DEFRA Contract No. CR0300

DNA Profiling of Birds of Prey – Fluorescent Multiplexing

Final Report

31st August 2006

R. Ogden, R. McEwing and N. Dawnay

Wildlife DNA Services Limited
Science Campus, University of Wales, Bangor, LL57 2UW

Executive Summary

The aim of this project was to develop a series of genetic tests for the identification of birds belonging to six species of raptor: goshawk, golden eagle, peregrine falcon, gyr falcon, saker falcon and merlin. The tests were commissioned in order to allow individual identification, sex determination and species identification of birds, from blood and feather samples. The resulting techniques were required to be validated for use in forensic investigations with the intention of providing evidence in criminal prosecutions involving bird of prey persecution.

Sample collection targeted captive bred birds of each species, with the aim of obtaining both unrelated birds and birds that constituted family groups. Wild bird samples were also collected for comparative purposes. Over 1000 samples were collected in total.

The construction of the DNA profiling systems for each of the six target species was performed in three principal phases. Potential microsatellite DNA markers were assessed on the basis of their amplification efficiency, variability and heterozygosity, in order to ascertain there suitability for incorporation into the profiling systems. Selected markers were subsequently validated according to forensic guidelines. Validated markers were combined to form PCR multiplexes to allow rapid and cost effective production of DNA profiles.

The resulting DNA profiles were analysed in order to assess their ability to distinguish individual birds. Statistical results indicate that the systems are powerful enough to provide strong evidence of identity in a legal framework.

A sex determination test was developed based on PCR amplification of a region of the avian sex chromosome. This test produces DNA fragments that enable male and female birds to be distinguished. The results can be visualised on agarose gel or incorporated into the individual and species identification tests.

A species identification test was designed that differentiates between each species with the exception of gyr and saker falcons. Due to the similarity of these two species at the DNA sequences analysed, it was not possible to design a robust forensic test capable of distinguishing them. Methods for gyr and saker discrimination at a research level are discussed.

In order to allow the rapid dissemination of the DNA profiling techniques produced in this project, a database containing all the relevant result data has been compiled and is freely available to other researchers in the field. Peer-reviewed publication of the project results is planned.

The report includes a discussion of the issues which arose during the project, including the selection of microsatellite loci and statistical considerations on the use of the profile systems in a forensic context, together with summary conclusions.

Contents

Executive summary		2
1. Introduction	1.1 Background	5
	1.2 Aims and Objectives	5
	1.3 Project Outputs	5
	1.4 Purpose and End Users	5
2. Description of Methods	Standard Techniques	6
	2.1 Sample Collection	7
	2.2 Assessment of Microsatellite Loci	7
	2.3 Forensic Validation of Markers	8
	2.4 Probability of Identity Assessment	9
	2.5 Development of PCR Multiplexes	9
	2.6 Validation of Moulted Feathers	10
	2.7 Development of Genetic Sexing Test	10
	2.8 Development of Species ID Test	11
	2.9 Database Production	12
3. Results	3.1 Sample Collection	14
	3.2 Assessment of Microsatellite Loci	15
	3.3 Forensic Validation of Markers	16
	3.4 Probability of Identity Assessment	20
	3.5 Development of PCR Multiplexes	21
	3.6 Validation of Moulted Feathers	22
	3.7 Development of Genetic Sexing Test	22
	3.8 Development of Species ID Test	23
	3.9 Database Production	24
4. Discussion	4.1 Individual Identification	25
	4.2 Sex Determination	28
	4.3 Species Identification	28
5. Conclusions		30
Acknowledgements		31
References		32
Appendix I	List of Sample Contributors	
Appendix II	List of Microsatellite Loci	
Appendix III	Protocol for Microsatellite Validation	
Appendix IV	Multiplex Design	
Appendix V	Multiplex PCR Conditions	
Appendix VI	Examples of PCR Multiplex Results	
Appendix VII	Species Identification Test Sequence Data	

List of Figures and Tables

Tables

Table 2.1	Details of database design	13
Table 3.1	Summary samples collected	14
Table 3.2	Summary of microsatellite assessment	15
Table 3.3	Details of HWE results for two peregrine loci	16
Table 3.4	Validation results for the Mendelian inheritance study	18
Table 3.5	Validation results for the sensitivity study	19
Table 3.6	Validation results for cross-species amplification	19
Table 3.7	Validation results for allele size verification	20
Table 3.8	Probability of Identity estimates	21
Table 3.9	List of microsatellite loci included in final profiling systems	21

Figures

Figure 3.1	Validation results for PCR annealing temperature variation	17
Figure 3.2	Validation results for PCR cycle number variation	17
Figure 3.3	Validation results for PCR magnesium concentration variation	18
Figure 3.4	Results of the sex determination test	22
Figure 3.5	Design of the species identification test	24

1. Introduction

1.1 Background

The theft of birds of prey from the wild for trade and falconry remains a serious threat to their natural population both in the UK and internationally. Investigations into alleged offences involving birds of prey often require identification of an individual bird or group of birds, in order to establish the species, population or family to which they belong. Individual evidence samples may also need to be matched to confirm or refute identity.

The use of DNA evidence to provide information for bird of prey identification has developed in the UK over the past decade. DNA profiling methods based upon minisatellite SLPs were developed under a previous DoE contract and have proved critical in securing convictions in more than 15 court cases against raptor keepers who had illegally attempted to launder wild birds into the captive population in Britain.

Under a more recent DETR-funded project, a number of raptor microsatellite DNA markers were isolated in order to enable DNA evidence to be obtained from feathers, as opposed to blood samples.

The current project was designed to build on this previous research to produce a series of validated DNA profiling systems capable of individual, species and sex identification in six bird of prey species: Golden eagle, Goshawk, Gyr falcon, Peregrine falcon, Saker falcon and Merlin.

1.2 Aims and Objectives

- i) Development of microsatellite DNA profiles capable of individual discrimination in each of the six target species.
- ii) Development of a genetic sex determination test for each of the six target species.
- iii) Development of a species identification test, capable of distinguishing the six target species.
- iv) Validation of each test to allow application to forensic investigation.

1.3 Project Outputs

- i) Validated peer-reviewed protocols for the genetic identification of bird of prey species commonly traded in the UK.
- ii) A database of DNA profiling markers, validated conditions and allele frequencies for distribution to forensic practitioners.

1.4 Purpose and End Users

The project will increase the ability of UK authorities to enforce wildlife legislation designed to protect native UK species and regulate the international trade in birds of prey. The resulting protocols will be applied by forensic scientists in the UK and abroad as part of investigations led by DEFRA, HMRC, Police and the RSPB, as well as overseas agencies.

2. Description of Methods

In order to meet the aims of the project the work was divided into eight key stages:

1. Sample collection
2. Assessment of existing microsatellite loci
3. Development of novel markers (as necessary)
4. Forensic validation of microsatellite markers
5. Development of a PCR multiplex for each species
6. Development of a genetic sexing test
7. Development of a species identification test
8. Production of a microsatellite database

The eight stages were approached in order. Specific methods are described in the following sections.

Standard Techniques

The genetic analyses performed during the project used standard equipment and reagents following WDNAS SOPs. Unless otherwise specified, the following methods were used throughout:

DNA extraction
Qiagen DNEasy extraction kit following manufacturers instructions.

DNA quantification

Reagents:	PicoGreen® fluorescent dye (Molecular Probes, Inc.)
Detector:	Galaxy Fluostar fluorometer (BMG Labtechnologies LTD).

PCR amplification

Thermocycler:	MJ Research Peltier P-200 or MJ Research Dyad
Reagents:	ABgene PCR 1.1X master mix (various MgCl concentrations)
	Qiagen PCR multiplex kit
Primers:	Operon Biotechnologies (unlabelled) and Proligo (labelled)

Genotyping and Sequencing

Genetic analyser:	Beckman CEQ8000 and proprietary software

2.1 Sample collection

In order to design and validate the DNA profiling assays, a range of samples were required for each species, including samples from family groups, samples from unrelated individuals and samples consisting of different source materials (e.g. blood feather etc.). A target of twenty unrelated captive bred birds and two family groups per each species was set at the start of the project, in order to allow for assessment of microsatellite loci. A comparable number of wild birds from each species were also sought.

Collection of samples was largely reliant upon the assistance of breeders, keepers and researchers involved with UK birds of prey. Over 60 individuals were contacted around the UK and internationally between June and December 2004. Bird keepers were provided with collection envelopes and a FREEPOST address for return of feather samples. Other samples were collected from around the UK by Wildlife DNA Services staff or DEFRA wildlife inspectors. Due to difficulties in obtaining sufficient samples belonging to family groups, a second phase of targeted sampling was undertaken in which specific birds were identified using the DEFRA bird registry (BirdLog), with their respective bird keepers subsequently contacted by DEFRA, rather than WDNAS, in order to maintain bird registry confidentiality. A list of sample contributors is given in Appendix I.

Validating the use of feathers as a DNA source required the collection of blood and feather samples from the same bird. WDNAS staff therefore undertook a Home Office training course to obtain personal licences to take blood and plucked feather samples from birds (Bioscientific Events Ltd, July 2004). In addition staff also completed the Police Wildlife Crime Officers Bird of Prey Handling Course, run by the National Bird of Prey Centre, Newent, Gloucestershire.

As an extension to the original sampling plan, it was decided at the second steering group meeting (December 2004) to increase the number of wild bird samples for at least the four UK native species in the project. In order to achieve this, bird ringers were contacted to request the collection of moulted feathers during routine annual ringing. As certain bird ringers were already involved in collecting DNA samples from newborn chicks, these persons were written to individually. Samples previously collected by ringers and held by J. Wetton (FSS) were also made available to the project. Other ringers were contacted via the British Trust for Ornithology (BTO). Further samples from wild birds were sought from an ongoing research project conducted by the NERC CEH facility at Monkswood (see Appendix I for full list of contributors).

2.2 Assessment of existing microsatellite loci

A list of potential microsatellite loci was generated from previously published studies and a number of unpublished sources. Primers for the amplification of each locus were obtained and a set of unrelated birds of each target species was selected for screening (Appendix II).

Prior to screening, the PCR amplification conditions were optimized for each locus following standard laboratory practices. Loci that were found not to amplify, or that could not be sufficiently optimised, were excluded from the project at this stage.

Species-specific loci were initially used to screen each set of unrelated birds (n~20) in order to assess levels of allelic variability. Additional cross-species loci were subsequently used to screen birds of closely related in order to examine the possibility of cross-amplification.

Throughout the duration of the project, new microsatellite loci were published by other researchers. In certain instances it was possible to incorporate these into the assessment process, however for those loci released during the latter stages, this was not practically possible.

The genotypes recorded at each locus were examined to assess allelic variability, allele frequencies, heterozygosity, linkage and the presence of null alleles. These analyses were performed using the programs Genepop (Raymond and Rousset 1995) and MSA (Dieringer & Schlotterer 2005).

2.3 Forensic validation of markers

i) Purpose

There are two types of validation required to implement or modify technologies for forensic DNA analysis - developmental and internal. Developmental validation is the demonstration of the accuracy, precision, and reproducibility of a procedure by the developmental laboratory. Developmental validation must precede the use of a novel forensic DNA technique and is the type of validation undertaken during this project. Internal validation is conducted by each forensic DNA testing laboratory and is the in-house demonstration of the reliability and limitations of the procedure.

The protocol used for developmental validation originates from a US FBI committee known as the Scientific Working Group on DNA Analysis Methods (SWGDAM). SWGDAM has issued guidelines that cover a broad range of genetic techniques and applications; therefore a customized set of protocols was designed for the current project, based on SWGDAM best practice.

The purpose of validation is to provide windows of optimal performance for each marker/multiplex. These windows allow other laboratories the ability to adjust their protocol while still providing valid results.

ii) Scope

Developmental validation of the microsatellite loci within this project requires multiple investigations of locus characteristics, amplification conditions and the effects of non-standard sample treatment. Key areas of validation are as follows:

1. Locus characteristics: a) Mode of inheritance
   b) Detection – i.e. PCR conditions
   c) Level of polymorphism

2. Species specificity: cross amplification tests

3. Sensitivity: the range of DNA quantities the method will amplify

4. Stability: the ability to obtain results from DNA recovered from samples subject to various environmental insults

5. Reproducibility: ensuring consistency of results replicates

6. Case-type samples: the ability to obtain results from samples that are typically used in casework i.e. feathers, blood, etc

7. Mixture studies: the ability to obtain reliable results from a mixed sample needs to be assessed if mixed samples are common

iii) Approach
A set of protocols was established to address each of these validation parameters within the context of the current project (Appendix III). Validation was subsequently undertaken by applying these protocols to the six target species.

2.4 Probability of Identity assessment

Loci were combined to form potential DNA profiling systems for each species, in order to assess their statistical power to identify individual birds. This was achieved by calculating the probability of identity (PI) for each profile system using the program API-CALC (Ayres and Overall 2004). Levels of population structure, inbreeding and relatedness between birds were simulated to examine the effect that these would have on the ability of the profiling systems to differentiate between related birds within a population.

Simulated values of F_{ST} and F_{IS} were at 0.1. The level of relatedness between birds in the population was defined as 20% of individuals being full siblings and a further 20% being half siblings. These values were designed to represent a relatively inbred population in which multiple birds from the same family were often present. When all the simulated parameters are employed, the resulting PI calculation is considerably more conservative than for human profiling (F_{ST}~0.01, F_{IS}~0), however captive bred bird of prey populations are expected to consist of much more closely related individuals.

2.5 Development of PCR multiplexes

Background
One of the advantages of using microsatellite loci to create DNA profiles is that it is often possible to reduce laboratory time and costs by analysing multiple loci simultaneously (‘multiplexing’). Traditionally, a profiling system consisting of ten loci would require ten separate PCR reactions and ten sample wells on a genetic analyser to create a profile for one individual. The ability to multiplex samples means

that a profile could theoretically be produced from a single PCR reaction and only one well on a genetic analyser, reducing time and reagent costs considerably.

In order to allow the simultaneous amplification of multiple microsatellite loci, it is necessary to use different coloured fluorescent dyes to label individual microsatellites. On most analysis platforms, there is a limit of three different colours available. To multiplex more than three loci it is therefore necessary to select microsatellites of different size, so that two loci of the same colour may still be distinguished. The relative sizes of the microsatellite DNA fragments is therefore the key factor limiting the ability to multiplex samples in this project.

Design

Following the results of the microsatellite assessment and forensic validation, a set of microsatellites was identified for the construction of PCR multiplexes. For each species, dye colours were assigned to the loci in order to maximize the number of loci that could be analysed in a single multiplex, thus minimizing the number of individual analyses required to produce the DNA profile.

PCR Optimization

Once the multiplex had been designed, PCR optimization was performed by altering relative primer concentrations within each reaction until the amplification strength of each locus was approximately equal.

2.6 Validation of moulted feathers as a DNA source

The use of feather material for the recovery and amplification of DNA has previously been demonstrated, however it was formally examined here in order to investigate the ability to generate DNA profiles from feather that were identical to those produced from DNA recovered from blood.

Moulted feathers and blood were collected from three individual birds (golden eagle, gyr falcon and saker falcon). DNA was extracted following standards conditions (p.6), with the elongation of the tissue lysis stage in the feather samples from 240 minutes to 90 minutes. DNA profiles were subsequently produced for each sample and compared.

2.7 Development of genetic sexing test

The objective for the sex determination test was to design a single PCR-based assay that would distinguish between males and females in all six species. In addition, the test would need to perform well from forensic type samples where the DNA size is often reduced as a result of fragmentation. In order to achieve this, a set of existing generic sexing PCR primers (Fridolfsson and Ellegren, 1999) was used as a starting point from which to design new primers that would be used to reliably amplify sex-specific DNA fragments across all species.

DNA from all six raptor species was amplified using the generic primers 2550F and 2718R generating either one or two amplification products dependant on sex: one

approximately 600bp from the Z chromosome (homozygous in males), the other 450bp from the W chromosome (heterozygotic with the Z in females). The single amplification product in males (ZZ) was sequenced using both amplification primers to reveal the DNA sequence of the Z chromosome for each raptor species. In females (ZW) both bands were independently excised from an agarose gel and again sequenced using both amplification primers.

The resulting sequences from each species Z and W chromosome were aligned and new raptor specific primers were designed based on regions of homology between species and chromosome region.

Details of the two new primers are as follows:

Forward: ACGTGGCAACAGAGTWCTGAT
Reverse: CCCCTTTTATTGATCCATCAAGTCTCT

The primers target a locus that varies in length between the W and Z forms of the avian sex chromosomes. In each species, the locus on the Z chromosome is longer than the locus on the W chromosome. Under PCR amplification, this resulted in male birds (with two Z chromosomes) showing a single long DNA fragment, whilst female birds (with one Z and one W chromosome) show one long DNA fragment and one short DNA fragment. The difference in fragment sizes could then be genotyped under electrophoresis on a simple agarose gel system, or on a genetic analyser as part of a larger fluorescent-dye based multiplex.

2.8 Development of species id test

The objective for the species determination test was to design a PCR-based assay that would distinguish between the six target species. Genetic species identification is normally performed by comparing DNA sequence variation among different species at a common gene region. Several genes within the mitochondrial genome were selected for assessment in the design of this test: cytochrome b (cyt b), cytochrome oxidase I (COI) and the control region (CR).

