Department of Forensic Medicine & Tocxicology
Faculty of Veterinary Medicine Benha University

# Application of molecular tools in parentage verification of Egyptian Arabian Horses

*Thesis submitteded by*

**Mary Aiad Nowar Sargious**
B.V.Sc., 2010, Benha University
M.V.Sc., 2014, Benha University

*FOR*

**PHD Degree (Forensic Medicine & Toxicology)**

*Under supervision of*

**Prof. Dr.**
**Ragab Mahmoud El-Shawarby**
Prof. of Forensic Medicine & Toxicology.
Faculty of Veterinary Medicine,
Benha University

**Prof. Dr.**
**Mohamed El-said Sobhy abou- Salem**
Professor of forensic Medicine & Toxicology.
Faculty of Veterinary Medicine,
Benha University.

**Prof. Dr. Elham Abd El-Monaem Elsheiwy**
Professor and head of Department of forensic Medicine & Toxicology. Faculty of Veterinary Medicine, Benha University.

**Dr.**
**Hanaa Abdel Kader Ahmed**
Chief Researcher
Genome Research Unit
Animal Health Research Institute,
Dokki, Giza.

**2019**

# List of contents


| | |
|---|---|
| 1. **Introduction**. | **1** |
| **2. Review of literature.** | **4** |
| **3. Material and methods:** | **13** |
| **3.1. Material**: | **13** |
| **3.1.1.** Samples. | **13** |
| **3.1.2.** Kits. | **13** |
| **3.1.3.** Consumables. | **14** |
| **3.1.4.** Equipment. | **14** |
| **3.1.5.** Reagent. | **14** |
| **3.2. Methods**: | **15** |
| **3.2.1.** Extraction of DNA from hair. | **15** |
| **3.2.2.** Determination of DNA concentration. | **16** |
| **3.2.3.** Preparation of seventeen Plex PCR mix. | **17** |
| **3.2.4.** Thermal-cycling conditions of seventeen plex PCR mix. | **17** |
| **3.2.5.** Prepare the amplified equine DNA for analysis by the Genetic analyzer 3500. | **17** |
| **3.2.6.** Determining the size of amplified Equine DNA. | **18** |
| **3.2.7.** Preparation of 14 TKY Plex PCR mix. | **19** |
| **3.2.8.** Thermal-cycling conditions of 14 TKY plex PCR mix. | **19** |
| **3.2.9.** Prepare the amplified equine DNA for analysis by the Genetic analyzer 3500. | **19** |
| **3.2.10.** Data Acquisition. | **20** |
| | **21** |

| | |
|---|---|
| **3.2.11.** Data Analysis. | **21** |
| **3.3. Statistical analysis:** | |
| **3.3.1.** Marker polymorphisms and population diversity of the three studied populations. | **21** |
| **3.3.2.** Relationships and Population Structure of the three studied populations. | **21** |
| **3.3.3.** Marker polymorphisms and genetic diversity of ISAG and TKY panels for 41 Arabian samples. | **22** |
| **3.3.4.** Parentage test**.** | **22** |
| **4. Results**: | **24** |
| **4.1.** Concentration of dsDNA samples (ng/ml). | **24** |
| **4.2.** Marker polymorphisms and population diversity of the three studied populations. | **26** |
| **4.3.** Relationships and Population Structure of the three studied populations. | **37** |
| **4.4.** Marker polymorphisms and genetic diversity of ISAG and TKY panels for 41 Arabian samples. | **39** |
| **4.5.** Parentage test. | **41** |
| **5. DISCUSSION:** | **48** |
| **5.1.** Marker polymorphisms and population diversity of the three studied populations. | **48** |
| **5.2.** Relationships and Population Structure of the three studied populations. | **49** |
| **5.3.** Marker polymorphisms and genetic diversity of ISAG and TKY panels for 41 Arabian Horses samples. | **50** |

| | |
|---|---|
| **5.4.** Parentage test. | **50** |
| **6. Conclusion and recommendations.** | **53** |
| **7. Summary.** | **55** |
| **8. References.** | **59** |
| **9. Arabic Summary.** | **66** |

