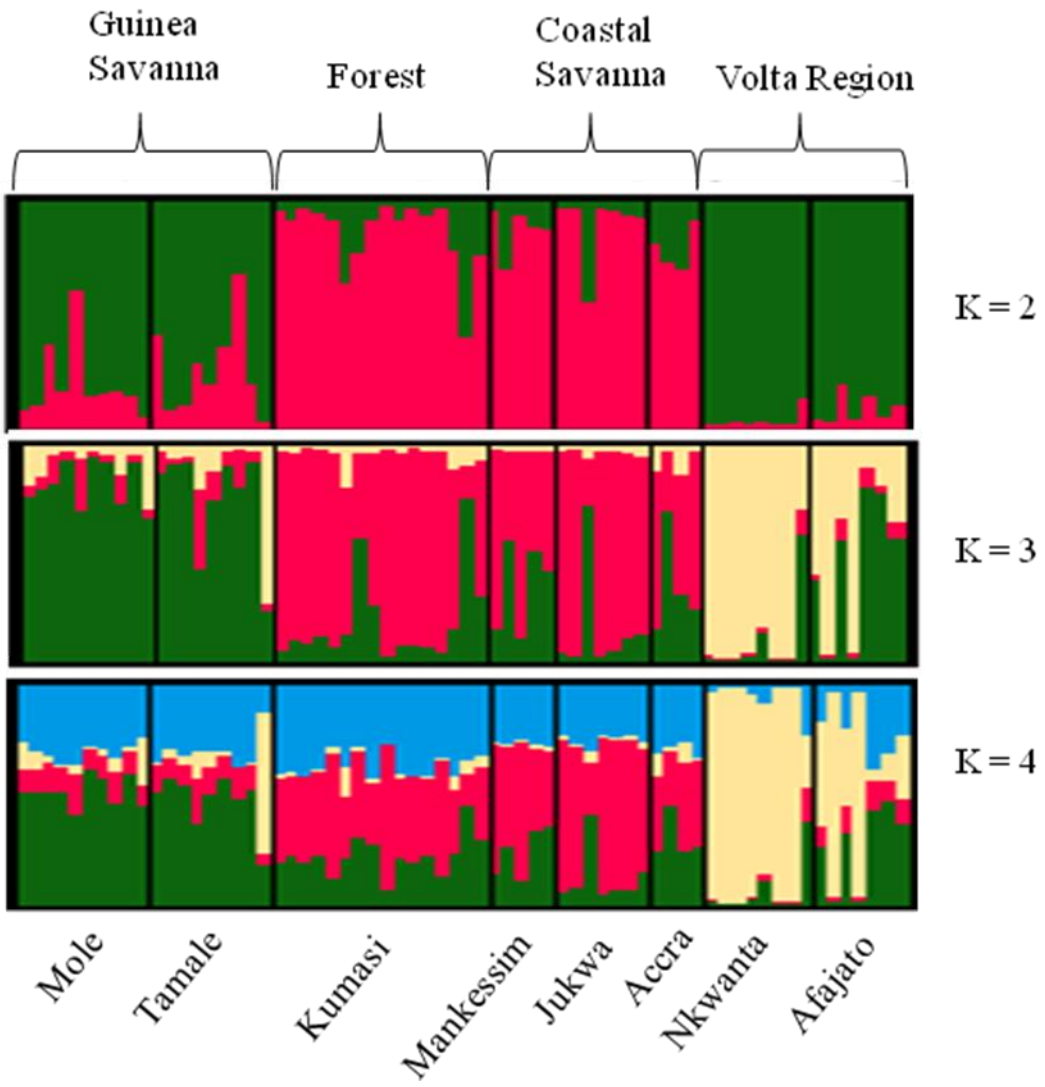

Fig 3.2.2 Structure clustering of grasscutter populations in Ghana. Sampling localities are shown below (separated by black vertical lines) while populations are indicated above the diagram. Each colour represents a cluster and individuals in each cluster are represented by bars.

4. GENERAL DISCUSSION AND FINAL REMARKS

Genetic variation is required for populations to remain viable in the face of natural selection. It has been exploited by geneticists for selection and genetic improvement of many species. Breeders have made tremendous gains in livestock species because genetic variation exists both within and between populations. Quantifying variation is a basic element of animal breeding programs. With the advent of DNA techniques, the field of animal breeding has had unprecedented growth. Genetic markers such as microsatellites and single nucleotide polymorphisms (SNPs) have been used extensively to make genetic gains through marker assisted selection (Eggen, 2012). The grasscutter which is being promoted as a protein resource has the potential to alleviate poverty in Sub-Saharan Africa especially in rural and peri-urban areas. Compared to conventional livestock such as sheep, goat and cattle, it has less space requirement and can easily be kept in the backyard. However, there are challenges regarding the grasscutter rearing chief of which is high mortality due to aggression. Selection for docility and other productive traits is therefore necessary for the promotion of domestication. In this study, I have developed markers and quantified the genetic variation in grasscutter populations in Ghana which can be used as basis for genetic improvement.

4.1 Microsatellite marker development

In order to develop the markers, partial genome data were screened for microsatellite repeats. In most vertebrate genomes, dinucleotide repeats are the most frequent followed by tetranucleotide repeats (Toth et al., 2000). Trirepeats and hexarepeats are usually less abundant compared to direpeats and tetrarepeats because they appear mostly in coding regions, which supposes that they may be part of genes (Toth et al., 2000). The results presented in this study indicate that the nucleotide repeat structure of the grasscutter follows a similar pattern as most vertebrates. However, apart from AC repeats which were the most abundant, the repeat types for each repeat class were different from other rodents as reported by Toth et al. (2000).

Comparing the chromosome numbers and the genome size of related species such as mouse (2n = 40 with estimated genome size of 2.75 Gb), rat (2n = 42 with estimated genome size of 2.75 Gb) and guinea pig (n = 64 with estimated genome size of 3.4 Gb) (Worley and Gunaratne, 2006; Hubbard et al, 2007; Eppig et al., 2012), an estimated genome size of 3.0 Gb for the grasscutter (2n = 44) (Deuve et al., 2008) is plausible. This genome size will

require approximately 300 markers in order to create an interval map of 10 cM (1 cM = 1 Mb). Through this study, 116 microsatellite markers have been developed, which is about one-third of the 300 estimated markers. These markers will be useful for linkage analysis and subsequent marker assisted selection as we advance the course of the grasscutter domestication. In this study several pairs of loci were found to be in linkage disequilibrium (LD), which is a good signal that they will be very useful for linkage analysis.

Genome comparative mapping of model organisms such as the rat, mouse and other mammals has provided insights into genome architecture and evolution (Copeland et al., 1993; Barendse et al., 1994). Comparative mapping has led to the understanding of the genetic basis (causal mutations) of certain diseases and genetic disorders especially in humans since identification of such mutations is easier in model organisms than in humans (Copeland et al., 1993). Genetic and physical mapping of the grasscutter genome will benefit immensely from model species such as rat and mouse whose genomes have been extensively studied. The microsatellite markers developed in this study may help us to map disease loci in the future since the grasscutter was found to be a suitable model to study a human disease (Addo et al., 2007).