Cytochrome b

Cyt b is the most widely used gene for species identification. A number of studies that have used cyt b to examine raptor phylogeny were examined in an attempt to identify potential species-specific sequence regions within this gene. While studies have shown sequence differences among most of the target species (e.g. Wink and Sauer-Gürth 2004), there has been a lack of categorical differentiation between species within the Hierofalcon complex (saker, gyr, lanner and laggar falcons). The saker falcon was observed to form paraphyletic clades with respect to the lanner falcon, while the difference between gyr falcons and the nearest saker falcon clade was 'minimal'. In addition, although a number of highly relevant studies have been undertaken in this area (Wink *et al.* 2004; Wink and Sauer-Gürth 2000), no sequence data has been submitted to public access databases from this work. Given these issues, alternative candidate genes for the species test were examined.

Cytochrome oxidase subunit I
COI is the gene that has been selected as the basis for genetic species identification in the Bar-Coding-of-Life project and was an obvious second choice as a candidate gene for species identification. Previous studies have demonstrated its utility in identification of birds (Hebert *et al.* 2004) and both sequence data and primers for PCR amplification have been published. Based on this study, additional samples were sequenced from each target species and the total sequence information was aligned in order to identify species-specific regions in COI. As with cyt b, suitable sequence differences were observed for the discrimination of all species with the exception of gyr and saker falcons.

Control Region
In an attempt to discriminate more clearly between gyr and saker falcons, a study by Nittinger *et al.* (2005) examined several gene regions in and around the control region for sequence information that would categorically distinguish the two species. This showed saker falcons to be divided into two clades, one of which was paraphyletic with respect to gyrs.

Nuclear loci
The sex determination test was examined to see whether species-specific PCR product was generated during the amplification of sex-linked loci.

Microsatellite loci were also investigated as a potential source of markers for species discrimination. An assessment of all the loci originally tested in these species was undertaken in an attempt to identify any differences in amplification size between gyr and saker falcons at polymorphic and monomorphic loci.

Microsatellite loci were subsequently used to investigate the possibility of species assignment based on species diagnostic allele frequencies. Allele frequency training sets were generated using DNA profiles of wild gyr and saker falcons. These were analysed using the programme STRUCTURE, to determine whether allele frequency data could form the basis of a species identification approach.

2.9 Database production

In order to enable the techniques developed during this project to be readily used by other researchers, a database containing information relevant to the application of each genetic assay was produced. The key elements of the database are listed in Table 2.1. The objective was to produce a system that returns the necessary information to allow other researchers to repeat the techniques and produce their own results with confidence. The database was designed for the Microsoft Access database programme because this software is sufficiently powerful to provide the necessary functionality and has the advantage of being available to almost all PC users, enabling easy dissemination of the project results. The database can also be easily updated, allowing additional data to be incorporated on a continual basis.

Table 2.1 *Database contents by category, describing the information to be made available by species, by journal reference and by microsatellite locus. The database is designed to be fully cross-referenced to allow searching by multiple categories.*

Searchable category	**Variables**	**Format**
Species	6 target species	Drop-down menu
Journal reference	Individual references	Drop-down menu
Microsatellite locus	Species	Text
	Reference	Text
	Allele ladder	Table
	Allele frequencies	Table
	Validated PCR conditions	Table
	Inheritance data	
	Example genotypes	Pictures of homo & heterozygotes
	Reference samples	WDNAS database sample number
	Genbank Accession	Hyperlink to NCBI locus info

3. Results

3.1 Sample Collection

Over 1000 birds were sampled throughout the project (Table 1). In terms of absolute numbers, sampling targets were met for captive bred and wild caught birds in all species, however sourcing specific sample types was problematic.

Table 3.1 Summary of the samples collected during the project. 'Known unrelated' refers to birds known not to share any parents, 'known families' refers to a group consisting of two parents and a minimum of two offspring.

	Total	**Wild**	**Captive**	**Known unrelated**	**Known families**
Goshawk	305	202	103	22	2
Golden Eagle	117	57	60	21	2
Peregrine	152	51	101	20	2
Gyr	67	27	40	4	2
Saker	116	40	76	5	1
Merlin	251	216	35	18	2
Total	1008				

Captive bred birds

Despite the numbers of birds sampled, there remained a lack of known unrelated samples from captive bred gyr and saker falcons (Table 1). The situation was due in part to the high proportion of existing samples originating from birds imported into the UK, for which it was not possible to verify relatedness. These imported birds are often used repeatedly as breeding stock, resulting in many UK gyr and saker falcons being closely related.

In order to derive a large enough sample set for microsatellite assessment in these two species, birds were included that were known not to have been bred from the same stock within the UK, but which might have originated from the same importation stock.

Wild birds

For the four native species (goshawk, golden eagle, peregrine, merlin) a sufficient number of samples were obtained to allow a parallel study of genetic diversity to that of the captive bred collection. For gyr and saker falcons, wild samples were collected to examine the ability to differentiate between these closely related species.

Validation samples

Blood and feather samples from the same individual were collected in three species (golden eagle, gyr and saker) in order to validate the use of moulted feathers as a DNA source suitable for the reliable production of DNA profiles.

3.2 Assessment of existing microsatellite loci

Microsatellite assessment progressed in a series of examinations, with loci being excluded at each stage. A summary of this process is shown in 3.1, with full results detailed in Appendix II.

Amplification

All of the microsatellite markers amplified in the species for which they were originally designed, with the exception of those designed for the gyr falcon, published by Nesje & Roed (2000). Despite repeated optimization, several of these loci failed to yield PCR product. On examination of the original sequence data submitted to GenBank, it was revealed that the published primer combinations were incorrect, preventing amplification at these loci. Primers were subsequently redesigned based on the GenBank sequences.

Cross-amplification of loci in other target species was variable. Within the *Falco* genus, most loci cross amplified among all species, but very few *Falco* loci could be amplified in the golden eagle or goshawk (Table 3.1). In contrast, many of the eagle and goshawk loci did amplify in the falcons. Where sufficient loci were available for some species, further cross-amplification of all loci was not always attempted.

Polymorphism

Microsatellite polymorphism varied markedly between species. Within the falcons, several markers observed to be polymorphic in peregrine and saker falcons were found to be monomorphic in gyr falcons and merlins. Eagle and goshawk loci that cross-amplified in falcons were generally found to be monomorphic, although two of the goshawk loci (NAGE5 and NAGE7) were polymorphic and amplified particularly well (Table 3.2).

Table 3.2 Summary of microsatellite screening results showing the numbers of loci carried forward at each stage of the assessment. See Appendix X for detailed results.

Species	Number tested	Number amplified	Number polymorphic	No. showing no HWE deviation
Goshawk	19	16	10	10
Golden eagle	51	39	23	19
Peregrine	59	30	12	10
Gyr	45	26	12	12
Saker	27	21	13	7
Merlin	35	22	11	10

Population genetic characteristics

Genotypic data for several of the polymorphic loci revealed deviation from Hardy-Weinberg equilibrium (HWE), indicating the possible presence of null alleles or selection. This led to the exclusion of several polymorphic loci in each species, however in both peregrine and saker falcons, deviations from HWE were examined further in order to assess the possible causes.

In the peregrine falcon, microsatellites NVH02 and NVH89 both deviated from HWE, showing statistically significant heterozygote deficit (P<0.05). However given the number of observed heteroygotes (Table 3.3), it was considered likely that the observed deviation was due to insufficient sampling at hypervariable loci rather the presence of null alleles, and that the markers would therefore still be suitable for inclusion in the DNA profiling system. This conclusion is supported by the fact that marker 'Falco 89' was originally isolated in peregrine falcons and does not show any sign of null alleles when cross-amplified in sakers, gyrs or merlins.

Table 3.3: Heterozygosity data for markers falco 02 and falco 89 in the peregrine falcon, based on 36 samples and 39 samples respectively.

Marker	**No. alleles**	**Expected no. heterozygotes**	**Observed no. heterozygotes**	**P-value (1-tailed)**
Falco 02	18	33.27	30	0.0012
Falco 89	10	26.18	24	0.0189

In the saker falcon, seven of fourteen markers displayed significant heterozygote deficit (6 loci 0.01<P<0.05, 1 locus P<0.01). Given that six of these loci did not display departure from HWE in other falcon species, it was suspected that the results were an effect of the sample set used for screening. Testing for the Wahlund effect using the population assignment program, STRUCTURE (Pritchard 2000), revealed no evidence of population substructure within the sample dataset, however when the same set of loci were used to screen a set of ten wild saker falcons from Pakistan, no departures from HWE were observed.

As a further investigation of the possible presence of null alleles in the saker falcon, primers were redesigned for a single marker, Falco 54, which showed highly significant heterozygote deficit in saker falcons, but no HWE departure in peregrine or gyr falcons. The redesigned primers were based on the original peregrine falcon sequence available from GenBank. Applying the new primers in all combinations with the old primers yielded no additional alleles when amplifying twenty captive-bred saker falcons. These results led to the conclusion that six of the seven loci for which HWE deviation was originally observed should be carried forward to the forensic validation stage.

3.3 Forensic validation of microsatellite markers

The results of the developmental validation of the microsatellite loci selected for use in each species are documented in the database that accompanies this report. A set of example results is provided below, based on the validation of loci selected for the goshawk.

Across all species, two loci failed validation due to displaying apparent non-Mendelian inheritance: goshawk – locus NAGE6, *Falco* species – locus Msfp01. In each case the familial genotypes were consistent with the locus being located on the male 'Z' sex chromosome.

Example developmental validation results:

Protocol 1: Optimization and selection of positive validation controls

The following conditions were tested and used as optimal PCR conditions for all subsequent validation protocols:

PCR master mix (inc Taq):	ABgene
Magnesium concentration:	2mM
Primer concentration:	4pM
Reaction volume:	10ul
Annealing temperature:	55°C

A single annealing temperature was used for all loci due to their subsequent inclusion in a single multiplex reaction. Two positive controls were selected for inclusion in subsequent validation stages.

Protocol 2: Magnesium concentrations and thermocycling parameters

The range of magnesium concentrations and thermocycling parameters described in Protocol 2 were examined for each locus. The results of these tests were documented as images of agarose gels showing PCR product for each locus, e.g. Figures 3.1-3.3.

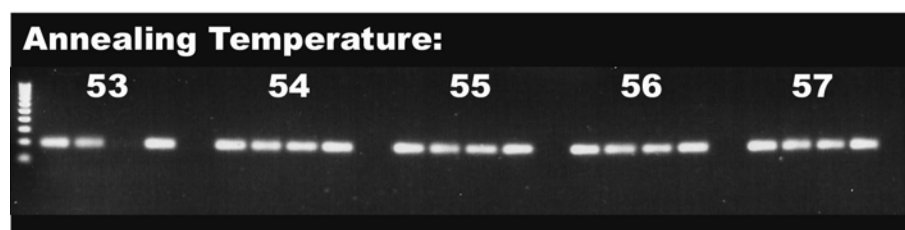


Figure 3.1: PCR results for locus NAGE 9 across a range of PCR annealing temperatures, around the optimum temperature of 55°C. Each temperature includes a –ve control, a +ve control and three test samples.

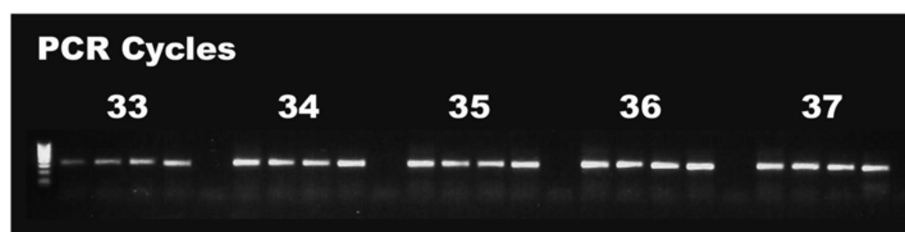


Figure 3.2: PCR results for locus NAGE 4 across a range of cycle numbers, around the optimum number of 35 cycles. Each cycle set includes a –ve contro,l a +ve control and three test samples. Note the increasing intensity of PCR product with cycle number.

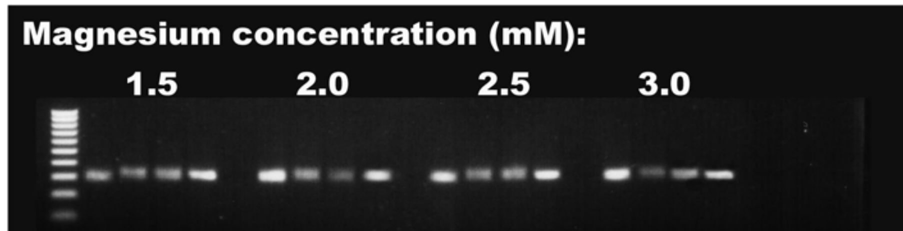


Figure 3.3: PCR results for locus NAGE 6 across a range of magnesium chloride concentrations, around the optimum temperature of 2mM. Each concentration includes a –ve contro,l a +ve control and three test samples.

All loci were found to work over the full range of values tested for each variable, validating these conditions as suitable for the amplification of the nine goshawk loci.

Protocol 3: Characterization of genetic markers

Each locus was tested to confirm that it displayed Mendelian inheritance by examining the genotypes of two family groups, each consisting of a mother, father and two offspring. Both family groups chosen for the validation study originated from samples belonging to DEFRA. The family relationships were provided with the samples.

From the genotypes generated for each family, deviation from Mendelian inheritance was only observed at locus NAGE6, as detailed in Table 3.4.

Table 3.4: Genotypes for the family members of locus NAGE 6 showing deviation from expected Mendelain (bi-parental) inheritance (*).

Locus	Individual		Sex	Allele 1	Allele 2
NAGE 6 (family 1)	Father	(837)	M	287	301
	Mother	(838)	F	291	291
	Offspring 1	(839)	M	291	301
	Offspring 2*	(840)	F	301	301
NAGE 6 (family 2)	Mother	(860)	F	301	301
	Father	(861)	M	297	311
	Offspring 1*	(862)	F	297	297
	Offspring 2	(863)	M	301	311

The results suggest are consistent with NAGE6 being sex-linked (i.e. positioned on the Z chromosome). To investigate this further, an additional six males and six females were genotyped for NAGE 6. Five of the six males were heterozygous, but all six females were homozygous. It is concluded that NAGE 6 is a sex-linked marker.

Protocol 4: Sensitivity study

The results of the sensitivity study show the level at which the quantity of DNA template begins to impact on the ability to successfully genotype each locus. The results presented in Table 3.5 show that the impact of reduced template DNA varies

among loci, indicating the difference in amplification strength of each optimized PCR. Although loading volumes for fragment analysis were controlled, second variable, dye strength, should also be taken into consideration when examining these results (intensity of blue>green>yellow).

Table 3.5: The effect of reducing DNA template on the success of genotyping PCR product. 'Y' = genotyping possible, 'N' = genotyping failure due to insufficient product, 'X' = genotyping failure due to failed PCR.

Locus	**Dye colour**	**Amount of DNA template (ng)**						
		10.00	5.00	2.50	1.25	0.62	0.31	0.15
AGE 1A	Blue	Y	Y	Y	Y	Y	Y	N
NAGE 1	Yellow	Y	Y	Y	Y	Y	Y	Y
NAGE 2	Blue	Y	Y	Y	Y	Y	Y	Y
NAGE 4	Green	Y	Y	Y	Y	Y	N	X
NAGE 5	Green	Y	Y	Y	Y	Y	Y	N
NAGE 6	Blue	Y	Y	Y	Y	Y	Y	Y
NAGE 7	Blue	Y	Y	Y	Y	Y	Y	Y
NAGE 9	Green	Y	Y	Y	Y	Y	Y	Y
NAGE 10	Yellow	Y	Y	Y	Y	Y	Y	Y

The results indicate that the majority of loci produce genotypes at a concentration of 0.15ng DNA. AGE1A and NAGE 5 should also work at this level if more PCR product is used for genotyping, however NAGE 4 did not amplify with this level of template. All loci were successfully validated using 0.62 ng of template DNA.

Protocol 5: Species specificity

Loci were tested for cross-amplification with three other species: golden eagle, barnacle goose and human. The results show cross-species amplification for the golden eagle in eight of the nine loci, amplification in the barnacle goose for NAGE 5 only and the amplification of a single band at around 450 base pairs in NAGE 2 for human DNA (Table 3.6). An understanding of cross amplification allows issues of contamination to be fully accounted for.

Table 3.6: Results of cross-species amplification tests in three non-target species.

Locus	Golden eagle	Barnacle goose	Human		Locus	Golden eagle	Barnacle goose	Human
AGE 1A	Y	N	N		NAGE 6	Y	N	N
NAGE 1	N	N	N		NAGE 7	Y	N	N
NAGE 2	Y	N	Y		NAGE 9	Y	N	N
NAGE 4	Y	N	N		NAGE 10	Y	N	N
NAGE 5	Y	Y	N					

Protocol 6: Stability studies and Case-type studies

The representative microsatellite used to amplify DNA recovered following various environmental treatments was successfully amplified in all conditions. As this microsatellite was of a larger size than the goshawk loci selected, these results suggest that the goshawk loci should be successfully amplified following similar environmental degradation of sample material.

Protocol 7: Precision, Accuracy, and Reproducibility

Results for reproducibility were produced by conducting separate PCR reactions using the same sample four times. The resulting genotypes were assessed for differences in exact fragment size. The mean estimated fragment sizes and standard deviation for each allele at each locus was recorded (Table 3.7).

Table 3.7: Mean fragment sizes and standard deviation for each allele of a single sample that has undergone four separate PCR reactions.

Locus	Allele 1		Allele 2	
	Mean	Std dev.	Mean	Std dev.
AGE 1A	161.436	0.201	169.735	0.201
NAGE 1	227.483	0.111	280.363	0.120
NAGE 2	203.360	0.034	223.470	0.056
NAGE 4	318.473	0.296	337.757	0.104
NAGE 5	161.250	0.042	170.355	0.049
NAGE 6	287.270	0.026	301.633	0.023
NAGE 7	220.730	0.018	232.300	0.014
NAGE 9	187.313	0.064	189.387	0.058
NAGE 10	165.925	0.173	171.828	0.177

The highest standard deviation recorded (s.d.=0.296, NAGE 4) is considered low enough to prevent mis-scoring when applying automated genotyping software to the data and indicates that all peaks produced for this locus can be confidently assigned to a particular allele by manual scoring.