# List of Tables


| Table No. | Title | Page |
|---|---|---|
| 1 | List of fourteen equine TKY microsatellite loci. | 9 |
| 2 | Genotypes of 13 TKY markers of control samples of HCT 2019. | 13 |
| 3 | Preparation of seventeen Plex PCR mix. | 17 |
| 4 | Thermal-cycling conditions of seventeen plex PCR mix. | 17 |
| 5 | Equine specific loci and their labels and types of filter. | 18 |
| 6 | Preparation of 14 TKY Plex PCR mix. | 19 |
| 7 | Thermal-cycling conditions of seventeen plex PCR mix. | 19 |
| 8 | Equine specific TKY loci and their labels and types of filter. | 20 |
| 9 | Concentration of Non Arabian Samples. | 24 |
| 10 | Concentration of Arabian samples. | 25 |
| 11 | Number of alleles (Na) effective alleles (Ne), Observed heterozygosity ($H_O$), expected heterozygosity ($H_E$), polymorphic information content (PIC), $F_{IS,}$ and Hardy | 27 |

| | | |
|---|---|---|
| | Weinberg equilibrium (HWE) across the Arabian population. | |
| 12 | Number of alleles (Na) effective alleles (Ne), Observed heterozygosity ($H_O$), expected heterozygosity ($H_E$), polymorphic information content (PIC), $F_{IS,}$ and Hardy Weinberg equilibrium (HWE) across the three studied populations. | 28 |
| 13 | Observed (*Na)* and effective (*Ne*) number of alleles, observed ($H_O$) and expected ($H_E$) heterozygosity, and fixation coefficient of an individual within a subpopulation ($F_{IS}$) per population for three studied populations. | 29 |
| 14 | Pairwise Reynolds's genetic distance (above diagonal) and pairwise $F_{ST}$ (below diagonal) between the 3 studied populations. | 37 |
| 15 | Number of alleles (Na) effective alleles (Ne), Observed heterozygosity ($H_O$), expected heterozygosity ($H_E$), polymorphic information content (PIC), and the probability of exclusion (PE) of the 12 ISAG markers. | 40 |
| 16 | Number of alleles (Na) effective alleles (Ne), Observed heterozygosity ($H_O$), expected heterozygosity ($H_E$), polymorphic information content (PIC), and the probability of exclusion (PE) of the 12 TKY markers. | 40 |
| 17 | Parentage assignment of eight sire families by 12 microsatellite loci of ISAG panel. | 44 |

| | | |
|---|---|---|
| 18 | Parentage assignment of eight sire families by 12 microsatellite loci of TKY panel. | 45 |
| 19 | GRU result of ISAG panel for Horse Comparison Test (HCT 2018-2019). | 46 |
| 20 | GRU result of TKY panel for Horse Comparison Test (HCT 2018-2019). | 47 |

# List of Figures


| Figure No. | Title | Page |
|---|---|---|
| 1 | Allele frequencies of locus *TKY337* in the three studied populations. | 29 |
| 2 | Allele frequencies of locus *TKY344* in the three studied populations. | 30 |
| 3 | Allele frequencies of locus *TKY343* in the three studied populations. | 30 |
| 4 | Allele frequencies of locus *TKY321* in the three studied populations. | 31 |
| 5 | Allele frequencies of locus *TKY287* in the three studied populations. | 31 |
| 6 | Allele frequencies of locus *TKY312* in the three studied populations. | 32 |
| 7 | Allele frequencies of locus *TKY301* in the three studied populations. | 32 |
| 8 | Allele frequencies of locus *TKY337* in the three studied populations. | 33 |
| 9 | Allele frequencies of locus *TKY374* in the three studied | 33 |

| | | |
|---|---|---|
| | populations. | |
| 10 | Allele frequencies of locus *TKY297* in the three studied populations. | 34 |
| 11 | Allele frequencies of locus *TKY333* in the three studied populations. | 34 |
| 12 | Allele frequencies of locus *TKY341* in the three studied populations. | 35 |
| 13 | Allele frequencies of locus *TKY325* in the three studied populations. | 35 |
| 14 | Allele frequencies of locus *TKY294* in the three studied populations. | 36 |
| 15 | Allele frequencies of locus *TKY394* in the three studied populations. | 36 |
| 16 | Neighbor Joining Phylogenetic tree for the three studied populations based on 14 TKY markers. | 38 |
| 17 | Structure clustering of the three studied populations obtained for $K = 2$. | 38 |
| 18 | Comparison between the genetic diversity parameters means of ISAG and TKY panel. | 41 |

| | | |
|---|---|---|
| 19 | Combined probability of exclusion of 12 ISAG loci for the eight sire families. | 43 |
| 20 | Combined probability of exclusion of 12 TKY loci for the eight sire families. | 43 |

# 7. Summary

The present study was carried out at the Genome Research Unit (GRU), Animal Health Research Institute (AHRI), to identify the Egyptian Arabian horses in Egypt using fourteen TKY microsatellite markers, aiming to establish TKY system and support equine parentage test using highly polymorphic and discriminatory loci.