The markers may also serve as effective tools to understand the phylogeography, dispersal patterns and genetic structure of wild grasscutter populations. There remains a lot to be known about the grasscutter ecology. Even though the grasscutter is not an endangered species, understanding of the population dynamics may help to monitor them for future conservation. Population level studies across the species range may identify unique populations that will inform conservation decisions in the future if the need arises. Microsatellite analysis of groups in the wild will help us to understand the social structure of the grasscutter which will be useful for the management of the grasscutters kept under domestic conditions.

4.2 Genetic diversity of grasscutter in Ghana

In this study, I determined genetic diversity of grasscutter populations in Ghana using mitochondrial and microsatellite markers developed *de novo*. This was done in order to identify populations from which individuals can be selected to set up a reference family for linkage analysis. I found high mitochondrial haplotype diversity across zones (0.853 – 0.978) with Guinea Savanna having the highest diversity which also agrees with the microsatellite

results where mean number of alleles (MN_A) was found to be considerably high (6.5 – 8.3). This could be due to the very high populations of grasscutter that presently exist or probably existed in the recent past in the country. This is supported by very low level of inbreeding (measured by F_{IS}) detected in all the populations in the microsatellite analysis (0.007 – 0.031). This in essence indicates that there is a lot of genetic potential which could be exploited for genetic improvement. The implication of these results for our purpose is that we can select individuals from any population to set up a reference population which will be valuable for studying the effect of domestication. The Forest population was found to be significantly differentiated from both Guinea and Coastal Savanna populations indicating two main populations of grasscutter in Ghana. The microsatellite results however showed significant differentiation among all populations. This discordance could be a population effect or marker effect. Since mitochondrial DNA is maternally inherited, it is possible that the difference observed in the two markers is due to dispersal of females from the Guinea Savanna to the Coastal Savanna or vice versa. The former is more plausible because of the harsh conditions in the Guinea Savanna zone. This is also supported by the fact that two common mitochondrial haplotypes were found among the Guinea Savanna, Forest and Coastal Savanna populations. It also suggests that there is to some extent gene flow among the various populations. A similar trend was observed with the microsatellite markers where at least one migrant from one of the studied populations was found in each population except the Guinea Savanna zone. Alternatively, the difference in the results of the two markers could be attributed to the fact that mitochondrial markers have slower mutation rate compared to microsatellite markers. It seems from the microsatellite results that the genetic differentiation reflects more of geographic distance even though a test of isolation-by-distance was not significant. This is probably due to limited sampling locations involved in this study.

Furthermore, a test of population neutrality using mitochondrial D-loop indicated a past expansion event in the Guinea Savanna population whilst the other populations have been stable over the years. The reason why this expansion occurred could be the abundance of feed resources in the Guinea Savanna zone in the past when the weather conditions were not as severe as they are at present. The most preferred feed resources of grasscutter are guinea grass and elephant grass (Ajayi and Tewe, 1980; Adu et al., 1999) which are savanna species, supporting the fact that the grasscutter is predominantly a savanna species. As feed resources are depleted due to harsh conditions, it is usual that grasscutters would disperse to regions

where preferred feed resources are available. It is therefore not surprising that they invade cultivated areas and are considered as pests in such areas.

In the microsatellite analysis, it was found that the Volta population was probably originally part of the Guinea Savanna but a recent split occurred which could be attributed to the damming of the Volta River for hydro-electric power generation, making it an effective barrier to dispersal of grasscutters. Documentary evidence exists to show that the Volta River is indeed a barrier to dispersal of a multimammate mouse (Brouat et al., 2009). This is speculated to be the case for many other mammals inhabiting Ghana but will require further studies to prove. These findings provide a glimpse into the population history, dispersal, diversity and structure of grasscutter populations for the first time in Ghana.

4.3 Future perspective

This study serves as a first step in a long term study of the grasscutter genome which will involve genetic and possibly physical mapping. The repeat structure of the grasscutter genome has been characterized using partial genome data and sets the foundation upon which other genomic studies on grasscutter may be based. The long term plan is to be able to map genes that underlie traits of economic importance to enhance selection and domestication through marker assisted selection. Marker assisted selection has the advantage of faster genetic gain when combined with traditional selection by increasing selection differential through higher accuracy of selection (Van Arendonk et al., 1994). Markers may be in linkage disequilibrium with quantitative trait loci (QTL) and therefore can be used to effectively enhance genetic improvement programs (Dekkers, 2004). In genetic evaluation of livestock for instance, markers could be added to best linear unbiased prediction (BLUP) models as fixed or random effects to estimate breeding values with the potential of increasing the accuracy of the estimated breeding values (Van Arendonk et al., 1994). Marker assisted selection is even more advantageous for traits that have low heritability, difficult to measure or can only be measured later in life or after the animal is slaughtered (e.g. carcass traits) (Dekkers, 2004).

Many markers have been developed which will serve as invaluable tools not only for genetic improvement but also for phylogeographic and genetic diversity studies on grasscutter. Due to declining cost of whole genome sequencing, it is envisaged that in the not too distant future, whole genome assembly may be carried out. Then other markers such as

SNPs can be developed to make marker rich maps to facilitate gene discovery especially when the grasscutter is considered not only as a micro-livestock but also as a model species. One other approach that could be used is restriction site associated DNA sequencing (RAD). This involves digesting genomic DNA with restriction endonucleases and sequencing regions surrounding all restriction sites using next generation sequencing approach (Davey et al., 2011). It is a cost effective way of developing thousands of SNPs and effectively creating linkage groups and identifying QTLs (Baird et al., 2008; Hohenhole et al., 2010; Davey et al., 2011). For instance we could identify extreme phenotypes (e.g. docile and aggressive) and cross them to get F1 individuals. We could then apply RAD sequencing to identify SNPs that may be associated with tameness and aggression. Similar studies have been conducted in mouse and rat. For instance Brodkin et al. (2002) reported two significant QTLs for aggression in mice; one on the distal end of chromosome 10 and the other on the proximal end of chromosome X. The authors suggested two candidate genes, diacylglycerol kinase α subnit gene (*Dagk1*) and glutamate receptor subunit AMPA3 gene (*Gria3*), as underlying genes that are responsible for aggressive behaviour in mice. A similar study was conducted in rats where two lines selected over several generations for aggression and tameness were analysed and chromosomal region responsible for such traits were fine mapped (Albert et al., 2009; Albert et al., 2011)

Finally, a new paradigm shift in animal breeding is genomic selection where selection is based on genome-wide estimated breeding values (GEBV) using tens of thousands of SNPs without prior knowledge of their location in the genome (Meuwissen et al., 2001; Dekkers 2004; Eggen, 2012). With the plummeting cost of whole genome sequencing and tremendous increase in computational methods capable of handling huge data sets, the use of genomic selection will become a common practice which will be applied to all livestock species including the grasscutter which may be recognised as a livestock species in the not too distant future. In a survey conducted during 2008 – 2009 by Ahenkan and Boon (2011) in Bibiani-Bekwai and Sefwi Wiawso districts in the Western Region of Ghana, it was reported that 39.8% of respondents (households) were engaged in grasscutter rearing. This is a clear indication that more and more people are getting interested in grasscutter rearing in Ghana. The grasscutter is gaining recognition as a micro-livestock with the potential to alleviate poverty and to increase protein supply in Sub-Saharan Africa. It is not so surprising that it attracted many interest groups, donor agencies and development partners such as German

Technical Cooperation (GTZ), ActionAid, Japan International Cooperation Agency (JICA), Heifer International and many others. Their continuous support is making it possible for many people to be involved in grasscutter rearing to enhance its domestication.