3.4 Probability of identity (PI) calculations

Calculation of the Probability of Identity (PI) for the validated loci in each species indicated that the profiling systems generate sufficient statistical power to enable positive identification of individual birds, when all birds are assumed to be unrelated (Table 3.8).

When population substructure and inbreeding parameters were included in the PI calculations (F and R > 0), the probability increased across all species, but remained lower than 10^{-6} for all species except the gyr falcon (10^{-5}) (Table 3.8).

Subsequent inclusion of parameter R, modeling familial relatedness between individuals in a population, further increased the PI. When 20% of the population

were assumed to be siblings and a further 20% half-siblings (R=0.2), PI values increased to around 10^{-5} in five species, with gyrs again higher still (~10^{-4}).

Table 3.8 *Probability of identity (PI_{ave}) generated by DNA profiling systems in each species, under varying levels of population substructure (F_{ST}), inbreeding (F_{IS}) and relatedness between individuals (R), described as the proportion of sibs and half sibs in the population. Estimates for F_{ST} and F_{IS} are based on those calculated from other non-human profiling systems for which data is available. The values of F_{ST} and F_{IS} should be treated as indicative estimates only.*

Species	**No. of loci in profile**	**Probability of identity (PI_{ave})**		
		F_{ST} & F_{IS} = 0, R = 0	**F_{ST} & F_{IS} = 0.1, R = 0**	**F_{ST} & F_{IS} = 0.1, R = 0.2**
Golden Eagle	11	7.63E-10	5.10E-08	1.57E-06
Peregrine	12	7.25E-09	2.00E-07	4.10E-06
Gyr	11	6.99E-07	1.04E-05	9.90E-05
Saker	11	2.10E-09	2.09E-07	6.45E-06
Merlin	10	2.00E-08	4.93E-07	9.62E-06
Goshawk	8	2.05E-10	7.14E-08	3.69E-06

3.5 Development of a PCR multiplex for each species

On the basis of the microsatellites selected for the DNA profiles, a single PCR multiplex was designed and optimised for the goshawk loci and two multiplexes produced for the each of the remaining five species (Table 3.9; Appendix IV).

Table 3.9 *Microsatellites loci chosen selected for the PCR multiplexes in each species. See Appendix II for original locus references and Appendix IV for details of allele size ranges and selected dye colours.*

Multiplex	**Goshawk**	**G. eagle**	**Peregrine**	**Gyr**	**Saker**	**Merlin**
1	AGE1A	AA02	UFPE01	UFPE01	UFPE01	UFPE01
	NAGE1	AA15	NVH13	NVH13	NVH13	NVH31
	NAGE2	AA26	NVH31	NVH46	NVH31	NVH82
	NAGE4	AA39	NVH46	NVH89	NVH89	NVH92
	NAGE5	AA43	NVH86	NAGE5	NAGE5	NAGE5
	NAGE7		NVH89	NAGE7	NAGE7	ND58
	NAGE9		NAGE5			
	NAGE10					
2		AA04	UFPE02	NVH34	UFPE02	UFPE02
		AA27	NVH34	NVH54	NVH54	NVH46
		AA36	NVH54	NVH79-4	NVH79-4	NVH79-4
		NVH142	NVH79-4	NVH82	NVH82	NVH89
		IEEEA04	NVH92	NVH92	NVH92	
		IEEEA15				

Optimization of the PCR reaction conditions for each of these multiplexes was successful, resulting a set of PCR reagent conditions for use in each multiplex (Appendix V). However it should be noted that obtaining equal strength amplification of all loci in each multiplex is difficult, with results being highly sensitive to initial primer concentration and dye characteristics. Examples of each multiplex profile are provided in Appendix VI.

3.6 Validation of moulted feathers as a DNA source

The two sets of DNA profiles generated from moulted feather and blood samples were identical, supporting the use of moulted feather material as a sample source for DNA profiling.

3.7 Development of a genetic sexing test

PCR amplification fragments generated using the new primers clearly demonstrate the ability to distinguish male and female birds (Figure 3.4). Results of blind trials all resulted in the correct sex being assigned to test samples in each species.

Figure 3.4 *Results of the sex determination test for each species, showing a single DNA band in males and two bands in females. The higher band (present in both sexes) represents a DNA fragment around 580 base pairs long; the lower band (only in females) is around 380 base pairs.*

Although the size of the largest PCR product (~580bp) is longer than desired given the potential use of the test on degraded DNA, it was not possible to redesign the primers to amplify a shorter region of DNA due to the species specific nature of the sequence at this locus. However, it should be noted that as the female (W-chromosome) fragment is shorter than the Z-chromosome fragment, incorrect sex-determination due to DNA degradation should not occur (see Figure 1).

Given the size of the two sexing alleles, it is possible to analyse the PCR product for this locus together with the PCR product for the DNA profiles ('poolplexing'), reducing the cost of a combined test. Alternatively, the sex test could be poolplexed with the species identification test described in Section 3.8. Incorporation of the sex test into PCR multiplexes (either with profiling or species id) was not investigated. In addition to offering a sex-determination test, inter-specific variation in the size of the larger fragment provides the basis for an independent species identification test (see Section 3.7).

3.8 Development of a species identification test

Genetic variability among species
Results of the preliminary assessment of interspecific variability at mitochondrial DNA (mtDNA) loci revealed that these regions were suitable for the differentiation of all species except gyr and saker falcons. Subsequently, design of the species identification test was divided into two phases: i) the design of a test to distinguish between goshawk, golden eagle, peregrine, merlin and gyr/saker falcons using mtDNA loci, and ii) an investigation into methods for genetic differentiation of gyr and saker falcons using nuclear DNA markers.

i) Species identification test using mtDNA
Based on the initial assessment of mtDNA sequence diversity, the cytochrome oxidase I (COI) gene was chosen as the target locus for the species identification test. The test was designed to amplify species-specific PCR products of varying length, enabling the five species groups to be distinguished using a single PCR multiplex reaction. The PCR reaction includes one forward primer that attaches to the same region in all target species and five different reverse primers that are attached specifically to different regions in each of the target species. The details of the PCR primer sites in each species are given in Appendix VII. During the species identification test, the forward primer and only one of the five reverse primers will attach, depending on the species present. This results in the PCR amplification of a single DNA region of a species-specific size (Figure 3.5).

Fluorescently labelling of the forward primer allows the test to be run under fragment analysis on an automated sequencer (the same technique as DNA profiling). Owing to the difference in primer sites and fragment lengths, the test would also be suitable for conversion onto a real-time PCR platform.

ii) Distinguishing gyr and saker falcons using nuclear loci
The inability to distinguish gyr and saker falcons using mtDNA raises questions regarding the validity of these species (see Conclusions), however given their distinct geographic distributions (circumpolar and central Asian, respectively), it was considered likely that genetic differentiation of these species would be possible at some level. The nuclear markers developed for these species elsewhere in this project were therefore examined for their utility in gyr/saker discrimination.

Nuclear sexing locus:
During the development of the sex determination test, inter-specific variation in the size of the larger fragment was observed, suggesting that this sexing marker may also allow species identification. The larger fragment is amplified in both sexes and was found to vary in size by at least 20bp in all species, with the exception of gyr and saker falcons, which had fragments of identical size.

Nuclear microsatellites:
None of the microsatellite loci originally assessed as part of the individual DNA profiling work displayed discrete difference in allele size between gyr and saker falcons.

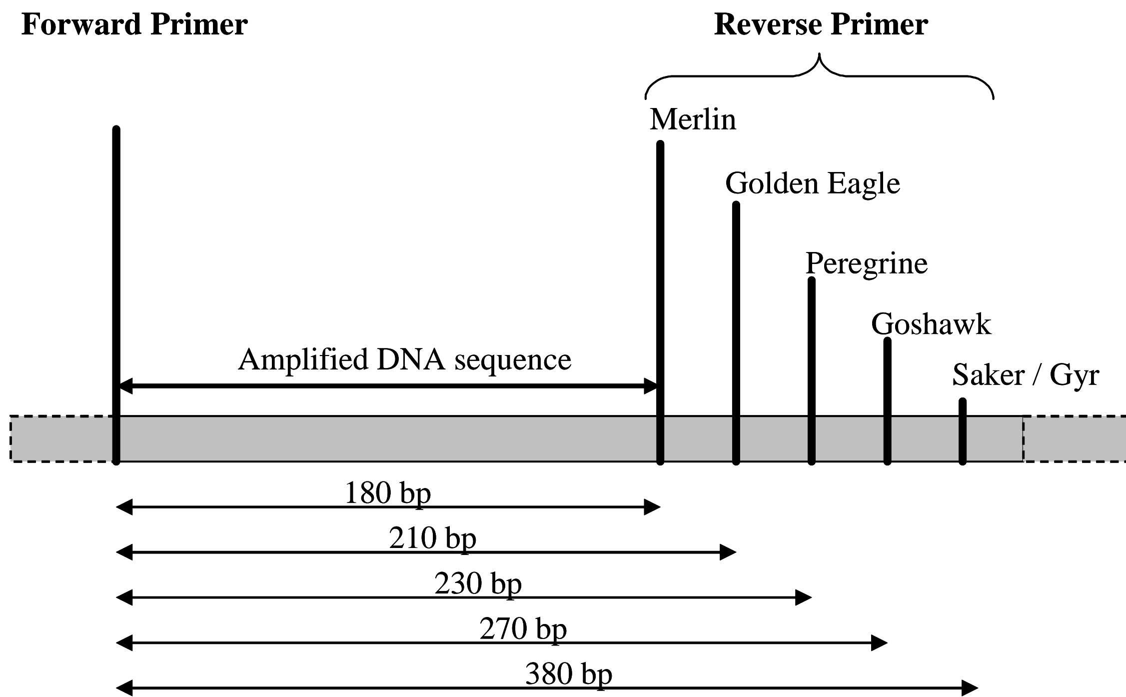


Figure 3.5: *Diagram of the species identification test. In a PCR reaction, two primers are used together as a pair, delimiting the beginning and end of the target DNA sequence. The primers are designed to attach to specific parts of this target sequence; one is known as the 'forward primer', the other the 'reverse primer'. The grey bar represents a section of the mtDNA COI region that is targeted during the PCR reaction. The inclusion of five different reverse primers will lead to the production of five different sized fragments depending on the species of bird present.*

The result of species assignment using allele frequencies was more encouraging. Analysis with STRUCTURE showed that individuals were correctly assigned to species in both wild and captive population samples more than 95% of the time. The establishment of training sets based on allele frequencies from wild birds allowed blind test samples to be correctly assigned using the program, suggesting this as a possible method for species discrimination. Such a method should be useful for allowing choice of profiling system, but could not be used as a forensic species identification test.

3.9 Production of a microsatellite database

The database is now available as a Microsoft Access .mdb file from the authors. Please contact enquiries@wdnas.com for a free copy. The database is designed to allow addition of further loci to account for the future design and validation of new markers. Submission of new markers to the database is encouraged.

4. Discussion

4.1 Individual Identification – DNA Profiling

Sample Collection

The collection target of twenty unrelated captive bred birds of each species was set in order to allow a comprehensive assessment of the microsatellite loci under assessment. This target proved difficult to achieve in both the gyr and saker falcons, where the availability of suitable birds with sufficient pedigree information was limited. The combination of importation from unknown sources and the relatively small number of UK breeders supplying the market, meant that the sample sets in these species could not be certified to contain solely unrelated individuals. The implications of this for the microsatellite assessment are that loci may exhibit lower allelic variability and heterozygosity than would be found in a totally unrelated captive group, or in natural populations. This would generally have a conservative effect on the statistical power of the DNA profiles calculated in this report and should not invalidate any of the project findings.

The limited number of gyr and saker breeders in the UK does have implications for the levels of genetic variability present in the UK captive populations. The combination of a relatively small founder population and repeated use of stud males for breeding is likely to have resulted in a significant reduction in genetic diversity in these species compared to wild populations. In addition, the UK captive populations are likely to have relatively high levels of inbreeding. This in turn affects the ability to use genetic markers to discriminate between individuals as discussed below.

Microsatellite Loci

Of the 100 microsatellite loci assessed across the six species, relatively few (n=31) were eventually used in the six DNA profiling systems. While an excess of potential loci existed for the golden eagle, those available for the goshawk and falcon species were extremely limited. It should be noted that none of the microsatellite loci developed during the preceding DETR project (CR0202) were found to be suitable for inclusion in the profile systems developed here.

The limited number of loci from which to construct the DNA profiling systems resulted in certain compromises being made in the choice of loci used. Forensic microsatellite (STR) profiling systems are ideally composed of loci displaying a tetranucleotide repeat (or greater), in order to remove possible ambiguity in allelic identification. Similarly, the use of compound microsatellites is avoided where possible to simplify the construction of allelic ladders and prevent mistyping. The DNA profiling systems developed here include both dinucleotide loci and compound microsatellites through necessity. While the use of these markers does allow the production of valid DNA profiles, accurate genetic analysis requires a significant level of genotyping experience and careful use of positive controls. Furthermore, the number of available loci has somewhat limited the statistical strength of the DNA profiling systems developed.

With the continual isolation of new microsatellites, the opportunity exists to update the DNA profiling systems, either through addition of markers, or exchange of new markers with the weakest loci currently employed. These processes would be

relatively rapid and inexpensive to perform and the future addition of microsatellite loci to the profiling systems presented here is therefore recommended.

The results of the validation studies provided a great deal of information necessary for the application of the microsatellite loci in a forensic context. The ability to recover DNA and generate profiles from a number of sources following various environmental treatments demonstrates the utility of the markers in a range of case-type scenarios. The range of validated PCR conditions allows laboratories to apply the profiling systems with some flexibility. Data on amplification variation with DNA concentration and data on possible cross-species amplification will aid laboratory interpretation of results. The identification of loci exhibiting non-Mendelian inheritance was also a crucial result of this study as this information had not been supplied when the markers were published.

Statistical Considerations

Forensic DNA profiling may be applied on the basis of exclusion or positive identification. Exclusion occurs where two DNA profiles are incompatible with each other and therefore the identification of an individual or family relative can be refuted. This type of application is relatively straightforward to perform as the forensic scientist essentially only needs to have information concerning the two DNA profiles being compared. In contrast, positive identification, e.g. the matching of a bird's feather to a blood sample, requires information concerning the genetic characteristics of the population to which the bird is presumed to belong, as well as the two DNA profiles being compared. The population information required for positive identification consists of population allele frequencies for each of the loci in the DNA profile.

The DNA profiling systems developed during this project may be used for either exclusion or positive identification, however the scope of the current project did not include the production of an allele frequency dataset that would be required for positive identification in a legal context. At this time, the profiles can therefore only be used for exclusion and to infer positive identification with a high degree of confidence. In order to generate an allele frequency dataset, it will be necessary to genotype additional UK captive bred birds. Ideally, the dataset should consist of at least 100 birds per species.

The ability to discriminate between individual birds using the DNA profiling systems developed in this project was simulated by calculating the average probability of identity (PI) for each species. The PI estimates the probability of observing two identical genotypes in a population. In order to account for population substructure, inbreeding and relatedness between individuals, further simulations were run incorporating these parameters. The PI values suggested that on average, the largest PI value (weakest evidence) would be less than 0.0001 (10^{-4}) for gyr falcons, with those for other species being below 0.00001 (10^{-5}). Despite these values being substantially higher than those generally used in human DNA profiling ($\sim10^{-9}$), they indicate that the profiles will provide strong evidence for individual identification in court, particularly when considering the size of the captive populations (<10.000).

The use of DNA profiles to positively match two samples for forensic investigation requires the calculation of match probabilities (rather than a PI) and the construction a

likelihood ratio. Match probabilities are typically used to estimate the probability that two samples chosen at random from the population have the same DNA profile. The equation incorporates information on population substructure as well as population allele frequencies. A likelihood ratio is the standard method used in the UK for presenting DNA profile evidence and is based on the principle of evaluating the relative likelihood of two competing hypotheses, usually referred to as the defence hypothesis (H_d) and the prosecution hypothesis (H_p). Methods for calculating match probabilities and constructing likelihood ratios are not discussed in depth here (see Evett and Weir (1998) for further information), but some important points are raised concerning the differences between their use for DNA profiling of humans compared to birds of prey.

Match probability calculations include a parameter to account for population substructure known as theta. Theta is estimated from population genetic data and is a measure of the genetic differentiation between subpopulations within the total population sampled. In human profiling, theta values used in the UK are generally around 0.01, with 0.03 sometimes used as a conservative value where insufficient data is available to make a precise estimate. For birds of prey, the estimation of theta is more complicated. If all the captive UK birds of prey of any one species are assumed to belong to a single interbreeding population, then theta would be set to zero. This is clearly an invalid assumption as population substructure generated by independent breeding stocks will exist. However, given the artificial constraints over gene exchange between breeding stocks in the UK, it is not clear how captive birds could be divided into meaningful subpopulations in order to estimate theta.

A previous study examining the same problem in captive bred dogs based its estimate of theta on measuring F_{ST} between breeds of dog (Halverson and Basten 2005). Although a similar subspecies classification is not available for bird species, the resulting theta value (0.11) for dog breeds does provide a reference point of some sort. Another option would be to refer to F_{ST} values calculated for wild bird of prey populations, based on the same microsatellite loci. While this would allow a more meaningful comparison than that with dog profiling, there is no reason why levels of substructure in wild and captive populations should be closely related. The production of accurate theta values for captive bred UK bird of prey species therefore remains problematic.