**To achieve this objective, the following methodology was performed:**

- Total of 112 hair samples were pulled by their root and collected at El Zahraa stud and brought to the lab. Samples were taken from the mane hair. The total 112 samples are classified into: 71 unrelated individuals and 41 in form of families (8 sires, 18 foals and 16 dams).
- DNA was extracted from hair root.
- Also, Thirty Equine DNA samples were brought for the ISAG Horse Comparison Test (HCT) 2017 and 2019, including 19 samples of Thoroughbred and 11 samples of Nooitgedacht breeds. Extraction of DNA was done by duty lab in South Africa and Germany for ISAG samples 2017 and 2019 respectively. Two reference DNA genotypes were provided for TKY panel by HCT 2019 only.
- The DNA concentration was determined and samples were diluted with PCR grade water.
- A set of 17 microsatellites (VHL20, HTG4, AHT4, HMS7, HTG6, HMS6, HTG7, HMS3, AHT5, ASB2, HTG10, HMS2, ASB17, LEX3, HMS1, CA425, and ASB23) were selected according to the auspices of the International Society for Animal Genetics (ISAG) for equine genotyping and parentage test.
- A set of 14 microsatellites of (*TKY287*, *TKY294*, *TKY301*, *TKY312*,

*TKY321, TKY325, TKY333, TKY341, TKY343, TKY344, TKY 337, TKY 297, TKY374 and TKY394)* were selected according to the recommendations of the International Society for Animal Genetics (ISAG) for single system exclusion cases within the ISAG panel.

- The microsatellite loci of TKY markers were amplified in one multiplex reaction by PCR using locus specific primers by standard PCR protocol. This PCR protocol was carried out with genomic DNA samples of 71 Arabian, 19 Thoroughbred and 11 Nooitgedacht.
- These 12 microsatellites were amplified in one multiplex reaction using Stockmarks®: horse genotyping kit. This PCR protocol was carried out with genomic DNA samples of 41 Arabian families and showing cases of exclusion in one marker within ISAG panel.
- Amplification of TKY markers of these families (dam, sire and foal) to confirm exclusion or inclusion.
- Fragment sizes of microsatellite alleles were determined using Genetic analyzer 3500 (Applied Biosystem-USA) with the aid of Liz standard.
- TKY Panel was adjusted by using known reference samples brought for the ISAG Horse Comparison Test 2019.
- The data obtained is further analyzed using Gene Mapper V 4.1 software.
- Statistical analysis using several softwares was done to estimate the allele frequencies observed and expected heterozygosity, polymorphic information content, Hardy Weinberg Equilibrium, Probability of exclusion and combined probability of exclusion.
- The genetic diversity of the three populations: Arabian, Thoroughbred and Nooitgedacht were measured.

- Parentage analysis was performed with CERVUS program version 3.0 to assign the putative sires, putative dams and putative parent pairs to their offspring.

**Accordingly, it was concluded that:**

- The polymorphism of 14 TKY markers across the Arabian population showed moderate values for genetic diversity parameters (*Na* =8.143, $N_e$ = 3.694, $H_O$, = 0.599, $H_E$ = 0.691 and *PIC* = 0.636).

- On contrast, Nooitgedacht and Thoroughbred populations showed higher genetic diversity (($N_a$ = 6.286 and 5.571; $N_e$ = 4.105 and 3.829; $H_O$ = 0.727 and 0.737; $H_E$ =0.760 *and* 0.741 respectively).

- The Arabian population was assigned independently into its own cluster while the remaining two populations (Nooitgedacht and Thoroughbred) were clustered together forming one admixed cluster.

- Locus *TKY337* should be interpreted with caution and should be analyzed in further studies based on different Arabian populations to test if it is linked to any morphological traits or not.
- The high $F_{IS}$ (0.129) of El Zahraa Arabian horses should be corrected by modifying mating system through avoiding excessive use of certain sires.

- TKY panel showed higher genetic diversity parameters than the ISAG panel, but TKY panel will continue to be an additional and confirmatory panel in case of single system exclusion within ISAG panel.

- Mismatch in one marker within the ISAG panel might be occurred due to either error in the amplification of the STR, or confusion with the stutter

peak, or false binning due to absence of numeric allele design. But whatever the reason of the mismatch, we had the possibility for assurance of inclusion or exclusion by using TKY panel.

- Cervus software provides an accurate and rapid assessment in the assignment of Arabian horse paternity and maternity.

**Finally**: We believe that this is the first study about the genetic diversity of Egyptian Arabian horse using TKY microsatellite panel and also the first report about the genetic characterization of the two panels across Egyptian Arabian horse.

Moreover, we present an applicable economic solution for single system exclusion cases within the ISAG panel by compatibility of at least twelve additional TKY markers parentage assignment of closely related Arabian horse families.