5. REFERENCES

Abyankar A, Park H-B, Tonolo G, Luthman H (2009) Comparative sequence analysis of the non-protein-coding mitochondrial DNA of inbred rat strains. *PLoS ONE,* **4**, e8148.

Addo P, Adu-Addai B, Quartey M, Abbas M, Okang I, Owusu E, Ofori-Adjei D, Awumbila B (2007) Clinical and histopathological presentation of Buruli ulcer in experimentally infected grasscutters (*Thryonomys swinderianus*). *Internet Journal of Tropical Medicine*, **3**, e2.

Addo P, Dodoo A, Adjei S, Awumbila B, Awotwi E (2002) Determination of the ovulatory mechanism of the grasscutter. *Animal Reproduction Science*, **71**, 125 – 137.

Adenyo C, Hayano A, Inoue E, Kayang BB, Inoue-Murayama M (2012) Development of microsatellite markers for grasscutter (*Thryonomys swinderianus*, RODENTIA) using next-generation sequencing technology. *Conservation Genetic Resources,* **4**, 1011-1014.

Adenyo C, Hayano A, Inoue E, Kayang BB, Owusu EH, Inoue-Murayama M (2013) Mitochondrial D-loop diversity of grasscutter (*Thryonomys swinderianus*, RODENTIA, HYSTRICOMORPHA) in Ghana. *Open Journal of Animal Sciences* **3**, 145 – 153.

Adu EK (2003) Patterns of parturition and mortality in weaned greater cane rats (*Thryonomys swinderianus*, Temminck). *Tropical Animal Health and Production,* **35**, 425 – 431.

Adu EK, Alhassan WS, Nelson FS (1999) Smallholder farming of the greater cane rat, *Thryonomys swinderianus,* Temminck, in southern Ghana: a baseline survey of management practices. *Tropical Animal Health and Production,* **31**, 223 – 232.

Adu EK, Yeboah S (2000) The efficacy of the vaginal plug formation after mating for pregnancy diagnosis, and embryonic resorption in utero in the greater cane rat (*Thryonomys swinderianus*, Temminck). *Tropical Animal Health and Production*, **32**, 1 – 10.

Adu EK, Yeboah S (2000) The efficacy of the vaginal plug formation after mating for pregnancy diagnosis, and embryonic resorption in utero in the greater cane rat

(*Thryonomys swinderianus*, Temminck). *Tropical Animal Health and Production*, **32**, 1 – 10.

Ahenkan A, Boon E (2011) Improving nutrition and health through non-timber forest products in Ghana. *Journal of Health, Population and Nutrtition*, **29**, 141 – 148.

Ajayi SS, Tewe OO (1980) Food preferences and carcass composition of the grasscutter (*Thryonomys swinderianus*) in captivity. *African Journal of Ecology*, **18**, 133 – 140.

Akramov K, Asante F (2008) Decentralization and local public services in Ghana: Do geography and ethnic diversity matter? [cited 2014/01/20]. Available from www.ifpri.org/sites/default/files/publications/gsspwp16.pdf.

Albert FW, Carlborg Ö, Plyusnina I, Besnier F, Hedwig D, Lautenschläger S, Lorenz D, McIntosh J, Neumann C, Richter H, Zeising C, Kohzhemyakina R, Shschepina O, Kratzsch J, Trut L, Teupser D, Thiery J, Schönenberg T, Andersson L, Pääbo S (2009) Genetic architecture of tameness in a rat model of animal domestication. *Genetics*, **182**, 541 – 554.

Albert FW, Hodges E, Jensen JD, Besnier F, Xuan Z, Rooks M, Bhattacharjee A, Brizuela L, Good JM, Green RE, Burbano HA, Plyusnina IZ, Trut L, Andersson L, Schönenberg T, Carlborg Ö, Hannon GJ, Pääbo S (2011) Targeted resequencing of a genomic region influencing tameness and aggression reveals multiple signals of positive selection. *Heredity*, **107**, 205 – 214.

Annor SK, Adu EK, Donkor J, Otsyina HR, Ahiaba J (2009) Grasscutter production: A handbook. GTZ/MOAP, Accra.

Annor SY, Ahunu BK, Aboagye GS, Boa-Amponsem K, Djang-Fordjour KT, Cassady JP (2011) The genetics of docility of the grasscutter (*Thryonomys swinderianus*). *Livestock Research for Rural Development*, Volume 23, Article #176. [cited 2013/11/11]. Available from http://www.lrrd.org/lrrd23/8/anno23176.htm.

Annor SY, Ahunu BK., Aboagye GS, Boa-Amponsem K, Djang-Fordjour KT, Cassady, JP (2011) The genetics of morphological traits in the grasscutter (*Thryonomys swinderianus*). *Livestock Research for Rural Development*, Volume 23, Article #16. [cited 2013/11/11]. Available from http://www.lrrd.org/lrrd23/8/Anno23167.htm.

Annor SY, Kagya-Agyemang JK, Abbam JEY, Oppong SK and Agoe IM (2008) Growth performance of grasscutter (*Thryonomys swinderianus*) eating leaf and stem fractions of Guinea grass (*Panicum maximum*). *Livestock Research for Rural Development* Volume 20, Article #125. [cited 2011/10/25]. Available from http://www.lrrd.org/lrrd20/8/anno20125.htm.

Asibey EOA (1974) Wildlife as a source of protein in Africa south of the Sahara. *Biological Conservation*, **6**, 32 – 39.

Asibey EOA (1981) Maternal and neo-natal weight in the grasscutter, *Thryonomys swinderianus* (Temminck) in Ghana. *African Journal of Ecology*, **19**, 355 – 360.

Asibey EOA, Child GS (1990) Wildlife management for rural development in sub-Saharan Africa. *Unasylva*, #161. [cited 2014/01/20]. Available from www.befac.net/pdf/t044_asibey_1990_unasylva.pdf.

Assogbadjo AE, Codjia JTC, Sinsin B, Ekue MRM, Mensah GA (2005) Importance of rodents as a human food source in Benin. *Belgian Journal of Zoology*, **135**, 11 – 15.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific Phylogeography: The Mitochondrial DNA Bridge between Population Genetics and Systematics. *Annual Review of Ecology and Systematics*, **18**, 489 – 522.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE, **3**: e3376.

Barendse W, Armitage SM, Kossarek LM, Shalom A, Kirkpatrick BW, Ryan AM, Clayton D, Li L, Neibergs HL, Zhang N, Grosse WM, Weiss J, Creighton P, McCarthy F, Ron M, Teale AJ, Fries R, McGraw RA, Moore SS, Georges M, Soller M, Womack JE, Hetzel DJS (1994) A genetic linkage map of the bovine genome. *Nature Genetics*, **6**, 227.