One possible approach to this issue is to make no direct estimation of population substructure (theta value) but instead assume that all the individuals in the population are related to the level of full siblings (Waits et al. 2001). This reduces the match probability to an expression for the probability that samples originating from two siblings have the same profile. This is clearly an extremely conservative approach and the use of such an equation will weaken the evidence for a profile match considerably (e.g. PI_{SIBS} for gyr falcons = 10^{-3}). However use of a sibling match probability does provide a framework for evaluating DNA profile evidence and has been successfully applied to individual identification of wildlife samples in the USA (Waits et al. 2001).

A further issue relating to the calculation of match probabilities arises where the DNA profile is being applied to a dispute between a breeder who claims a bird is captive bred and the prosecution who claim the bird was stolen from a wild nest site. In this

situation it is necessary to decide how to choose the appropriate allele frequency dataset and theta value where there are potentially two totally isolated breeding populations with two sets of allele frequencies and different levels of subpopulation structure. Where the wild and captive bred allele frequencies differ substantially, the choice of allele frequency dataset may have a significant impact on the match probability generated. Given that the likelihood ratio framework requires two hypotheses to be evaluated (H_d and H_p), we suggest using different sets of allele frequencies for generating the different match probabilities relevant to each hypothesis.

4.2 Sex Determination

The sex determination test offers a straightforward method for sexing birds in all six species. In addition to revealing gender, the test effectively provides an additional marker for subsequent individual match probability calculations, increasing the power of the DNA profiles to discriminate between individuals. However it should be noted that the usual assumption of 1:1 sex ratios may not hold for captive bred birds and therefore data on actual sex ratios should be sought in order to use this marker in probability calculations.

The size of the DNA fragments produced in the test allow the PCR product to be poolplexed together with either the microsatellite profiles or the species identification test, reducing time and costs associated with the test. One drawback of the system is that the size of the target sequence may limit its amplification in degraded samples, however the design of the test will prevent misidentification of sex in such cases.

4.3 Species Identification

Application of the Species Identification Test
The species identification test will allow an informed choice of DNA profiling system to be made in situations where the sample is known to have come from one of the target species, but it is not known which. In practice, this situation is likely to be quite rare, with the species either known, or unknown, in which case universal primers would be employed and the species identified by nucleotide sequencing. However, the work on species identification has produced interesting results regarding the species status of gyr and saker falcons

Species status of gyr and saker falcons
Given the degree of similarity in the sequence data at all three mitochondrial gene regions between gyr and saker falcons, the status of *Falco rusticolus* and *Falco cherrug* as discernable species might be questioned. Historically identified as a single species (*Falco hierofalco*), the gyr and saker falcons appear to be difficult to separate genetically, with other studies indicating that sakers are themselves paraphyletic with respect to Lanner falcons (Wink et al. 2004) Although this project has demonstrated that genetic species assignment may be possible from a research standpoint, there do not appear to be sufficient interspecific sequence differences from which to design a robust forensic species identification test. Following communication with researchers in the US who have sequenced larger regions of mtDNA in these species (J. Wells, U.

Michigan, pers comm.), it appears that a single nucleotide polymorphism (SNP) in cytochrome b has been found to be conserved between sakers and gyrs. However a single SNP is considered to be an insufficient basis for the species test required here.

It is not within the remit of the current project to assess the species status of gyr and saker falcons, however the data gathered here does indicate that there is very little genetic difference between them. This is clearly of relevance to future conservation measures and enforcement of wildlife regulations.

Hybrid Identification

The identification of hybrid birds and determination of their species ancestry has been a problem in bird registration in general and of enforcement in particular. Although this project did aim to investigate methods for the identification and characterisation hybrid birds, this was always likely to be difficult given the lack of any clear theoretical framework for doing so. Mitochondrial DNA markers only carry information concerning the maternal parent, preventing hybrid identification. Nuclear DNA is only likely to detect hybridisation in the previous generation and this approach suffers from the lack of common species-specific nuclear DNA loci.

Given this situation, the results of the sex determination test were quite interesting, as they allow for the identification of the paternal parent through variation in fragment length on the Z chromosome. When combined with mitochondrial species identification this allows identification of both parental species (and thus F1 hybrids). However, this technique would not allow complex or second generation hybrids to be confidently identified, nor would it discriminate between gyr and saker falcons or therefore identify gyr x saker hybrids.

Conclusions

- DNA profiling systems were successfully developed for each of the six target species using validated microsatellite DNA markers.

- PCR multiplexes were designed for the simultaneous amplification of fluorescent labelled markers in each profiling system.

- Data regarding the microsatellite loci and their associated validation results are available in a searchable database allowing simple dissemination of DNA profiling methods to end users.

- The power of the profiling systems to identify individual birds was assessed based on the probability of identity, an average measure of the likely match probability. Values are considered high enough to provide strong evidence of identity in a legal framework.

- Allele frequency datasets for captive populations in each species should be increased prior to use in match probability equations.

- All six of the profiling systems may be considered as initial platforms from which to improve the techniques, either through the addition of further markers or exchange of markers.

- A genetic sex determination test was devised for the six target species. This test can be screened together with either the DNA profiling or species identification tests.

- A species identification test based on the mitochondrial cytochrome oxidase I gene was designed to discriminate between all target species with the exception of gyrs and sakers. Microsatellite results suggest that gyr and saker falcons may be distinguished using allele frequency assignment methods.

Acknowledgements

The authors would like to thank everyone who contributed to the project, from sample collection through to statistical analysis. In particular, we are grateful to the following: Nick Williams (DEFRA Chief Wildlife Inspector and project manager) and his team in Bristol; Prof. Mike Bruford (Cardiff University and scientific consultant to the project); Dr. Jon Wetton (Forensic Science Service); Dr. Nick Fox (International Wildlife Consultants); Jemima Parry-Jones (National Bird of Prey Centre); Johan Ekenstedt (Projekt Jaktfalk, Sweden); Prof. Terry Burke's group (University of Sheffield); Dr. Chris 'Shep' Bryant; Helen Briggs and Liz Heap (WDNAS).

References

Ayres KL and Overall ADJ (2004) API-CALC 1.0: a computer program for calculating the average probability of identity allowing for substructure, inbreeding and the presence of close relatives. *Molecular Ecology Notes,* **4,** 315-318.

Busch JD, Katzner TE, Bragin E, Keim P. 2005. Tetranucleotide microsatellites for *aquila* and *haliaeetus*. *Molecular Ecology Notes* **5**:39-41.

Dieringer D & Schlötterer C (2002) Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. Molecular Ecology Notes 3 (1), 167-169

Double MC, Dawson D, Burke T, Cockburn A. 1997. Finding the fathers in the least faithful bird: a microsatellite-based genotyping system for the superb fairy-wren Malurus cyaneus. *Molecular Ecology* **6**:691-693.

Evett IW & Weir BS (1998) Interpreting DNA evidence: statistical genetics for forensic scientists. Sinauer Associates, Sunderland, MA, USA.

Fridolfsson AK, Ellegren H. 1999. A simple and universal method for molecular sexing of non-ratite birds. *Journal of Avian Biology* **30**(1):116-121.

Gautschi B, Tenzer I, Muller JP, Schmid B. 2000. Isolation and characterization of microsatellite loci in the bearded vulture (Gypaetus barbatus) and cross-amplification in three old world vulture species. *Molecular Ecology* **9**:2193-2195.

Griffith SC, Stewart IRK, Dawson DA, Owens IPF, Burke T. 1999. Contrasting levels of extra-pair paternity in mainland and island populations of the house sparrow (*Passer domesticus*): is there an 'island effect'? *Biological Journal of the Linnean Society* **68**:303-316.

Halverson J, Basten C. 2005. A PCR multiplex and database for forensic DNA identification of dogs. *Journal of Forensic Science* **50**(2):1-12.

Hanotte O, Zanon C, Pugh A *et al.* 1994. Isolation and characterization of microsatellite loci in a passerine bird: the reed bunting Emberiza schoeniclus. *Molecular Ecology* **3**:529-530.

Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM. 2004. Identificationof birds through DNA barcodes. *Public Library of Science* **2**(10):1657-1663.

Martinez JG, Soler JJ, Soler M, Moller AP, Burke T. 1999. Comparative population structure and gene flow of a brood parasite, the great spotted cuckoo (Clamator glandarius), and its primary host, the magpie (Pica pica). *Evolution* **53**:269-278.

Martinez-Cruz B, David VA, Godoy JA, Negro JJ, O'Brien J, Johnson WE. 2002. Eighteen polymorphic microsatellite markers for the highly endangered Spanish imperial eagle (Aquila adalberti) and related species. *Molecular Ecology Notes* **2**:323-326.

Mira S, Billot C, Guillemaud T, Palma L, Cancela ML. 2002. Isolation and characterization of polymorphic microsatellite markers in Eurasian vulture Gyps fulvus. *Molecular Ecology Notes* **2**:557-558.

Nesje M, Roed KH, Lifjeld JT, Lindberg P, Steen OF. 2000. Genetic relationships in the peregrine falcon (Falco peregrinus) analysed by microsatellite DNA markers. *Molecular Ecology* **9**:53-60.

Nesje M, Roed KH. 2000. Microsatellite DNA markers from the gyrfalcon (Falco rusticolus) and their use in other raptor species. *Molecular Ecology* **9**:1438-1440.

Nittinger F, Haring E, Pinsker W, Wink M, Gamauf A. 2005. Out of Africa? Phylogenetic relationships between Falco biarmicus and the other hierofalcons (Aves: Falconidae). *Journal of Zoology, Systematics and Evolutionary Research* **43**(4):321-331.

Parkin and Peck unpublished DEFRA Report. Feathers as a source of DNA CR0202: The isolation and testing of microsatellite markers from some birds of prey.

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

Raymond M, Rousset F (1995) Genepop, Institut des Sciences de l'Evolution. Universite de Montpellier, France.

Richardson DS, Jury F, Dawson D, Salgueiro P, Komdeur J, Burke T. 2000. Fifty Seychelles warbler (Acrocephalus sechellensis) microsatellite loci polymorphic in Sylviidae species and their cross-species amplification in other passerine birds. *Molecular Ecology* **9**:2226-2231.

Sonsthagen SA, Talbot SL, White CM. 2004. Gene flow and genetic characterization of Northern Goshawks breeding in Utah. *The Condor* **106**:826-836.

Topinka JR, May B. 2004. Development of polymorphic microsatellite loci in the northern goshawk (Accipiter gentilis) and cross-amplification in other raptor species. *Conservation Genetics* **5**(6):861-864.

Waits LP, Luikart G, Taberlet P. 2001. Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Molecular Ecology* **10**:249-256.

Wink M, Sauer-Gurth H. 2000. Advances in the molecular systematics of African raptors. In: Chancellor RD, Meyburg B-U, editors. *Raptors at Risk*. Berlin/Blaine: Hancock House/WWGBP. pp 135-147.

Wink M, Sauer-Gurth H, Ellis D, Kenward R. 2004. Phylogenetic relationships in the hierofalco complex (saker-, gyr, lanner, laggar falcon). In: Chancellor RD, Meyburg B-U, editors. *Raptors Worldwide*. Budapest/Berlin: MME/WWGBP. pp 499-504.

Appendix I - List of people and organisations who donated samples to the project

Contact Surname	Contact First Name	Organisation
Andriunas	Karen	Devon Bird of Prey Centre
Barnaby	Jonathan	DEFRA
Bourke	Brian	University of Nottingham/Sheffield
Burden	Terry	Kentish Falconry & Conservation Centre
Cole	R.	Private
Cookson	Steve & Mike	Private
Dillane	Eileen	University College Cork, Ireland
Downing	Steve	West Yorkshire Police
Ekenstedt	Johan	Sweden
Evans	Simon	Simon Evans
Fox	Nick	Falcon Research Institute
France	Janice	Welsh Hawk Board
French	Jim	Private
Frost	R. A.	Private
Garland	Ian	Raptor Propagation & Research Farm
Henderson	Malcolm	Private
Jackson	Nick	Welsh Mountain Zoo
Jones	Dave	Welsh Hawking Club
Jones	Martin	Private
Kanes	Mick	Private
Kirby	Mr.	Private
Knowles-Brown	Andrew	Private
Lindberg	Peter	University of Gothenburg
MacDonald	Kim	Guild of Taxidermists
Masden	Amy	University of Nottingham
Millar	Annie	National Bird of Prey Centre
Moores	S. F.	Private
Morris		British Hawking Association, East Sussex
Morton	Keith	RSPB, Edinburgh
Murn	Campbell	Hawk Conservancy
Mussard	Mussard	DEFRA known breeders
Norman	Wilf	Private
Norman	David	Private
O'Toole	Lorcan	Ireland Golden Eagle Project
Parry Jones	Jemima	National Bird of Prey Centre
Prytherch	Robin	Private
Rix	Mr.	Private
Roe	Peter	Private
Samworth	Stephen	Private
Scott	Peter	Zoo & Aquatic Veterinary Group, Winchester
Signer	Esther	University of Leicester
Smith	Ken	Hawks Lodge, Dunlavin
Stirling-Aird	PK	Scottish Raptor Study Groups
Tarr	D. J.	Private
Walker	Lee	CEH Monkwood
Walker	James	LGC
Watson	Mark	Private
Wetton	Jon	University of Nottingham, DEFRA stocks (FSS) Forensic Science Service

Appendix II – List of all microsatellite loci assessed during the project with summary results by species

	Locus	Species	Authors	Journal	Year	Page_Vol	Amplification					
							GE	Gos	Per	Sak	Gyr	Mer
1	Aa 02	Imperial eagle	Martinez-Cruz et al.	Molecular Ecology Notes	2002	2: 323-326	YES	X	YES	YES	YES	YES
2	Aa 04	Imperial eagle	Martinez-Cruz et al.	Molecular Ecology Notes	2002	2: 323-326	NO	X	NO	NO	NO	NO
3	Aa 11	Imperial eagle	Martinez-Cruz et al.	Molecular Ecology Notes	2002	2: 323-326	POOR	X	POOR	X	X	X
4	Aa 12	Imperial eagle	Martinez-Cruz et al.	Molecular Ecology Notes	2002	2: 323-326	POOR	X	YES	X	X	X
5	Aa 15	Imperial eagle	Martinez-Cruz et al.	Molecular Ecology Notes	2002	2: 323-326	YES	X	YES	YES	YES	X
6	Aa 26	Imperial eagle	Martinez-Cruz et al.	Molecular Ecology Notes	2002	2: 323-326	YES	X	NO	POOR	POOR	POOR
7	Aa 27	Imperial eagle	Martinez-Cruz et al.	Molecular Ecology Notes	2002	2: 323-326	YES	X	POOR	NO	NO	NO
8	Aa 35	Imperial eagle	Martinez-Cruz et al.	Molecular Ecology Notes	2002	2: 323-326	YES	X	X	X	X	X
9	Aa 36	Imperial eagle	Martinez-Cruz et al.	Molecular Ecology Notes	2002	2: 323-326	YES	X	X	X	X	X
10	Aa 39	Imperial eagle	Martinez-Cruz et al.	Molecular Ecology Notes	2002	2: 323-326	YES	X	X	X	X	X
11	Aa 41	Imperial eagle	Martinez-Cruz et al.	Molecular Ecology Notes	2002	2: 323-326	POOR	X	NO	X	X	X
12	Aa 43	Imperial eagle	Martinez-Cruz et al.	Molecular Ecology Notes	2002	2: 323-326	YES	X	X	X	X	X
13	Aa 49	Imperial eagle	Martinez-Cruz et al.	Molecular Ecology Notes	2002	2: 323-326	YES	X	X	X	X	X
14	Aa 50	Imperial eagle	Martinez-Cruz et al.	Molecular Ecology Notes	2002	2: 323-326	POOR	X	NO	X	X	X
15	Aa 51	Imperial eagle	Martinez-Cruz et al.	Molecular Ecology Notes	2002	2: 323-326	YES	X	NO	X	X	X
16	Aa 53	Imperial eagle	Martinez-Cruz et al.	Molecular Ecology Notes	2002	2: 323-326	YES	X	YES	X	X	X
17	Aa 56	Imperial eagle	Martinez-Cruz et al.	Molecular Ecology Notes	2002	2: 323-326	YES	X	NO	X	X	X
18	Aa 57	Imperial eagle	Martinez-Cruz et al.	Molecular Ecology Notes	2002	2: 323-326	YES	X	YES	X	X	X
19	BV11	Bearded vulture	Gautschi *et al.*	Molecular Ecology	2000	9: 2193 - 2195	YES	X	NO	X	X	X
20	BV12	Bearded vulture	Gautschi *et al.*	Molecular Ecology	2000	9: 2193 - 2195	NO	X	NO	X	X	X
21	BV13	Bearded vulture	Gautschi *et al.*	Molecular Ecology	2000	9: 2193 - 2195	YES	X	NO	X	X	X
22	BV16	Bearded vulture	Gautschi *et al.*	Molecular Ecology	2000	9: 2193 - 2195	YES	X	NO	X	X	X
23	GF11A4	Eurasian vulture	Mira *et al.*	Molecular Ecology Notes	2002	2: 557 - 558	YES	X	NO	X	X	X
24	GF3F3	Eurasian vulture	Mira *et al.*	Molecular Ecology Notes	2002	2: 557 - 558	YES	X	YES	X	X	X
25	GF3H3	Eurasian vulture	Mira *et al.*	Molecular Ecology Notes	2002	2: 557 - 558	YES	X	NO	X	X	X
26	GF8G1	Eurasian vulture	Mira *et al.*	Molecular Ecology Notes	2002	2: 557 - 558	YES	X	YES	X	X	X
27	GF9C1	Eurasian vulture	Mira *et al.*	Molecular Ecology Notes	2002	2: 557 - 558	YES	X	NO	X	X	X
28	NAGE 1	Goshawk	Topinka & May	Conservation Genetics	2004	5 (6): 861-864	X	YES	X	X	X	X
29	NAGE 2	Goshawk	Topinka & May	Conservation Genetics	2004	5 (6): 861-864	X	YES	X	X	YES	X
30	NAGE 4	Goshawk	Topinka & May	Conservation Genetics	2004	5 (6): 861-864	X	YES	X	X	NO	X
31	NAGE 5	Goshawk	Topinka & May	Conservation Genetics	2004	5 (6): 861-864	X	YES	YES	YES	YES	YES
32	NAGE 6	Goshawk	Topinka & May	Conservation Genetics	2004	5 (6): 861-864	X	YES	X	X	NO	X
33	NAGE 7	Goshawk	Topinka & May	Conservation Genetics	2004	5 (6): 861-864	X	YES	YES	YES	YES	YES

	Locus	Species	Authors	Journal	Year	Page_Vol	Amplification					
							GE	Gos	Per	Sak	Gyr	Mer
34	NAGE 8	Goshawk	Topinka & May	Conservation Genetics	2004	5 (6): 861-864	X	YES	X	X	X	X
35	NAGE 9	Goshawk	Topinka & May	Conservation Genetics	2004	5 (6): 861-864	X	YES	X	X	NO	X
36	NAGE 10	Goshawk	Topinka & May	Conservation Genetics	2004	5 (6): 861-864	X	YES	X	X	POOR	X
37	NAGE 11	Goshawk	Topinka & May	Conservation Genetics	2004	5 (6): 861-864	X	YES	X	X	X	X
38	AGE7	Goshawk	Sonsthagen et al.	The Condor	2004	106: 826-836	X	YES	X	X	X	X
39	NVH 107	Peregrine	Nesje et al.	Molecular Ecology	2000	9: 53-60	YES	X	NO	X	YES	YES
40	NVH 13	Peregrine	Nesje et al.	Molecular Ecology	2000	9: 53-60	X	X	YES	YES	YES	YES
41	NVH 195	Peregrine	Nesje et al.	Molecular Ecology	2000	9: 53-60	X	X	NO	NO	NO	NO
42	NVH 203	Peregrine	Nesje et al.	Molecular Ecology	2000	9: 53-60	X	X	NO	NO	NO	NO
43	NVH 31	Peregrine	Nesje et al.	Molecular Ecology	2000	9: 53-60	X	X	YES	YES	YES	YES
44	NVH 46-1	Peregrine	Nesje et al.	Molecular Ecology	2000	9: 53-60	X	X	YES	YES	YES	YES
45	NVH 5	Peregrine	Nesje et al.	Molecular Ecology	2000	9: 53-60	X	X	YES	YES	YES	YES
46	NVH 54	Peregrine	Nesje et al.	Molecular Ecology	2000	9: 53-60	X	X	YES	YES	YES	NO
47	NVH 79-1	Peregrine	Nesje et al.	Molecular Ecology	2000	9: 53-60	X	X	YES	YES	YES	YES
48	NVH 79-4	Peregrine	Nesje et al.	Molecular Ecology	2000	9: 53-60	X	X	YES	YES	YES	YES
49	NVH 82	Peregrine	Nesje et al.	Molecular Ecology	2000	9: 53-60	X	X	YES	YES	YES	YES
50	NVH 86	Peregrine	Nesje et al.	Molecular Ecology	2000	9: 53-60	X	X	YES	YES	YES	YES
51	NVH 89	Peregrine	Nesje et al.	Molecular Ecology	2000	9: 53-60	X	X	YES	YES	YES	YES
52	NVH 92-1	Peregrine	Nesje et al.	Molecular Ecology	2000	9: 53-60	X	X	YES	YES	YES	YES
53	NVH 34	Gyr	Nesje & Roed	Molecular Ecology	2000	9: 1438-1440	X	X	YES	YES	YES	YES
54	NVH fr142	Gyr	Nesje & Roed	Molecular Ecology	2000	9: 1438-1440	YES	X	NO	X	X	X
55	NVH fr144-2	Gyr	Nesje & Roed	Molecular Ecology	2000	9: 1438-1440	YES	X	NO	X	NO	X
56	NVH fr164-1	Gyr	Nesje & Roed	Molecular Ecology	2000	9: 1438-1440	X	X	NO	X	NO	X
57	NVH fr190-2	Gyr	Nesje & Roed	Molecular Ecology	2000	9: 1438-1440	YES	X	POOR	NO	YES	NO
58	NVH fr195-2	Gyr	Nesje & Roed	Molecular Ecology	2000	9: 1438-1440	YES	X	NO	X	POOR	X
59	NVH fr206	Gyr	Nesje & Roed	Molecular Ecology	2000	9: 1438-1440	POOR	X	NO	X	POOR	X
60	uAch 3	Golden Eagle	Parkin & Peck	Unpublished DEFRA Report			YES	YES	X	X	X	X
61	uAch 5	Golden Eagle	Parkin & Peck	Unpublished DEFRA Report			YES	YES	X	X	X	X
62	uAch 7	Golden Eagle	Parkin & Peck	Unpublished DEFRA Report			YES	NO	X	X	X	YES
63	uAge 1	Goshawk	Parkin & Peck	Unpublished DEFRA Report			YES	YES	X	X	X	X
64	uAge 1a	Goshawk	Parkin & Peck	Unpublished DEFRA Report			YES	YES	X	X	X	X
65	uAge 4	Goshawk	Parkin & Peck	Unpublished DEFRA Report			YES	NO	X	X	X	X
66	uAge5	Goshawk	Parkin & Peck	Unpublished DEFRA Report			X	X	X	X	X	POOR
67	uMmi 09	Red Kite	Parkin & Peck	Unpublished DEFRA Report			YES	X	NO	X	NO	X
68	uMmi 10	Red Kite	Parkin & Peck	Unpublished DEFRA Report			POOR	X	X	X	X	X
69	uMmi 11	Red Kite	Parkin & Peck	Unpublished DEFRA Report			NO	X	X	X	X	X
70	uMmi 13	Red Kite	Parkin & Peck	Unpublished DEFRA Report			YES	X	X	X	X	X

	Locus	Species	Authors	Journal	Year	Page_Vol	Amplification					
							GE	Gos	Per	Sak	Gyr	Mer
71	uFPE 1	Peregine	Jon Wetton	Unpublished DEFRA Report			X	X	YES	YES	YES	YES
72	uFPE 2	Peregine	Jon Wetton	Unpublished DEFRA Report			X	X	YES	YES	YES	YES
73	ND58	Peregrine	Calonge E et al	Unpublished AF448411 - WDNAS designed primers			NO	NO	YES	YES	YES	YES
74	ND360	Peregrine	Calonge E et al	Unpublished AF448411 - WDNAS designed primers			X	X	YES	X	X	X
75	ND20	Peregrine	Calonge E et al	Unpublished AF448411 - WDNAS designed primers			X	X	YES	X	X	X
76	ND184	Peregrine	Calonge E et al	Unpublished AF448411 - WDNAS designed primers			X	X	YES	X	X	X
77	ND354	Peregrine	Calonge E et al	Unpublished AF448411 - WDNAS designed primers			X	X	NO	X	X	X
78	ND347	Peregrine	Calonge E et al	Unpublished AF448411 - WDNAS designed primers			X	X	YES	X	X	X
79	Esu6	Reed Bunting	Hanotte et al	Molecular Ecology	1994	3: 529 - 530	X	X	X	X	POOR	POOR
80	Ase18	Seychelles warbler	Richardson et al	Molecular Ecology	2000	9: 2226 - 2231	X	X	X	X	YES	YES
81	Ase9	Seychelles warbler	Richardson et al	Molecular Ecology	2000	9: 2226 - 2231	X	X	X	X	YES	X
82	MCY	Malurus cyaneus	Double et al 1997	Molecular Ecology	1997	6: 691 - 693	X	X	X	X	NO	X
83	Ppi2	Pica pica	Martinez et al	Evolution	1999	53: 269 - 278	X	X	X	X	NO	X
84	Pdo5	Passer domesticus	Griffith et al	Biol. J. Linn. Soc.	1999	68: 303 - 316	X	X	X	X	NO	X
85	Le160		D.A. Dawson	Unpublished			YES	YES	YES	YES	YES	YES
86	Msfp01	Peregrine	Fickel et al	Unpublished AF218771			X	X	YES	YES	YES	YES
87	Msfp02	Peregrine	Fickel et al	Unpublished AF218772			X	X	POOR	X	X	X
88	Msfp03	Peregrine	Fickel et al	Unpublished AF218773			X	X	POOR	X	X	X
89	IEAAAGO4	Imperial eagle/Steppe eagle	Busch et al.	Molecular Ecology Notes	2005	5: 39-41	YES	X	X	X	X	X
90	IEAAAG15	Imperial eagle/Steppe eagle	Busch et al.	Molecular Ecology Notes	2005	5: 39-41	YES	X	X	X	X	X
91	ge27	Golden Eagle	Dillane	Unpublished			YES	X	X	X	X	X
92	ge36	Golden Eagle	Dillane	Unpublished			YES	X	X	X	X	X
93	ge48	Golden Eagle	Dillane	Unpublished			YES	X	X	X	X	X
94	ge55	Golden Eagle	Dillane	Unpublished			YES	X	X	X	X	X
95	NICK 3	Merlin	WDNAS	Unpublished			X	X	X	X	X	NO
96	NICK 5	Merlin	WDNAS	Unpublished			X	X	X	X	X	NO
97	NICK 6	Merlin	WDNAS	Unpublished			X	X	X	X	X	NO
98	NICK 8	Merlin	WDNAS	Unpublished			X	X	X	X	X	NO
99	ROSS 10	Gyr/Sak	WDNAS	Unpublished			X	X	X	X	POOR	X
100	ROSS 9	Gyr/Sak	WDNAS	Unpublished			X	X	POOR	X	X	X

Appendix III

PROTOCOL FOR SWGDAM VALIDATION OF MICROSATELLITES

DEVELOPMENTAL PROTOCOL 1

1.1 Validation of Positive and Negative Controls – A positive and negative control is chosen for PCR. The positive control is an individual (or number of individuals) that is shown to consistently amplify. The negative control is water.

Protocol.
1) DNA is extracted from 5 individuals (captive bred) following forensic extraction protocols. The sample codes are recorded.
2) DNA is visualised on a 1% agarose gel and quantified using a flourometer. Quantification data is recorded.
3) DNA extract is used to optimise PCR conditions (highest temperature, lowest $MgCl_{2,}$ lowest number of cycles) by performing PCR reactions under a range of thermocycling conditions.
4) DNA extract is chosen as a positive control based on its performance during optimisation. The negative control water is UV treated prior to each PCR
5) Following PCR optimisation the positive controls is genotyped and the allele sizes recorded.
6) Positive and negative controls are labelled and kept at -20^0C separate from other samples used in the validation experiments.

DEVELOPMENTAL PROTOCOL 2

2.1 Range of PCR conditions – It is necessary to demonstrate the effect that varying the reaction conditions has on the data produced. These include thermocycling parameters and magnesium chloride concentrations. 1 – 10ng/µl DNA is used.

Protocol.
$MgCl_2$ Concentrations.
1) To test $MgCl_2$ concentrations, separate master mixes containing 1.5mM, 2mM, 2.5mM and 3mM are made.
2) Three samples (together with a positive and negative control) undergo PCR under each of these different $MgCl_2$ concentrations, while other variables (temp, number of cycles, DNA concentration) are kept constant.
4) Amplification product is run on a 2% agarose gel and visualised with a 100bp ladder. A photo is taken and the success of the PCR is recorded.

Thermocycling parameters.
1) To test the effect of the thermocycling parameters the optimised master mix determined above is used.
2) Three samples (together with a positive and negative control) undergo PCR under varying thermocycling parameters (temp, number of cycles) while other variables are kept constant.

3) The thermocycling conditions tested are:
   i) Cycle number = optimised $\pm$ 1 and $\pm$ 2 cycles.
   ii) Annealing temperature = optimised $\pm$ 2 and $\pm$ 4 degrees C.

4) Samples are run on a 2% agarose gel and visualised with a 100bp ladder.

DEVELOPMENTAL PROTOCOL 3

3.1 Inheritance – The mode of inheritance of DNA markers demonstrated through family studies.

Protocol.
1) Family groups are identified for each species under study.
2) DNA is recovered from family samples following forensic extraction procedures.

DNA concentration is quantified as above.
3) The family samples (together with a positive and negative control) undergo PCR following the optimised conditions (see above).
4) Samples are run on a 2% agarose gel and visualised with a 100bp ladder. A photo is taken and the success of the PCR is recorded.
5) The family samples (together with a positive and negative control) are genotyped on the CEQ 8000 and the mode of inheritance and the level of polymorphism is examined.

DEVELOPMENTAL PROTOCOL 4

4.1 Sensitivity Studies – The range of DNA quantities able to produce reliable typing results is determined.

Protocol.
1) DNA is diluted to produce samples at the following concentrations: 10, 5, 2.5, 1.25, 0.62, 0.31, 0.16ng/ul.
2) The samples (together with a positive and negative control) undergo PCR following the optimised conditions (see above).
3) Samples are run on a 2% agarose gel and visualised with a 100bp ladder. A photo is taken and the success of the PCR is recorded.
4) The samples are run on the CEQ 8000 and the affect of DNA concentration on locus amplification is assessed.

DEVELOPMENTAL PROTOCOL 5

5.1 Species Specificity – The ability of loci to amplify DNA in non-target species is assessed. Species typically mistaken, encountered or likely to contaminate are used.

Protocol.
1) PCR amplification success of a target sample is compared with samples of:
   i) A near related species i.e same genus.
   ii) A more distant species i.e same order.
   iii) Human.
2) The samples (together with a positive and negative control) undergo PCR following the optimised conditions (see above).
3) Samples are run on a 2% agarose gel and visualised with a 100bp ladder. A photo is taken and the success of the PCR is recorded.

DEVELOPMENTAL PROTOCOL 6

6.1 Stability studies – The ability to obtain results from DNA recovered from biological samples deposited on various substrates and subjected to various environmental insults will be assessed.

Protocol.

Environmental stability.

1) Blood is subjected to a time series study where a single sample is subjected to environmental insults over a 4 - 6 week period. During this time a daily log of weather conditions (Temp, level of rain fall) is kept.

2) 50µl of blood is pipetted on to a swab head and allowed to dry. Twenty swabs provides five replicates for each time period of the experiment. Blood untreated negative control swabs are also required.

3) After drying, five swab heads and a negative is extracted from providing the week

0 data. The extraction from swab head protocol is used (thereby doubling up on reagents).

4) The other swab heads are placed out side. Once every two weeks 5 samples and a negative control are extracted from.

5) The DNA is quantified and the data recorded in a spreadsheet.

6) The samples (together with positive control DNA and the negative control extract) undergo PCR following the optimised conditions. The largest available marker for the species is used in the PCR reaction as an indication of DNA degredation.

7) Samples are run on a 2% agarose gel and visualised with a 100bp ladder. A photo is taken and the success of the PCR is recorded.

Substrate and chemical effects.

Protocol:

1) 50µl of fresh blood is placed on the 10 replicates of each item and left to dry. The items being denim, metal, suede, wood, and cotton swabs treated with 50µl bleach, soap, gasoline and NaOH. Negative control items (items untreated with blood) are also required.

2) When the blood is dry, DNA is recovered from five replicates by following standard swab extraction protocols.

3) The other five items and negative are stored in ambient conditions for 5 days and then extracted.

4) All extractions from items mixed with chemicals are done in the fume cupboard and left for 1 hour before the extraction commences.

5) The DNA is quantified and the data recorded in a spreadsheet.

6) The samples (together with positive control DNA and the negative control extract) undergo PCR following the optimised conditions. The largest available marker for the species is used in the PCR reaction as an indication of DNA degredation.

7) Samples are run on a 2% agarose gel and visualised with a 100bp ladder. A photo is taken and the success of the PCR is recorded.

DEVELOPMENTAL PROTOCOL 7

8.1 Precision, Accuracy, and Reproducibility – The amplification of each locus is evaluated in the laboratory to ensure the consistency of the result. The extent to which a given set of measurements of the same sample agree with their mean and the extent to which these measurements match the actual values being measured is determined.

Protocol.

1) A single individual undergoes PCR three independent times under the standard conditions.
2) Samples are run on a 2% agarose gel and visualised with a 100bp ladder. A photo is taken and the success of the PCR is recorded.
3) The samples are run on the CEQ 8000 and genotypic data recorded.
4) The data is used to infer the precision, accuracy and reproducibility of results generated from each locus.