Bennett-Lartey SO, Ayerno GS, Markwei CM, Asante IK, Abbiw DK, Boateng SK, Anchirinah, VM, Ekpe P (2002) Contribution of home gardens to *in situ* conservation

of plant genetic resources farming systems in Ghana. *Proceedings of the Second International Home Gardens Workshop*, 83 – 96.

Beysard M, Perrin N, Jaarola M, Heckel G, Vogel P (2011) Asymmetric and differential gene introgression at a contact zone between two highly divergent lineages of field voles (*Microtus agrestis*). *Journal of Evolutionary Biology*, **25**, 400 – 408.

Botchway PK, Adenyo C, Kayang BB, Hayano A, Inoue-Murayama M (2013) Development of 31 polymorphic microsatellite markers for guinea fowl (*Numida meleagris*) using next-generation sequencing technology. *Conseravation Genetics Resources*, doi 10.1007/s12686-013-9991-8.

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map using restriction fragment length polymorphisms. *American Journal of Human Genetics*, **32**, 314 – 331.

Braaker S, Heckel G (2009) Transalpine colonization and partial phylogeographic erosion by dispersal in the common vole (*Microtus arvalis*). *Molecular Ecology*, **18**, 2518 – 2531.

Brodkin ES, Goforth SA, Keene AH, Fossela JA, Silver LM (2002) Identification of quantitative trait loci that affect aggressive behaviour in mice. *The Journal of Neuroscience*, **22**, 1165 – 1170.

Bromham L, Rambaut A, Harvey PH (1996) Determinants of rate variation in mammalian DNA sequence evolution. *Journal of Molecular Evolution*, **43**, 610 – 621.

Brouat C, Tatard C, Ba K, Cosson J-F, Dobigny G, Fichet-Calvet E, Granjon L, Lecompte E, Loiseau A, Mouline K, Piry S, Duplantier J-M (2009) Phylogeography of the Guinea multimammate mouse (*Mastomys erythroleucus*): a case study for Sahelian species in West Africa. *Journal of Biogeography*, **36**, 2237 – 2250.

Bryja J, Granjon L, Dobigny G, Patzenhauerová H, Konečný A, Duplantier JM, Gauthier P, Colyn M, Durnez L, Lalis A, Nicolas V (2010) Plio-Pleistocene history of West African Sudanian savanna and the phylogeography of the *Praomys daltoni* complex (Rodentia): the environment/geography/genetic interplay. *Molecular Ecology*, **19**, 4783 – 4799.

Buschiazzo E, Gemmel NJ (2006) The rise, fall and renaissance of microsatellites in eukaryotic genomes. *Bioessays*, **28**, 1040 – 1050.

Chambers GK, MacAvoy ES (2000) Microsatellites: consensus and controversy. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, **126**, 455 – 476.

Chapuis M-P, Estoup A (2007) Microsatellites null alleles and estimation of population differentiation. *Molecular Biology Evolution*, **24**, 621 – 631.

Chavez AS, Saltzberg CJ Kenagy GJ (2011) Genetic and phenotypic variation across a hybrid zone between ecologically divergent tree squirrels (*Tamiasciurus*). *Molecular Ecology*, **20**, 3350 – 3366.

Chen JZ, Herbert PDN (1999) Intraindividual sequence diversity and a hierarchical approach to the study of mutations. *Mutation Research*, **434**, 205 – 217.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657 – 1660.

Clutton-Brock TH (1989) Female transfer and inbreeding avoidance in social mammals. *Nature*, **337**, 70 – 72.

Colbert J, Danchin E, Dhont AA, Nichols JD (2001) Dispersal. Oxford University Press, New York.

Copeland NG, Jenkins NA, Gilbert DJ, Eppig JT, Maltais LJ, Miller JC, Dietrich WF, Weaver A, Lincoln SE, Steen RG, Stein LD, Nadeau JH, and Lander ES (1993) A genetic linkage map of the mouse: current application and future prospects. Science, **262**, 57 – 66.

Ćosić N, Říčanová S, Bryja J, Penezić A, Cirović D (2013) Do rivers and human-induced habitat fragmentation affect genetic diversity and population structure of the European ground squirrel at the edge of its Pannonian range? *Conservation Genetics*, **14**, 345 – 354.

Cowlishaw G, Mendelson S, Rowcliff JM (2005) Structure and operation of a bushmeat commodity chain in southwestern Ghana. *Conservation Biology*, **19**, 139 – 149.

Cox JM, Marinier SL, Alexander AJ (1988) Auditory communication in the cane rat (*Thryonomys swinderianus*). *Journal of Zoology, London*, **216**, 141 – 167.

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Review Genetics*, **12**, 499 – 510.

Dekkers JC (2004) Commercial application of marker- and gene-assisted selection in livestock: Strategies and lessons. *Journal of Animal Science*, **82**: E313 – E328.

Deuve JL, Bennett NC, Britton-Davi J (2008) Chromosomal phylogeny and evolution of the African mole-rats (Bathyergidae). *Chromosome Research*, **16**, 57 – 74.

Dietrich WF, Miller JC, Steen RG, Merchant M, Damron D, Nahf R, Gross A, Joyce DC, Wessel M, Dredge RD, Marques A, Stein LD, Goodman N, Page DC, Lander ES (1994) A genetic map of the mouse with 4,006 simple sequence length polymorphisms. *Nature Genetics*, **7**, 220 – 245.

DiPetrillo K, Wang X, Stylianou IM, Paigen B (2005) Bioinformatics toolbox for narrowing rodent quantitative trait loci. *Trends in Genetics*, **21**, 683 – 692.

Dobigny G, Aniskin V, Granjon L, Cornette R, Volobouev V (2005) Very recent radiation in West African Taterillus: the concerted role of chromosome and climatic changes. *Heredity*, **95**, 358 – 368.

Drinkwater NR, Gould MN (2012) The Long Path from QTL to Gene. *PLoS Genetic*, **8**, e1002975.

Dwinell MR, Worthey EA, Shimoyama M, Bakir-Gungor B, DePons J, et al. (2009) The Rat Genome Database 2009: variation, ontologies,and pathways. *Nucleic Acids Research*, **37**, D744 – D749.

Earl DA and vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetic Resources*, **5**, 359 – 361.

Eggen A (2012) The development and application of genomic selection as a new breeding paradigm. Animal Frontiers, **2**, 1 – 15.

Eppig JT, Blake JA, Bult CJ, Kadin JA, Richardson JE (2012) The Mouse Genome Database Group 2012. The Mouse Genome Database (MGD): comprehensive resource for genetics and genomics of the laboratory mouse. *Nucleic Acids Research*, **40**, D881 – 886.

Ersts P (Internet) Geographic Distance Matrix Generator (ver 1.2.1) American Museum of Natural History, Center for Biodiversity and Conservation. [cited 2013/08/12]. Available from http://biodiversityinformatics.amnh.org/open-source/gdmg.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611 – 2620.

Ewer RF (1969) Form and function in the grasscutter, *Thryonomys swinderianus*, Temm. (Rodentia, Thryonomyidae). *Ghana Journal of Science*, **9**, 131 – 140.

Excoffier L (2004) Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Molecular Ecology*, **13**, 853 – 864.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Research*, **10**, 564 – 567.