Appendix IV **Multiplex design for each species (a-f).**

Shading represents different dye colours:

Light w/black text = green, medium w/white text = blue, dark w/white text =black

a) GOSHAWKS

Allele Size	Microsatellite locus							
	Nage5	Nage9	Nage7	Nage4	Nage10	Nage1	Age1a	Nage2
125 - 149					144			
150 - 174	156-171				172		162-170	
175 - 199		184-190						187
200 - 224			220			203		
225 - 249			242					238
250 - 274								
275 - 299								
300 - 324				304		318		
325 - 349								
350 - 374				352				

b) GOLDEN EAGLES

Set 1

Allele Sizes	Microsatellite locus				
	aa02	aa39	aa15	aa43	aa26
75 - 99					
100 - 124				107-117	
125 - 149	135-141				141
150 - 174					155
175 - 199		188			
200 - 224		202	200-206		
225 - 249					
250 - 274					
275 - 299					
300 - 324					

Set 2

Allele Sizes	Microsatellite locus					
	aa36	aa04	ieeea04	aa27	ieaaa15	nvh142
75 - 99	85			83-97		
100 - 124	105	123			113	
125 - 149					129	
150 - 174		155				
175 - 199						177-185
200 - 224						
225 - 249			228-246			
250 - 274						
275 - 299						
300 - 324						
325 - 349						

c) PEREGRINES

Set 1

Allele Sizes	Microsatellite locus						
	13	46	ufpe01	89	31	nvh86	Nage 5
51 - 75							
76 - 100							
101 - 125	101 - 107	122		118			
126 - 150		126		134		139 - 147	
151 - 175			153		153 - 159		164 - 170
176 - 200			181				
201 - 225							

Set 2

Allele Sizes	Microsatellite locus				
	54	NVH34	ufpe02	92	79-4
51 - 75					
76 - 100					
101 - 125	103 - 125			115 - 123	
126 - 150					148
151 - 175		154-156			166
176 - 200					
201 - 225					
226 - 250			223		
251 - 275					
276 - 300					
301 - 325			307		

d) GYRS

Set 1

Allele Sizes	Microsatellite locus					
	13	46	ufpe01	89	Nage 7	Nage 5
51 - 75						
76 - 100	99					
101 - 125	101	116		122 - 124		
126 - 150		130				
151 - 175			158			155 - 164
176 - 200						
201 - 225			210		204 - 210	
226 - 250						
251 - 275						

Set 2

Allele Sizes	Microsatellite locus				
	54	NVH34	82	92	79-4
51 - 75					
76 - 100	88				
101 - 125				102 - 110	
126 - 150	136	148	138 - 150		150
151 - 175		152			152
176 - 200					
201 - 225					
226 - 250					

e) SAKERS

Set 1

Allele Sizes	Microsatellite locus					
	13	ufpe01	89	31	Nage 7	Nage 5
51 - 75						
76 - 100						
101 - 125	101 - 105		120 - 122			
126 - 150		137		143		
151 - 175				151		152 - 170
176 - 200		193				
201 - 225					204-210	
226 - 250						
251 - 275						

Set 2

Allele Sizes	Microsatellite locus				
	54	82	ufpe02	92	79-4
51 - 75					
76 - 100					
101 - 125	102			107 - 117	
126 - 150	154	133			144
151 - 175		155			156
176 - 200			178		
201 - 225					
226 - 250			242		
251 - 275					

f) MERLINS

Set 1

Allele Sizes	Microsatellite locus					
	ufpe01	82	31	nd58	92	Nage 5
51 - 75						
76 - 100						
101 - 125		125			107-117	
126 - 150		127				
151 - 175	154 - 170		153 - 169			158 - 170
176 - 200				192-198		
201 - 225						
226 - 250						

Set 2

Allele Sizes	Microsatellite locus			
	46	89	ufpe02	79-4
51 - 75				
76 - 100				
101 - 125	123	123		
126 - 150	145	141		136
151 - 175				162
176 - 200				
201 - 225			212	
226 - 250				
251 - 275				
276 - 300				
301 - 325			316	

Appendix V Details of PCR reaction conditions for each PCR multiplex optimized for the six species

1) Volumes of forward and reverse primer (20mM) required for each multiplex primer mix:

Multiplex	Goshawk		G. eagle		Peregrine		Gyr		Saker		Merlin	
	Locus	Vol.	Locus	Vol.	Locus	Vol.	Locus	Vol.	Locus	Vol.	Locus	Vol.
1	AGE1A	0.35	AA02	0.75	UFPE01	6.00	UFPE01	6.00	UFPE01	6.00	UFPE01	6.00
	NAGE1	2.25	AA15	0.75	NVH13	3.40	NVH13	3.00	NVH13	3.00	NVH31	1.50
	NAGE2	0.70	AA26	0.75	NVH31	2.20	NVH46	4.20	NVH31	1.60	NVH82	3.00
	NAGE4	0.90	AA39	1.50	NVH46	4.20	NVH89	2.00	NVH89	1.60	NVH92	2.00
	NAGE5	5.20	AA43	1.50	NVH86	2.00	NAGE5	2.00	NAGE5	2.40	NAGE5	1.50
	NAGE7	0.35			NVH89	2.80	NAGE7	1.60	NAGE7	1.60	ND58	0.50
	NAGE9	0.90			NAGE5	2.40						
	NAGE10	0.90										
2			AA04	2.00	UFPE02	1.75	NVH34	1.00	UFPE02	1.50	UFPE02	3.00
			AA27	2.80	NVH34	1.50	NVH54	1.00	NVH54	2.00	NVH46	4.00
			AA36	4.00	NVH54	1.00	NVH79-4	1.50	NVH79-4	2.00	NVH79-4	1.00
			NVH142	1.60	NVH79-4	1.25	NVH82	1.50	NVH82	1.00	NVH89	1.50
			IEEEA04	1.60	NVH92	1.00	NVH92	1.50	NVH92	1.00		
			IEEEA15	3.00								

2) PCR reagents:

PCR mix (RED TOP)	6ul
Primer mix	1ul
Q Solution	1ul
H20	3ul
DNA Template	1ul
Total	12ul

3) PCR thermocycling conditions (AT=annealing temp.):

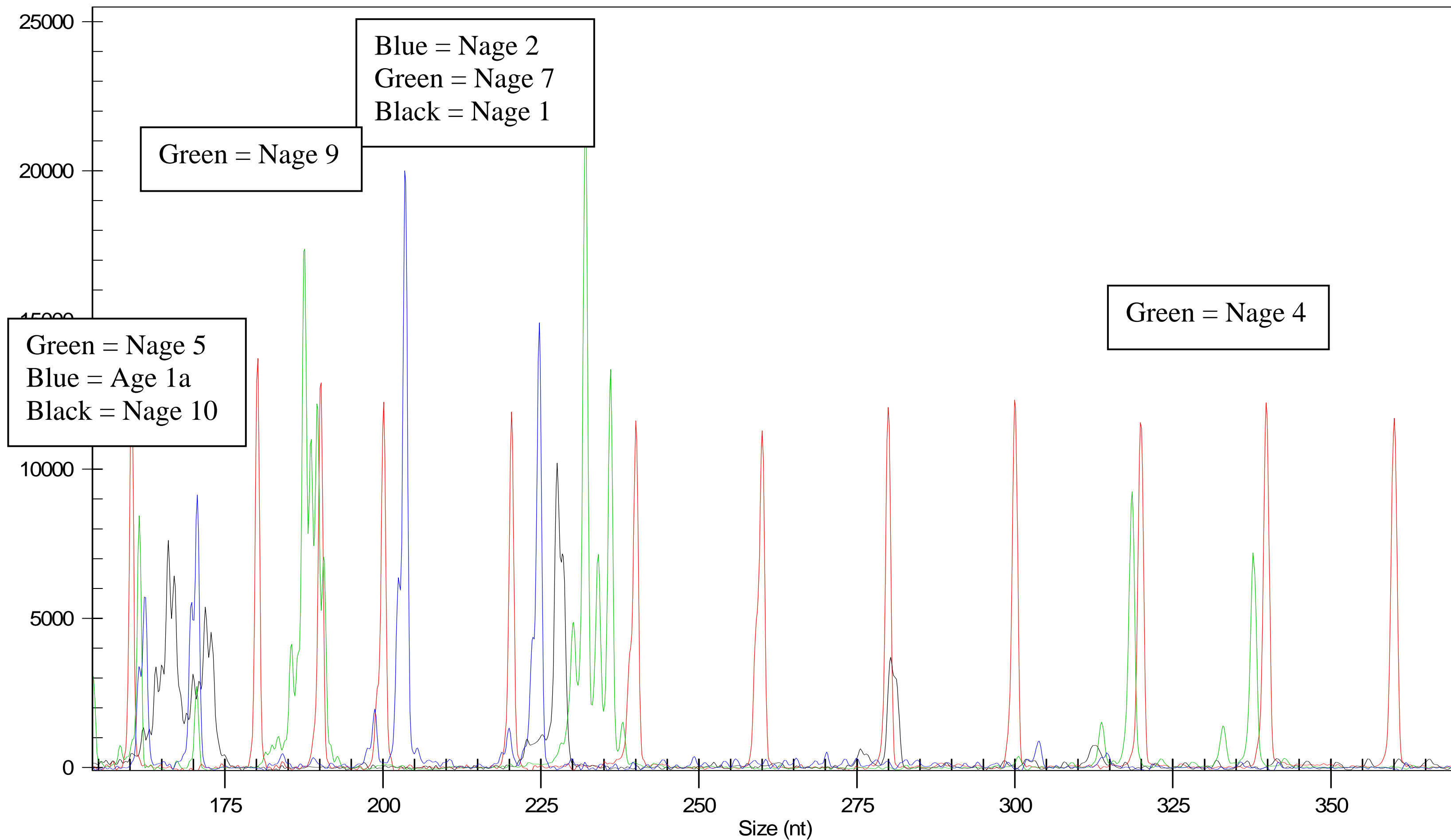
95°C	15 min	
95°C	30 sec	40 cycles
AT°C	30 sec	40 cycles
72°C	30 sec	40 cycles
72°C	5 min	
4°C	Hold	

4) Annealing temperatures:

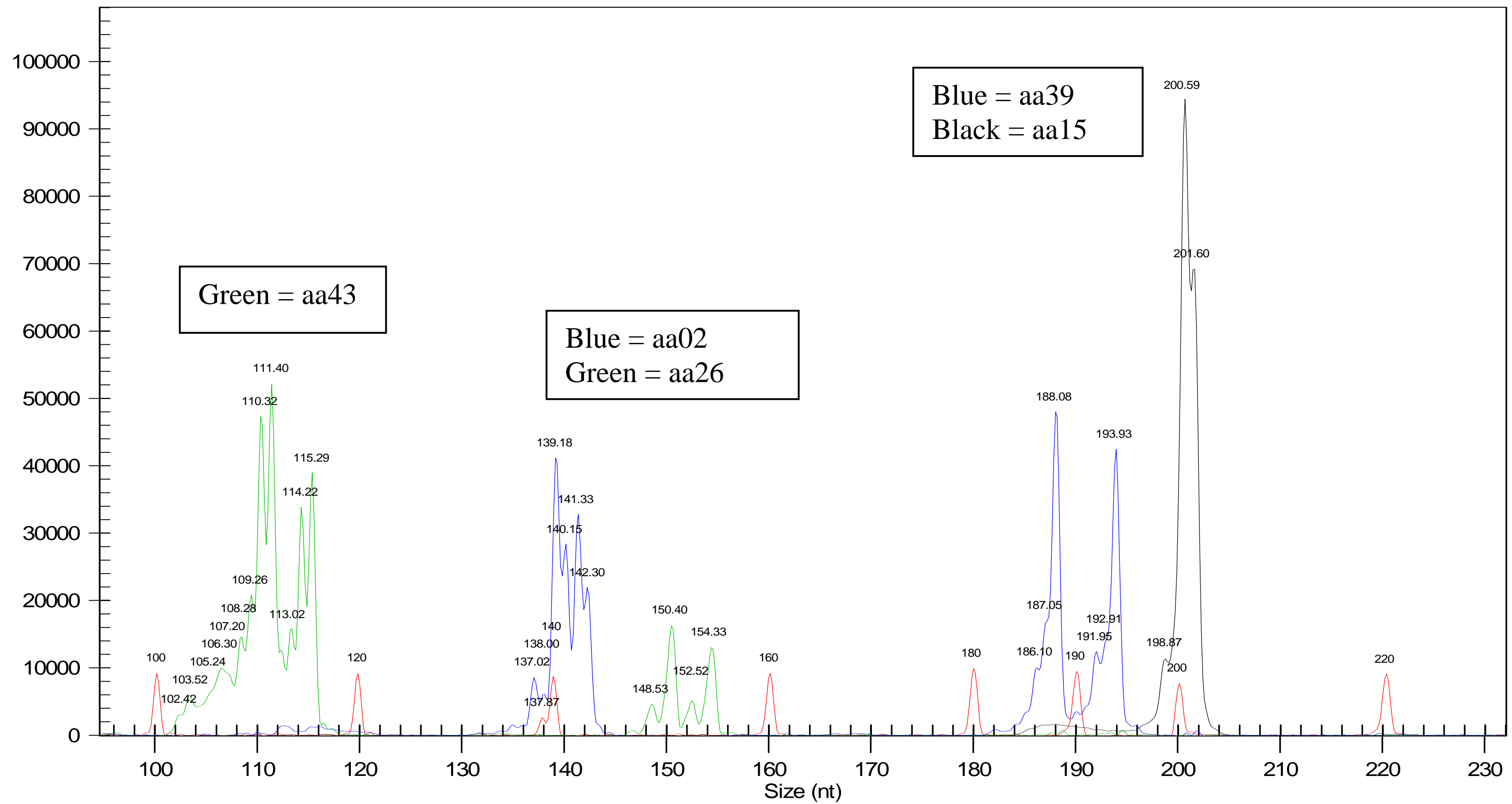
Multiplex	Goshawk	G. eagle	Peregrine	Gyr	Saker	Merlin
1	50°C	56°C	50°C	50°C	50°C	54°C
2		50°C	50°C	52°C	50°C	52°C

Appendix VI Examples of each PCR multiplex.

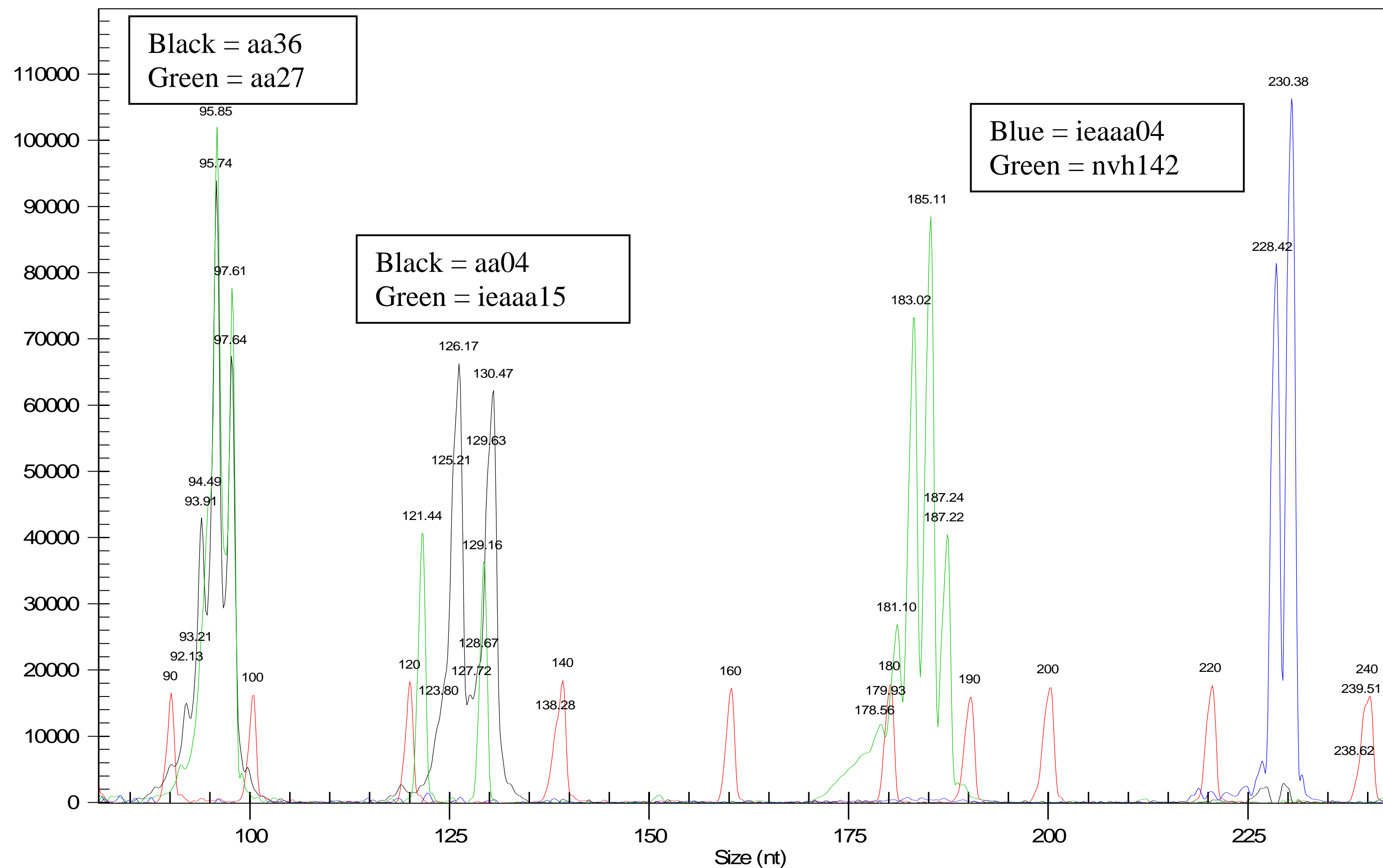
Goshawk multiplex screenshot (Beckman CEQ8000).
Loci identified by name and colour – see Appendix IV for corresponding allele size ranges