Faircloth BC (2008) Msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources*, **8**, 92 – 94.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567 – 1587.

Flanders J, Jones G, Benda P, Dietz C, Zhang S, Li G et al. (2009) Phylogeography of the greater horseshoe bat, *Rhinolophus ferrumequinum*: contrasting results from mitochondrial and microsatellite data. *Molecular Ecology*, **18**, 306 – 318.

Flint J, Valdar W, Shifman S, Mott R (2005) Strategies for mapping and cloning quantitative trait genes in rodents. *Nature Reviews Genetics*, **6**, 271 – 286.

Fu Y-X (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and backgroud selection. *Genetics*, **147**, 915 – 925.

Gailard C, Strauss F (1994) Association of poly(CA)poly(GT) DNA fragments into four-stranded complexes bound by HMG1and 2. *Science*, **264**, 433 – 436.

Gaines MS, Diffendorfer JE, Tamarin RH, Whittam TS (1997) The effects of habitat fragmentation on the genetic structure of small mammal populations. *Journal of Heredity*, **88**, 294 – 304.

Glenn TC, Schable NA (2005) Isolating microsatellite DNA loci. *Methods in Enzymology*, **395**, 202 – 222.

Goudet J (2002) FSTAT: a program to estimate and test gene diversities and fixation indices. [cited 2013/07/01]. Available from http://www2.unil.ch/popgen/softwares/fstat.htm

Hamada H, Petrino MG, Takunaga T (1982) A novel repeated element with Z-DNA forming potential is widely found in evolutionary diverse eukaryotic genomes. *Proceedings of the National Academy of Sciences of the United States of America*, **79**, 6465 – 6469.

Hamada H, Seidman M, Howard B, Gorman CM (1984) Enhanced gene expression by the poly(dT-dG) multiplied by poly(dC-dA) sequence. *Molecular Cell Biology*, **4**, 2622 – 2630.

Hargrove J, Gitzendanner M, Austin JD (2012) Isolation and characterization of 14 novel polymorphic loci for the Florida mouse (*Podomys floridanus*). *Conservation Genetic Resources*, **4**, 411 – 413.

Harpending RC (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology*, **66**, 591 – 600.

Hennequin C, Thierry A, Richard GF, Lecointre G, Nguyen HV, Gaillardin C, Dujon B (2001) Microsatellite typing as a new tool for identification of *Saccharomyces cerevisiae* strains. *Journal of Clinical Microbiology*, **39**, 551 – 559.

Henry AJ (2011) Reproductive performance of grasscutter does at first parity and growth performance of their F1 generation. *Asian Journal of Animal Science*, **5**, 289 – 295.

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quartenary. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 183 – 195.

Hirota T, Hirohata T, Mashima H, Satoh T, Obara Y (2004) Population structure of the large field mouse, *Apodemus speciosus* (Rodentia: Muridae), in suburban landscape, based on mitochondrial D-loop sequences. *Molecular Ecology*, **13**, 3275 – 3282.

Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA (2010) Population genomics of parallel adaptation in Threespine stickleback using sequenced RAD tags. *PLoS Genetics*, **6**, e1000862.

Hubbard TJP, Aken BL, Beall K, Ballester B, Caccamo M, Chen Y, Clarke L, Coates G, Cunningham F, Cutts T, Down T, Dyer SC, Fitzgerald S, Fernandez-Banet J, Graf S, Haider S, Hammond M et al (2007) Ensembl 2007. *Nucleic Acids Research*. 35: D610 – D617.

Inoue-Murayama M, Kayang BB, Kimura K, Ide H, Nomura A, Takahashi H, Nagamine Y, Takeda T, Hanada H, Tatsuda K, Tsudzuki M, Matsuda Y, Mizutani M, Murayama Y, Ito S (2001) Chicken microsatellite markers are not efficient markers for Japanese quail. *Animal Genetics*, **32**, 7 – 11.

Ito H, Adenyo C, Yoshikawa N, Murayama M (2013) Large scale development of microsatellite markers for wildlife and its application to conservation. *Biological Science*, **64**, 159 – 167 (in Japanese).

IUCN (2011) IUCN Red List of Threatened Species. Version 2011.1. [cited 2011/09/30]. Available from http://www.iucnredlist.org

Jacob HJ, Brown DM, Bunker RK, Daly MJ, Dzau VJ, Goodman A, Koike G, Kren V, Kurtz T, Lernmark A, Levan G, Mao Y, Pettersson A, Pravnec M, Simon JS, Szpirer C, Szpirer J, Trolliet MR, Winer ES, Lander ES (1995) A genetic linkage map of the laboratory rat, *Rattus norvegicus*. *Nature Genetics*, **9**, 63 – 69.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801 – 1806.

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution*, **11**, 424 – 429.

Jennings TN, Knaus BJ, Mullins TD, Haig SM, Cronn RC (2011) Multiplexed microsatellite recovery using massively parallel sequencing. *Molecular Ecology Resources*, **11**, 1060 – 1067.

Jensen J, Bohonak A, Kelly S (2005) Isolation by distance, web service. *BMC Genetics*, **6**, 13.

Jori F, Mensah GA, Adjanohoun E (1995) Grasscutter production: an example of rational exploitation of wildlife. *Biodiversity Conservation*, **4**, 257 – 265.

Kankam T, Adu EK, Awumbila B (2009) Gastrointestinal parasites of the grasscutter (*Thryonomys swinderianus*, Temminck 1827) on the Accra Plains of Ghana. *African Journal of Ecology*, **47**, 416 – 421.

Karikari PK, Nyameasem JK (2009) Productive performance and carcass characteristics of captive grasscutters (*Thryonomys swinderianus*) fed concentrate diets containing varying levels of guinea grass. *World Applied Science Journal*, **6**, 557 – 563.

Katti MV, Ranjekar PK, Gupta VS (2001) Differential distribution of simple sequence repeats in eukaryotic genome sequences. *Molecular Biology and Evolution*, **18**, 1161 – 1167.

Kayang BB, Youssao I, Inoue E, Naazie A, Abe H, Ito S, Inoue-Murayama M (2010) Genetic diversity of Helmeted guinea fowl based on microsatellite analysis. *Journal of Poultry Science*, **47**, 120 – 124.

Kruglyak S, Durrett RT, Schug MD, Aquadro CF (1998) Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. Proceedings of the *National Academy of Science*, **95**, 10774 – 10778.

Lacy RC (1987) Loss of genetic diversity from managed population: interacting effects of drift, mutation, immigration, selection and population subdivision. *Conservation Biology*, **1**, 143 – 158.

Lance SL, Light JE, Jones KL, Hagin C, Hafner JC (2010) Isolation and characterization of 17 polymorphic microsatellite loci in the kangaroo mouse, genus *Microdipodops* (Rodentia: Heteromyidae). *Conservation Genetic Resources*, **2**, 139 – 141.

Langella O (1999) POPULATIONS: Population genetic software (individuals or populations distances, phylogenetic trees). [cited 2013/10/12]. Available from http://bioinformatics.org/populations/.

Larizza A, Pesole G, Reyes A, Sbisa E, Saccone C (2002) Lineage specificity of the evolutionary dynamics of the mtDNA D-loop region in rodents. *Journal of Molecular Evolution*, **54**, 145 – 155.