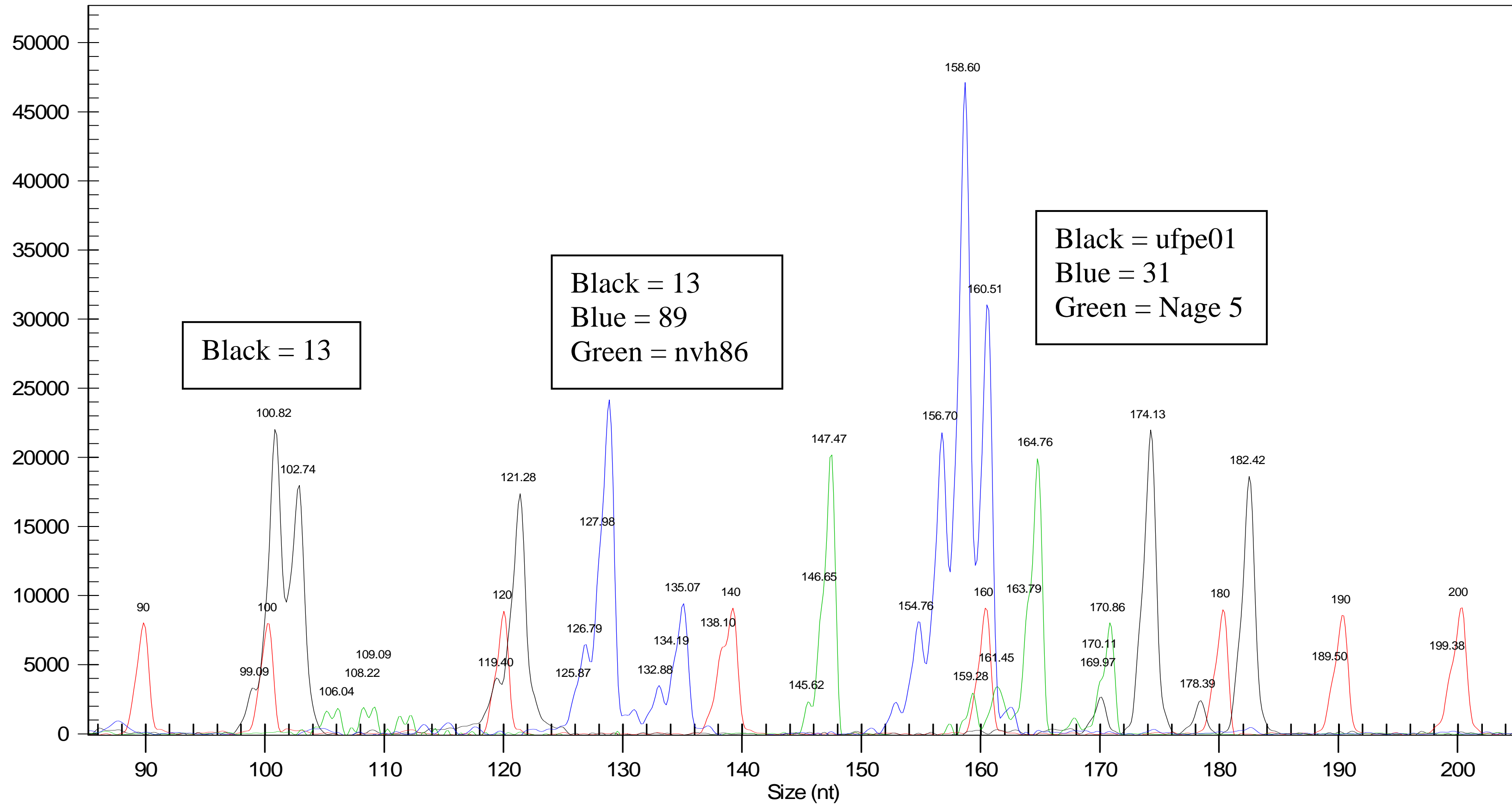
Golden eagle multiplex 1 screenshot (Beckman CEQ8000).
Loci identified by name and colour – see Appendix II for corresponding allele size ranges



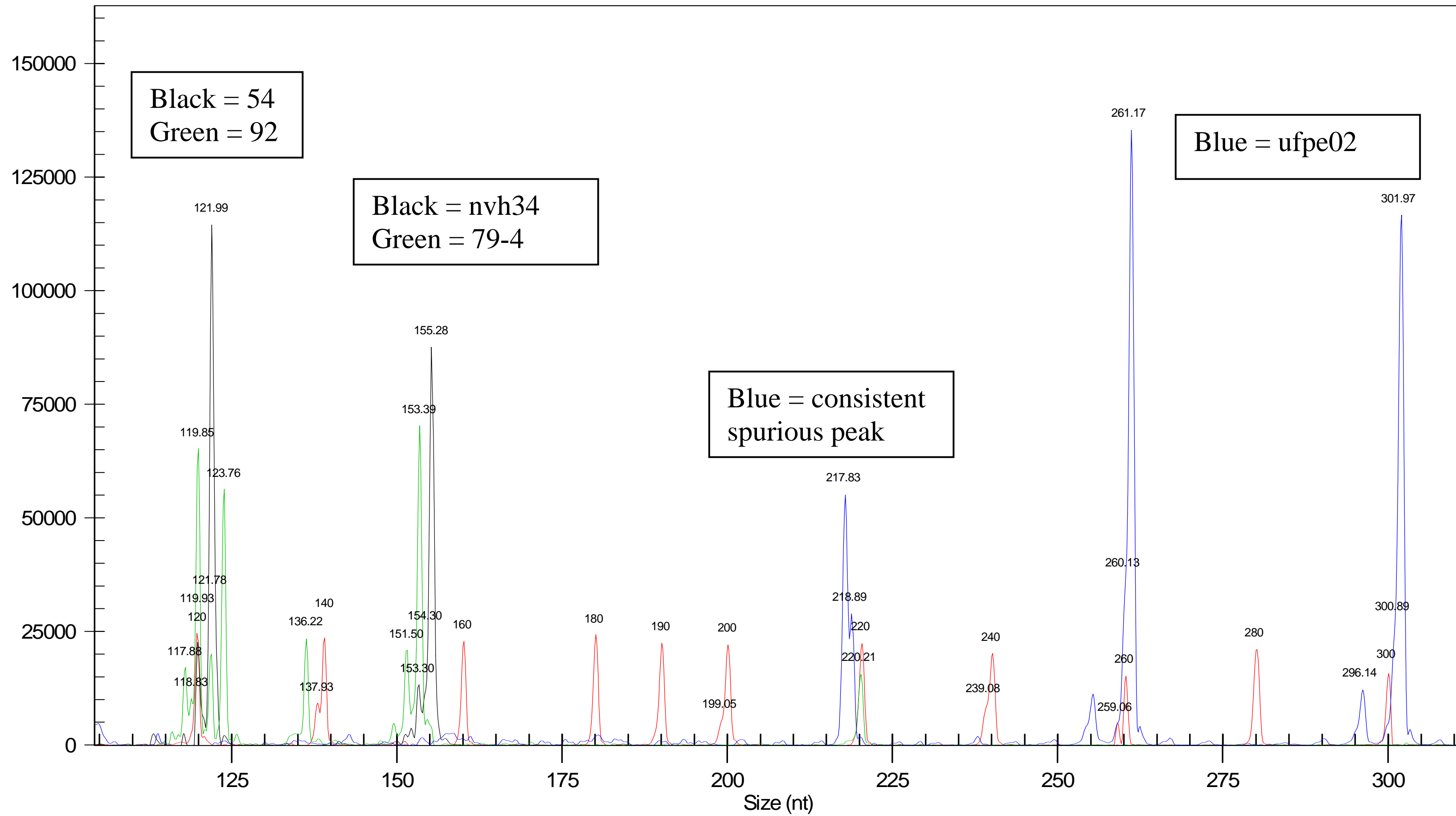
Golden eagle multiplex 2 screenshot (Beckman CEQ8000).
Loci identified by name and colour – see Appendix II for corresponding allele size ranges



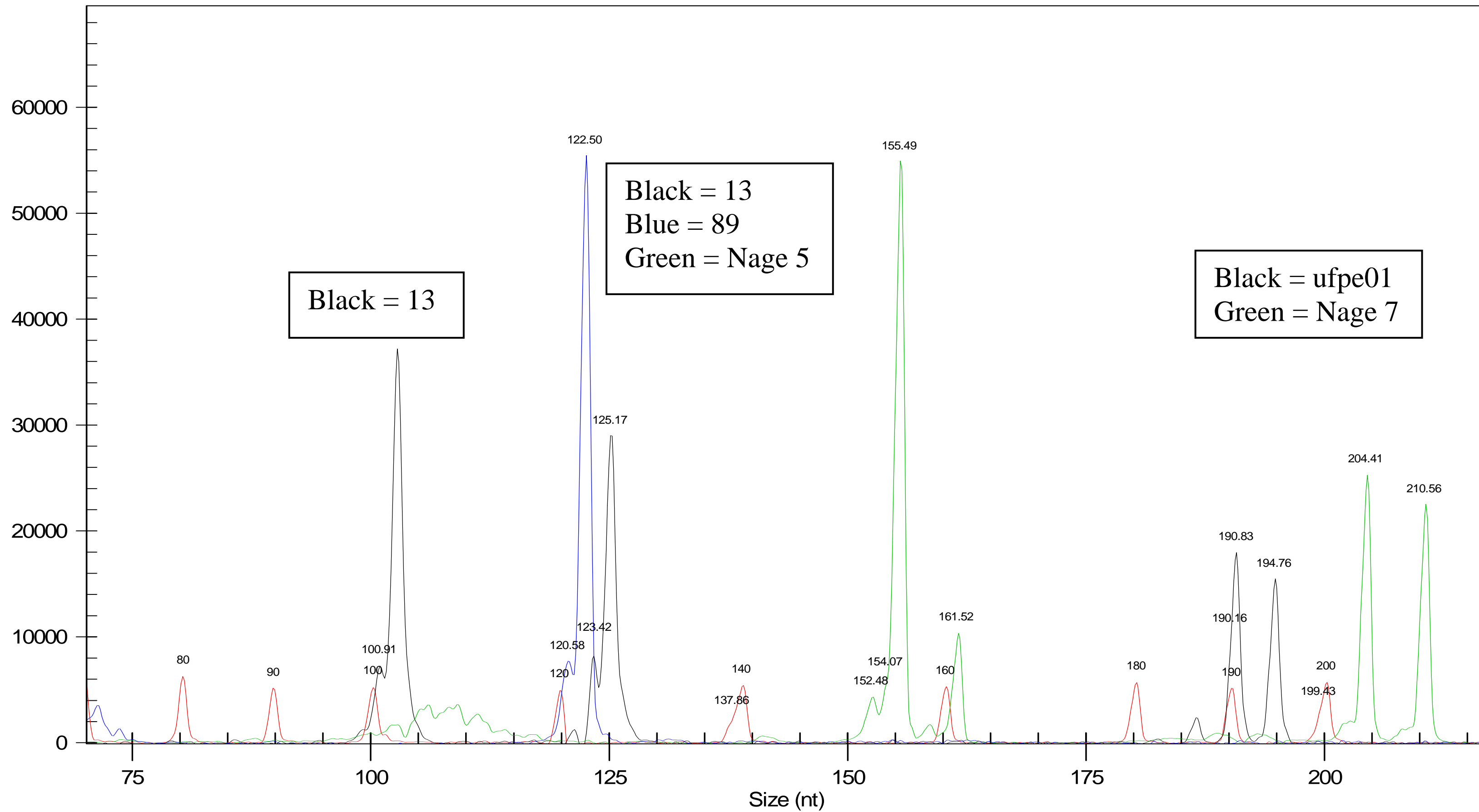
Peregrine falcon multiplex 1 screenshot (Beckman CEQ8000).
Loci identified by name and colour – see Appendix II for corresponding allele size ranges



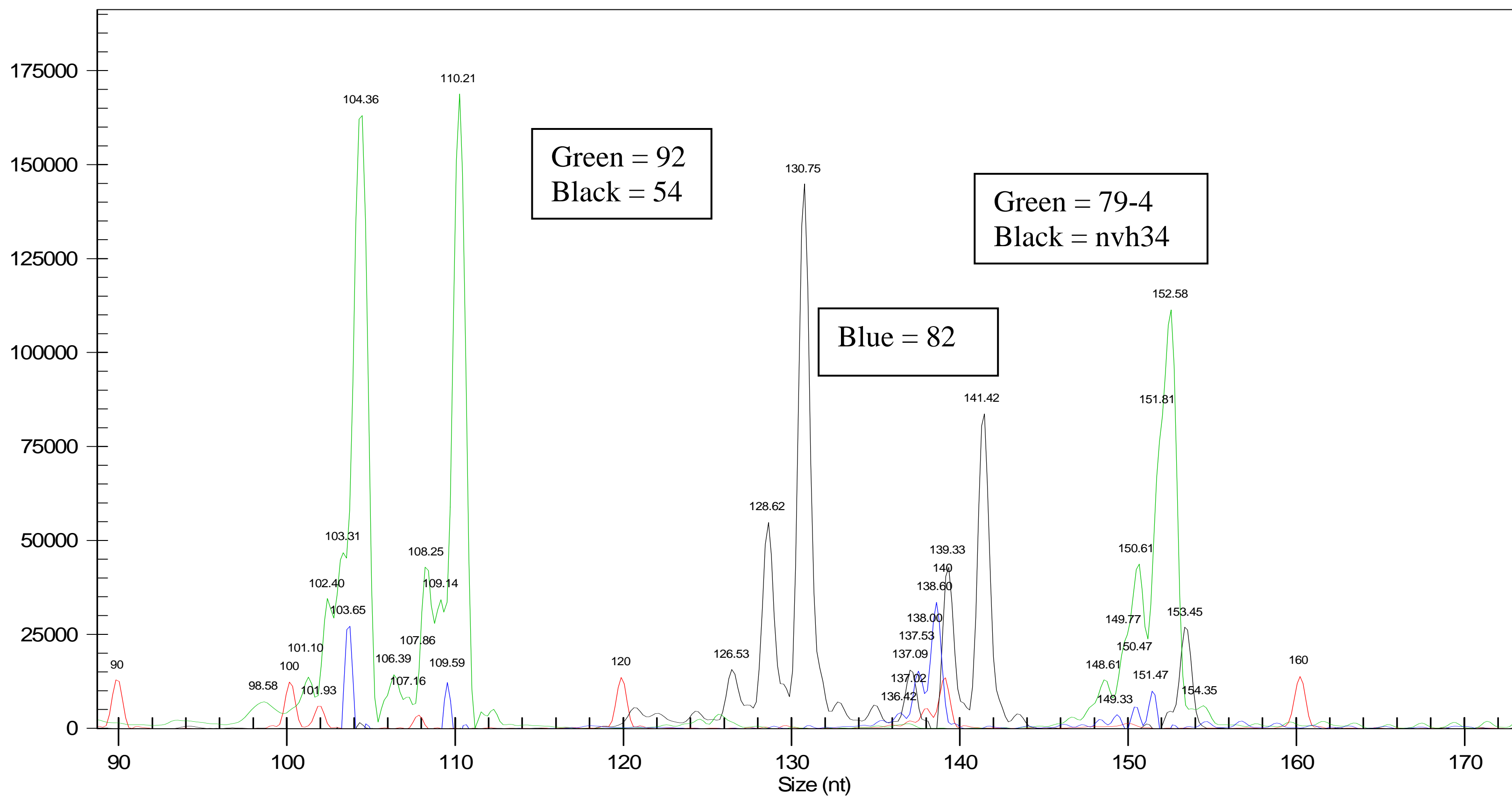
Peregrine falcon multiplex 2 screenshot (Beckman CEQ8000).
Loci identified by name and colour – see Appendix II for corresponding allele size ranges



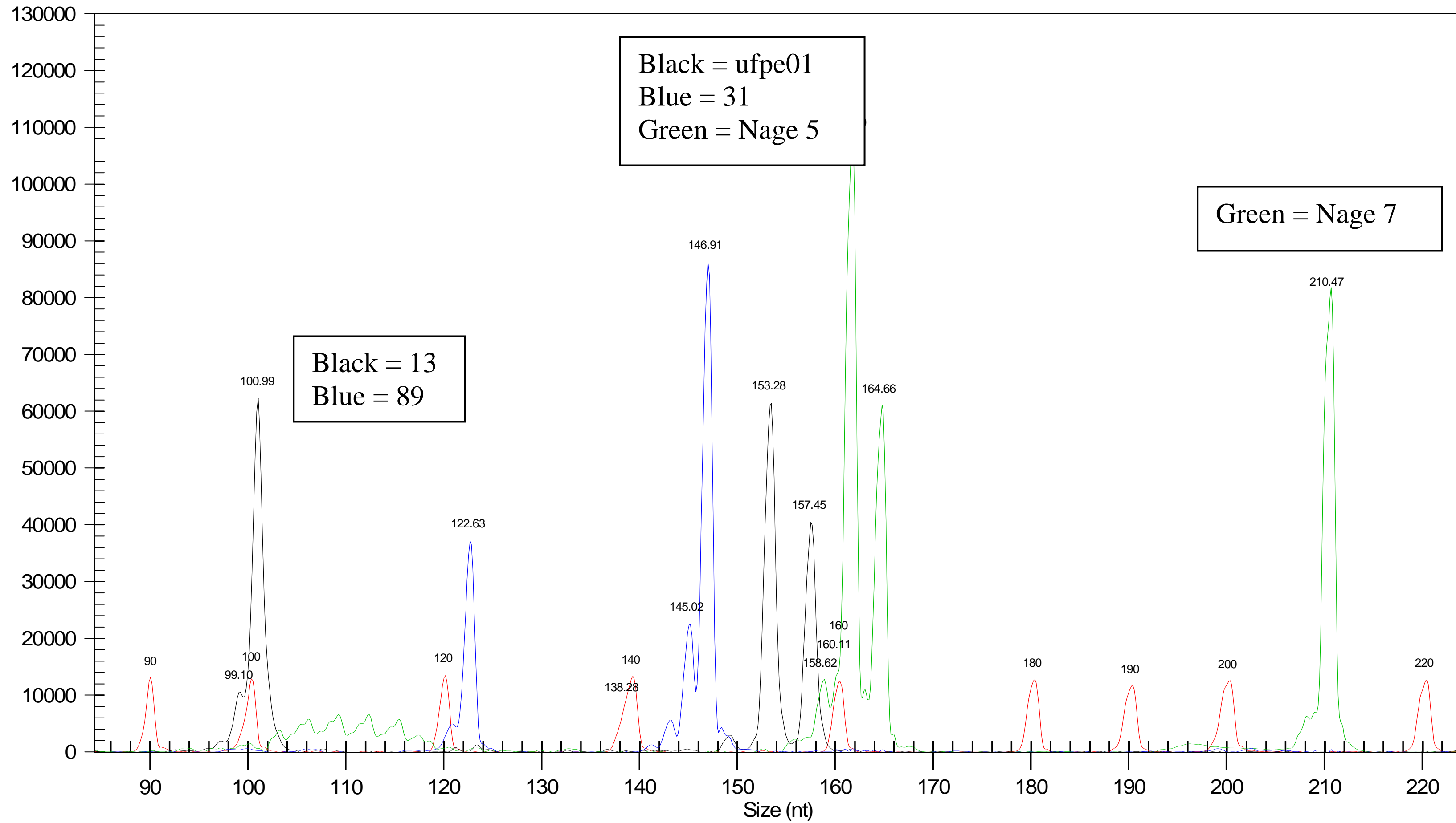
Gyr falcon multiplex 1 screenshot (Beckman CEQ8000).
Loci identified by name and colour – see Appendix II for corresponding allele size ranges