Latinne A, Waengsothorn S, Risterucci M, Michaux JR (2011) Isolation, characterization and PCR multiplexing of polymorphic microsatellite markers in the threatened murine rodent, Leopoldamys neilli. *Conservation Genetic Resources*, **4**, 511 – 513.

Li WH, Ellsworth DL, Krushkal J, Chang BH, Hewett-Emmett D (1996) Rates of nucleotide substitution in primates and rodents and the generation-time effect hypothesis. *Molecular Phylogenetic and Evolution*, **5**, 182 – 187.

Lim S, Notley-MacRobb L, Lim M, Carter DA (2004) A comparison of the nature and abundance of microsatellites in 14 fungal genomes. *Fungal Genetics and Biology*, **41**, 1025 – 1036.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, **4**, 981 – 994.

Martin J-F, Pech N, Meglecz E, Ferreira S, Costedoat C, Dubut V, Malausa T, Gilles A (2010) Representativeness of microsatellite distributions in genomes, as revealed by 454 GS-FLX Titanium pyrosequencing. *BMC Genomics*, **11**, 560 – 572.

Meglécz E, Nève G, Biffin E, Gardener MG (2012) Breakdown of phylogenetic signal: A survey of microsatellite densities in 454 shogun sequences from 154 non model eukaryotic species. *PLoS ONE*, **7**, e40861.

Mensah GA, Okeyo AM (2006) Continued harvest of the diverse African animal genetic resources from the wild through domestication as a strategy for sustainable use: A case of the larger Grasscutter (*Thryonomysswinderianus*). In: Animal Genetics Training Resource, version 2, 2006. Ojango, J.M., Malmfors, B. and Okeyo, A.M.

(Eds.).International Livestock Research Institute, Nairobi, Kenya, and Swedish University of Agricultural Sciences, Uppsala, Sweden.

Meuwissen THL, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, **157**, 1819 – 1829.

Meyer J, Kohnen A, Brandl R (2009) Genetic differentiation in an arboreal rodent from African savannas. *African Journal of Ecology*, **48**, 831 – 836.

Mignon-Grasteau S, Boissy A, Bouix J, Faure J-M, Fisher AD, Hinch GN, Jensen P, Neindre PL, Mormede P, Prunet P, Vandeputte M, Beaumont C (2005) Genetics of adaption and domestication in livestock. *Livestock Production Science*, **93**, 3 – 14.

Mouchaty SK, Catzeflis F, Jankel A, Arnason U (2001) Molecular evidence of African Phiomorpha-South American Caviomorpha clade and support for Hystricognathi based on the complete mitochondrial genome of the cane rat (*Thryonomys swinderianus*). *Molecular Phylogenetic and Evolution*, **18**, 127 – 135.

Neff BD, Gross MR (2001) Microsatellite evolution in vertebrates: inference from AC dinucleotide repeats. *Evolution*, **55**, 1717 – 1733.

Nicolas V, Bryja J, Akpatou B, Konecny A, Lecompte E, Colyn M, Lalis A, Couloux A, Denys C, Granjon L (2008) Comparative phylogeography of two sibling species of forest-dwelling rodent (*Praomys rostratus* and *P. tullbergi*) in West Africa: different reactions to past forest fragmentation. *Molecular Ecology*, **17**, 5118 – 5134.

Ntiamoa-Baidu Y (1997) Wildlife and food security in Africa. FAO Conservation Guide 33. FAO.

Ntiamoa-Baidu Y (1998) Wildlife Development Plan 1998 - 2003, Vol. 6: Sustainable Use of Bushmeat. Wildlife Department, Ministry of Lands and Forestry, Accra.

Oboegbulem SI and Okoronkwo I (1990) Salmonallae in the African great cane rat, (*Thryonomys swinderianus*). *Journal of Wildlife Disease*, **26**, 199 – 121.

Ojeda AA (2010) Phylogeography and genetic variation in the South American rodent *Tympanoctomys barrerae* (Rodentia: Octodontidae). *Journal of Mammalogy*, **91**, 302 – 313.

Opara MN (2010a) The Grasscutter I: A livestock of tomorrow. *Research Journal of Forestry*, **4**, 119 – 135.

Opara MN (2010b) Grasscutter: The haematology and major parasites. Research Journal of Parasitology, **5**, 214 – 223.

Opara MN, Fagbemi BO (2008) Occurence and prevalence of gastro-intestinal helminthes in the wild grasscutter (*Thryonomys swinderianus,* Temminck). *Life Science Journal*, **5**, 50 – 56.

Owusu BA, Adu EK, Awotwi EK, Awumbila B (2010) Embryonic resorption, litter size and sex ratio in the grasscutter, *Thryonomys swinderianus*. *Animal Reproduction Science*, **118**, 366 – 371.

Owusu EH, Ntiamoa-Baidu Y, Ekpe EK (2006) The dependence of local people on bushmeat in the Afadjato and Agumatsa Conservation Area, Ghana. *Nature & Faune*, **21**, 33 – 44.

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology*, **13**, 55 – 65.

Page RD (1996) TreeView: an application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences,* **12**, 357 – 358.

Pardue ML, Lowenhaupt K, Rich A, Nordheim A (1987) (dC-dA)n (dG-dT)n sequences have evolutionarily conserved chromosomal locations in *Drosophila* with implications for roles in chromosome structure and function. *EMBO Journal,* **6**, 1781 – 1789.

Park SDE (2001) Trypanotolerance in West African cattle and the population genetic effects of selection. PhD Thesis. University of Dublin.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in excel. Population genetic software for teaching and research. *Molecular Ecology Notes,* **6**, 288 – 295.

Pino JL, Ascunce MS, Reed D, Hugot J-P (2012) Development and characterization of microsatellite markers for the endangered Laotian rock-rat (*Laonastes aenigmamus*) using 454-sequencing technology. *Conservation Genetic Resources,* **4**, 999 – 1002.

Piry S, Alapetite A, Cornuet J-M, Paetkau D, Baodouin L, Estoup A (2004) GENECLASS 2: a software for genetic assignment and first generation migrant detection. *Journal of Heredity*, **95**, 536 – 539.

Pritchard JK, Stephens M, Donelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945 – 959.

Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19**, 2092 – 2100.

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 9197 – 9201.

Reynolds J, Weir BS, Cockerham CC (1983) Estimation for the coancestry coefficient: basis for a short-term genetic distance. *Genetics*, **105**, 767 – 779.

Rice WR (1989) Analysing tables of statistical tests. *Evolution*, **43**, 223 – 225.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137 – 138.

Rousset F (2008) GENEPOP'007: a complete re-implementation of GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103 – 106.

Santana QC, Coetzee MPA, Steenkamp ET, Mlonyeni OX (2009) Microsatellite discovery by deep sequencing of enriched genomic libraries. *BioTechniques*, **46**, 217 – 223.

Sassi PL, Chiappero MB, Borghi C, Gardenal CN (2011) High genetic differentiation among populations of the small cavy *Microcavia australis* occupying different habitats. J*ournal of Experimental Zoology*, **315**, 337 – 348.

Schneider S, Excoffier L (1999) Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: Application to human mitochondrial DNA. *Genetics*, **152**, 1079 – 1089.

Schuelke M (2000) An economic method for the fluorescent labelling of PCR fragments. *Nature Biotechnology*, **18**, 233 – 234.

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters*, **9**, 615 – 629.

Shalom A, Darvasi A (2002) Experimental designs for QTL fine mapping in rodents. In: Quantitative Trait Loci (pp. 199 - 223). Humana Press.

Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, **47**, 264 – 279.

Stallings RL, Ford AF, Nelson D, Torney DC, Hildebrand CE, Moyzis RK (1991) Evolution and distribution of (AC)n repetitive sequences in mammalian genomes. *Genomics*, **10**, 807 – 815.

Taberlet P, Waits LP, Luikart G (1999) Noninvasive genetic sampling: look before you leap. *Trends in Ecology and Evolution*, **14**, 323 – 327.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution*, **28**, 2731 – 2739.

Tautz D, Renz M (1984) Simple sequence are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Research*, **12**, 4127 – 4138.

Toews DPL, Brelsford A (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, **21**, 3907 – 3930.

Toth G, Gaspari Z, Jurka J (2000) Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Research*, **10**, 967 – 981.

Van Arendonk JAM, Tier B, Kinghorn BP (1994) Use of multiple genetic markers in prediction of breeding values. *Genetics*, **137**, 319 – 329.

Van der Merwe M (2007) Discriminating between *Thryonomys swinderianus* and *Thryonomys gregorianus*. *African Zoology*, **42**, 165 – 171.

Van Oosterhout C, Hutchinson WF, Wills DP, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535 – 538.

Wilson GA, Rannala B (2003) Bayesian influence of recent migration rates using multilocus genotypes. *Genetics*, **163**, 1177 – 1191.

Worley KC, Gunaratne P (2006) Rat Genome (*Rattus norvegicus*). Encyclopedia of Molecular Cell Biology and Molecular Medicine.

Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. *Molecular Ecology*, **11**, 1 – 16.

6. APPENDICES

Appendix A

List of pairs of loci that show tendency of linkage

Locus1	Locus2	*p*-value
Tsw08	Tsw11	0.0183
Tsw13	Tsw15	0.0466
Tsw05	Tsw18	0.0394
Tsw07	Tsw25	0.0107
Tsw05	Tsw25	0.0372
Tsw13	Tsw26	0.0007
Tsw18	Tsw28	0.0169
Tsw17	Tsw29	0.0084
Tsw12	Tsw30	0.0174
Tsw20	Tsw32	0.0052
Tsw15	Tsw34	0.0214
Tsw30	Tsw35	0.0069
Tsw05	Tsw35	0.0477
Tsw12	Tsw36	0.0077
Tsw23	Tsw36	0.0424
Tsw17	Tsw38	0.0491
Tsw17	Tsw40	0.0477
Tsw26	Tsw45	0.0354
Tsw10	Tsw46	0.0154
Tsw25	Tsw46	0.0227
Tsw25	Tsw47	0.0212
Tsw17	Tsw47	0.0238
Tsw20	Tsw49	0.0259
Tsw24	Tsw49	0.0466
Tsw08	Tsw51	0.0174
Tsw26	Tsw51	0.0257
Tsw08	Tsw52	0.0075
Tsw06	Tsw52	0.0229
Tsw34	Tsw52	0.0315
Tsw31	Tsw53	0.0452
Tsw45	Tsw54	0.0262
Tsw10	Tsw56	0.0253
Tsw14	Tsw57	0.0395
Tsw51	Tsw58	0.0058
Tsw46	Tsw58	0.0452
Tsw07	Tsw60	0.0451
Tsw17	Tsw61	0.0481
Tsw55	Tsw62	0.0049
Tsw26	Tsw63	0.0147

List of pairs of loci that show tendency of linkage (continued)

Locus 1	Locus 2	*p*-value
Tsw10	Tsw63	0.0220
Tsw60	Tsw64	0.0021
Tsw41	Tsw64	0.0123
Tsw32	Tsw64	0.0192
Tsw39	Tsw65	0.0095
Tsw18	Tsw67	0.0371
Tsw50	Tsw67	0.0403
Tsw50	Tsw69	0.0280
Tsw15	Tsw69	0.0348
Tsw65	Tsw69	0.0441
Tsw45	Tsw70	0.0089
Tsw51	Tsw70	0.0260
Tsw39	Tsw70	0.0295
Tsw65	Tsw70	0.0478
Tsw32	Tsw71	0.0171
Tsw07	Tsw71	0.0227
Tsw05	Tsw71	0.0279
Tsw59	Tsw73	0.0027
Tsw17	Tsw73	0.0491
Tsw53	Tsw74	0.0258
Tsw48	Tsw75	0.0023
Tsw72	Tsw75	0.0350
Tsw54	Tsw76	0.0147
Tsw34	Tsw76	0.0286
Tsw12	Tsw76	0.0344
Tsw30	Tsw77	0.0029
Tsw05	Tsw77	0.0204
Tsw35	Tsw77	0.0251
Tsw76	Tsw77	0.0321
Tsw41	Tsw77	0.0361
Tsw20	Tsw77	0.0461
Tsw64	Tsw77	0.0462
Tsw36	Tsw78	0.0126
Tsw08	Tsw78	0.0382
Tsw18	Tsw80	0.0079
Tsw28	Tsw80	0.0248
Tsw05	Tsw81	0.0044
Tsw23	Tsw81	0.0342
Tsw50	Tsw82	0.0281

List of pairs of loci that show tendency of linkage (continued)

Locus 1	Locus 2	*p*-value
Tsw18	Tsw83	0.0038
Tsw75	Tsw83	0.0307
Tsw72	Tsw83	0.0351
Tsw80	Tsw83	0.0415
Tsw01	Tsw84	0.0017
Tsw65	Tsw84	0.0165
Tsw38	Tsw84	0.0367
Tsw75	Tsw85	0.0083
Tsw81	Tsw85	0.0293
Tsw59	Tsw86	0.0155
Tsw12	Tsw86	0.0482
Tsw35	Tsw87	0.0335
Tsw05	Tsw87	0.0364
Tsw23	Tsw88	0.0066
Tsw29	Tsw88	0.0262
Tsw76	Tsw88	0.0266
Tsw54	Tsw89	0.0116
Tsw38	Tsw90	0.0218
Tsw81	Tsw91	0.0053
Tsw59	Tsw91	0.0269
Tsw73	Tsw91	0.0433
Tsw50	Tsw91	0.0490
Tsw25	Tsw92	0.0369
Tsw14	Tsw93	0.0473
Tsw48	Tsw94	0.0068
Tsw75	Tsw94	0.0083
Tsw85	Tsw94	0.0185
Tsw15	Tsw94	0.0219
Tsw57	Tsw94	0.0380
Tsw83	Tsw94	0.0470
Tsw17	Tsw95	0.0038
Tsw60	Tsw95	0.0110
Tsw73	Tsw95	0.0119
Tsw25	Tsw95	0.0187
Tsw88	Tsw95	0.0307
Tsw20	Tsw95	0.0428
Tsw40	Tsw97	0.0287
Tsw31	Tsw99	0.0113
Tsw06	Tsw99	0.0226

List of pairs of loci that show tendency of linkage (continued)

Locus 1	Locus 2	*p*-value
Tsw05	Tsw99	0.0379
Tsw05	Tsw100	0.0018
Tsw81	Tsw100	0.0377
Tsw17	Tsw100	0.0488
Tsw04	Tsw101	0.0029
Tsw23	Tsw101	0.0209
Tsw77	Tsw102	0.0243
Tsw96	Tsw102	0.0250
Tsw02	Tsw102	0.0437
Tsw74	Tsw103	0.0216
Tsw24	Tsw103	0.0241
Tsw82	Tsw103	0.0305
Tsw50	Tsw103	0.0466
Tsw80	Tsw104	0
Tsw82	Tsw105	0.0459
Tsw76	Tsw105	0.0490
Tsw30	Tsw106	0.0090
Tsw47	Tsw106	0.0312
Tsw46	Tsw106	0.0354
Tsw34	Tsw107	0.0037
Tsw27	Tsw107	0.0134
Tsw73	Tsw107	0.0216
Tsw32	Tsw107	0.0288
Tsw28	Tsw107	0.0394
Tsw13	Tsw108	0.0150
Tsw03	Tsw108	0.0175
Tsw65	Tsw108	0.0279
Tsw84	Tsw108	0.0279
Tsw101	Tsw108	0.0339
Tsw93	Tsw109	0.0001
Tsw59	Tsw109	0.0314
Tsw65	Tsw109	0.0441
Tsw88	Tsw109	0.0460
Tsw76	Tsw110	0.0415
Tsw75	Tsw112	0.0289
Tsw92	Tsw112	0.0381
Tsw32	Tsw114	0.0101
Tsw107	Tsw114	0.0109
Tsw70	Tsw114	0.0471

7. ACKNOWLEDGMENTS

I would like to first of all thank the Almighty God for the gift of life and His immense and sustaining grace. He has granted me good health to be able to carry out this study. I ascribe to Him all glory and honour.

My heartfelt thanks go to my supervisor Miho Inoue-Murayama for her tremendous support, critical but kind supervision and motherly care throughout my study period in Japan. Her concern and support included not only academics but my general wellbeing of which I am profoundly grateful. I would also like to appreciate the following teaching staff of Wildlife Research Center for their guidance; Shiro Koshima, Gen'ichi Idani, Hideyuki Sugiura, Masayuki Tanaka, Michio Nakamura and Tadamichi Morisaka. I thank the administrative staff members including Tomoko Kurihara, Tokunori Nanjo, Shigeki Fukugaki, Eiji Kotera, Sawako Takahashi, Takako Endoh and Miki Sugitani for their administrative support.

I am immensely grateful to the following people who assisted me in various ways as far as my research is concerned; Boniface Baboreka Kayang, Erasmus Henaku Owusu, Phyllis Addo, Sekyere Yaw Annor (University of Ghana), Eiji Inoue, Azusa Hayano, Shigenobu Yazawa (Kyoto University).

I would like to sincerely acknowledge the Ministry of Education, Culture, Sports, Science and Technology, Japan (MEXT), Asia and Africa Science Platform Program under the Japan Society for the Promotion of Science (JSPS), the Global Center of Excellence of Kyoto University and Environment Research and Technology Development Fund for providing funds and resources for this study.

Furthermore, I would like to express my appreciation to Azusa Hayano and Miyuki Sato (Kyoto University) for their dedicated work on maintenance of laboratory equipment and data management. My sincere thanks also go to all the graduate students and postdoctoral fellows in the Wildlife Research Center for the helpful suggestions and discussion. I would like to thank all members of onigashima including Eiji Inoue, Azusa Hayano, Akitsugu Konno, Hideaki Abe, Coline Arnaud, Kodzue Kinoshita, Sherif Ramadan, Hideyuki Ito, Etienne-François Okoue, Augustin Basabose, Chihiro Hiramatsu, Yusuke Hori, Saki Yasui, Hisayo Kishi and Akane Ishihara for their useful comments on my study during our weekly meetings.

I would like to express my gratitude to my family especially my parents for their support and encouragement. I would like to especially thank Bright Adenyo for his assistance during field visits for sample collection.

Lastly, I would like to thank all my friends in Japan especially members of Assembly Kyoto Church for providing a congenial atmosphere for weekly fellowship. I thank the Pastors for their prayer support and spiritual guidance.

8. PUBLICATIONS ASSOCIATED WITH THIS THESIS

This thesis consists of three parts of which two have been published in peer-reviewed scientific journals and the latter part is now in preparation to be submitted soon. Although I was assisted by co-authors, I was the coordinator and main contributor to all parts of this scientific research, including laboratory works, data analysis and manuscript writing.

The publications associated with this thesis are the following:

Chapter 2

Adenyo C, Hayano A, Inoue E, Kayang BB, Inoue-Murayama M (2012) Development of microsatellite markers for grasscutter (*Thryonomys swinderianus*, RODENTIA) using next-generation sequencing technology. *Conservation Genetic Resources,* **4**, 1011-1014.

Chapter 3

Adenyo C, Hayano A, Inoue E, Kayang BB, Owusu EH, Inoue-Murayama M (2013) Mitochondrial D-loop diversity of grasscutter (*Thryonomys swinderianus*, RODENTIA, HYSTRICOMORPHA) in Ghana. *Open Journal of Animal Sciences,* **3**, 145 – 153.

Adenyo C, Kayang BB, Owusu EH, Inoue-Murayama M. Genetic diversity of grasscutter (*Thryonomys swinderianus*) in Ghana based on microsatellite markers (in preparation).

9. LIST OF FIGURES

Figure 2.1 Frequency of different repeat types (di – hexa) of the grasscutter based on partial genome data.

Figure 3.1.1 The map of Ghana showing three agro-ecological zones.

Figure 3.1.2 Network among haplotypes.

Figure 3.1.3 Harpendings mismatch distribution showing observed and simulated frequencies of pairwise differences for each population.

Figure 3.2.1 Neighbour-joining tree of grasscutter populations in Ghana based on Nei's genetic distance, $D_{A.}$

Figure 3.2.2 Structure clustering of grasscutter populations in Ghana.

10. LIST OF TABLES

Table 2.1 Fragment read characteristics of the partial genome data of the grasscutter using three sequencing approaches.

Table 2.2 Polymorphism of different repeat loci from the partial genome data of the grasscutter.

Table 2.3 Characteristics of 116 microsatellites of the grasscutter.

Table 3.1.1 Haplotypes and their frequencies in each population.

Table 3.1.2 Genetic diversity and indices of population neutrality across the three agro-ecological zones.

Table 3.1.3 Results of analysis of molecular variance.

Table 3.1.4 Matrix of pairwise F_{ST} values for the three agro-ecological zones.

Table 3.2.1 Profile of 12 microsatellite loci.

Table 3.2.2 Characteristics of grasscutter populations.

Table 3.2.3 Matrix of pairwise F_{ST} (below diagonal) and Nei's genetic distance (above diagonal).

Table 3.2.4 Mean (95% confidence interval) migration rates among populations of grasscutter.

11. LIST OF APPENDICES

Appendix A List of pairs of loci that show tendency of linkage.