Gyr falcon multiplex 2 screenshot (Beckman CEQ8000).
Loci identified by name and colour – see Appendix II for corresponding allele size ranges

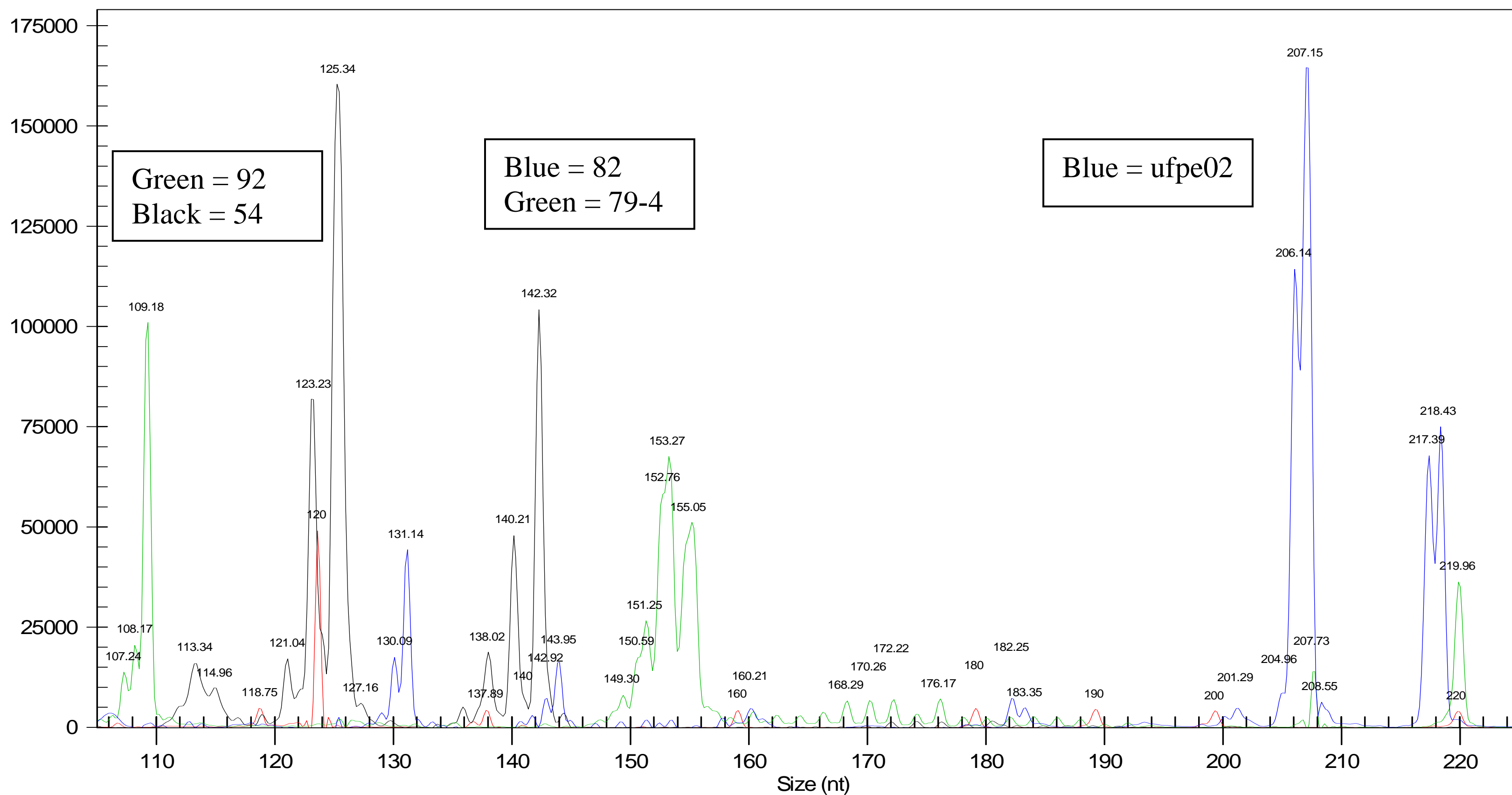


Saker falcon multiplex 1 screenshot (Beckman CEQ8000).
Loci identified by name and colour – see Appendix II for corresponding allele size ranges



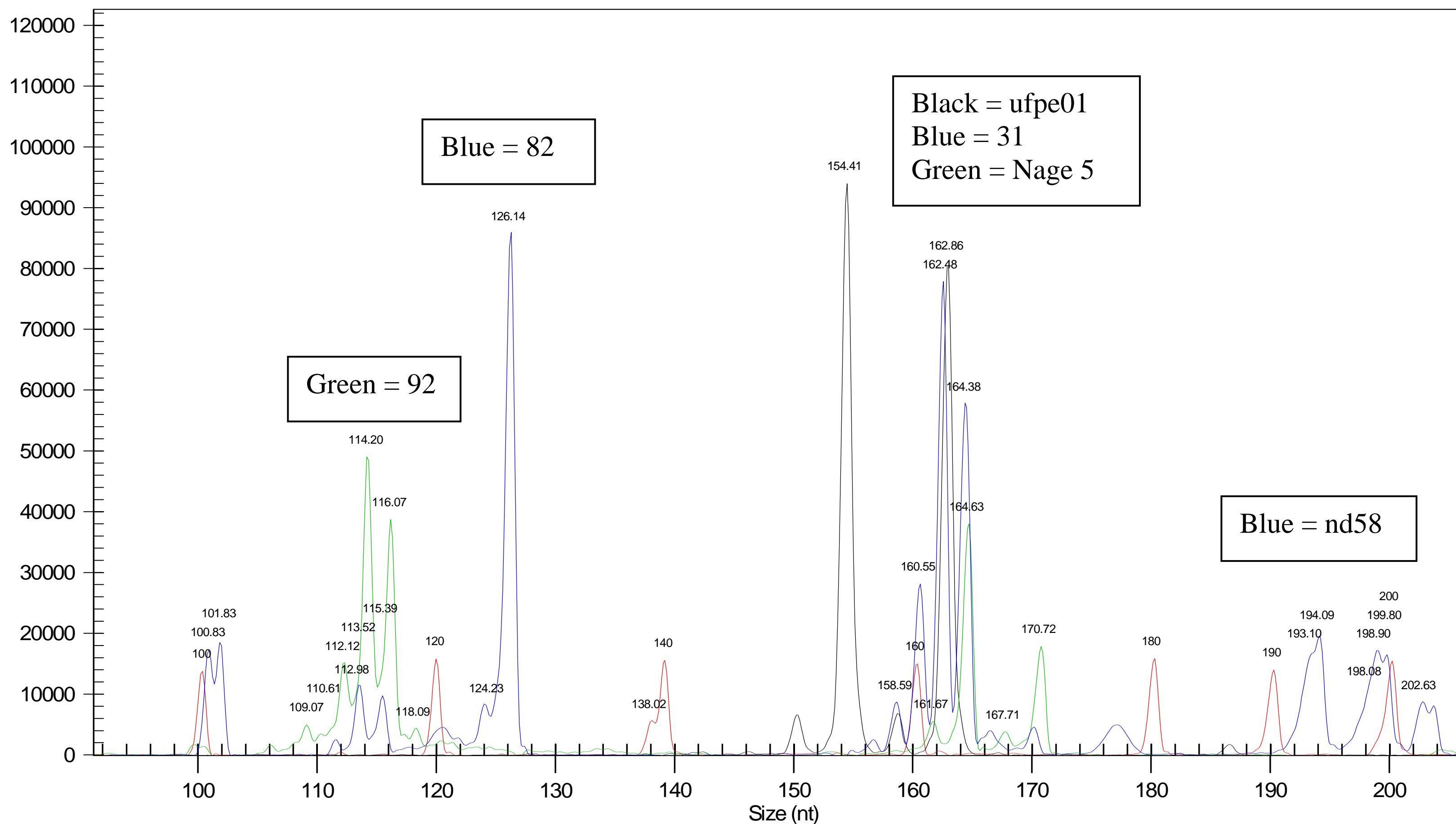
Saker falcon multiplex 2 screenshot (Beckman CEQ8000).

Loci identified by name and colour – see Appendix II for corresponding allele size ranges

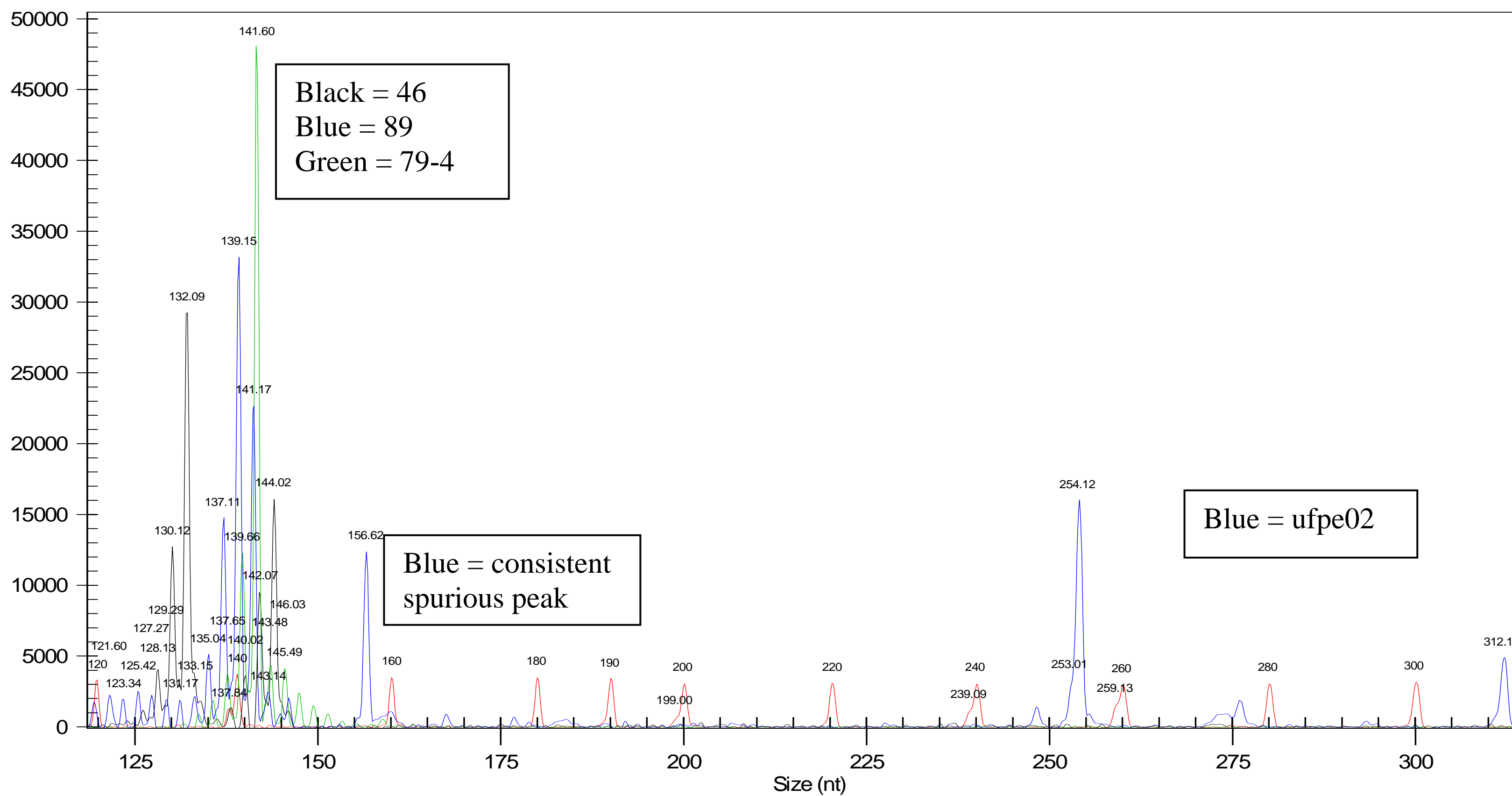


Merlin multiplex 1 screenshot (Beckman CEQ8000).

Loci identified by name and colour – see Appendix II for corresponding allele size ranges



Merlin multiplex 2 screenshot (Beckman CEQ8000).
Loci identified by name and colour – see Appendix II for corresponding allele size ranges



Appendix VII – Species identification test sequence data

DNA sequences and species-specific primer sites within mtDNA cytochrome oxidase subunit 1 for the six target species. Specific primer sites are highlighted in **bold**. Note that saker and gyrs falcons have the same primer site as no sequence differences were observed at this locus.

SAKER --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --t tag cct cct tat tcg aac aga act tgg cca acc agg gac tct cct agg aga [99]
GYR --- --- --- --- --- --- --- --- agc agg cat agt cgg cac tgc cct tag cct cct tat tcg aac aga act tgg cca acc agg gac tct cct agg aga [99]
PEREGRINE --- --- --- --- --- --- --- --- --c agg cat agt cgg tac tgc cct tag cct cct tat tcg aac aga act tgg cca acc agg aac tct cct agg aga [99]
MERLIN --- --- -ct act ctt cgg agc atg agc agg tat agc cgg cac tgc cct cag cct cct cat tcg agc aga act tgg aca acc agg aac tct cct agg aga [99]
GOLDENEAGLE --- --- --- --- --- --- --- --- --- --- --- --- --- --c cgc cct tag cct tct tat ccg cgc aga act tgg cca acc tgg cac cct ctt agg cga [99]
GOSHAWK cct ata ttt aat ctt tgg cgc ttg agc cgg cat agt agg cac tgc cct tag cct cct cat ccg cgc aga act cgg tca acc agg tac act cct agg cga [99]

SAKER tga cca aat cta caa tgt cat tgt cac cgc cca tgc ctt cgt aat aat ctt ttt cat agt cat acc cat tat gat cgg agg att tgg aaa ctg act agt [198]
GYR tga cca aat cta caa tgt cat tgt cac cgc cca tgc ctt cgt aat aat ctt ttt cat agt cat acc cat tat gat cgg agg att tgg aaa ctg act agt [198]
PEREGRINE tga cca aat cta caa tgt cat cgt cac tgc cca tgc ctt cgt aat aat ctt ttt cat agt tat acc cat tat gat cgg agg att tgg aaa ctg act agt [198]
MERLIN tga cca aat cta caa tgt cat cgt cac cgc cca tgc ttt cgt aat aat ctt ctt cat agt tat acc **cat cat gat tgg ggg att tgg** aaa ctg act agt [198]
GOLDENEAGLE tga cca aat cta caa tgt aat cgt cac cgc tca tgc ttt cgt aat aat ctt ctt cat agt cat acc aat cat aat tgg agg ctt tgg **aaa ctg act tgt** [198]
GOSHAWK cga cca aat cta caa cgt aat cgt cac cgc aca tgc ctt cgt aat aat ctt ctt cat agt tat acc aat cat aat tgg agg ctt cgg aaa ctg act cgt [198]

SAKER ccc cct tat aat tgg agc ccc aga cat agc att ccc ccg cat aaa caa cat gag ttt ctg act gct ccc ccc atc ctt tct act act cct agc atc ttc [297]
GYR ccc cct tat aat tgg agc ccc aga cat agc att ccc ccg cat aaa caa cat gag ttt ctg act gct ccc ccc atc ctt tct act act cct agc atc ttc [297]
PEREGRINE tcc cct tat aat **tgg agc tcc aga cat agc att tcc ccg** cat aaa caa cat gag ttt ctg act act ccc ccc atc ctt tct act gct cct agc atc ttc [297]
MERLIN ccc cct tat aat tgg agc ccc aga cat agc att ccc ccg cat aaa caa cat gag ctt ctg act act ccc tcc atc ctt cct act act cct agc atc ttc [297]
GOLDENEAGLE **ccc act cat** aat tgg cgc ccc cga cat agc ctt ccc acg cat aaa caa cat gag ctt ctg att act ccc ccc atc ctt cct cct cct act agc ctc ttc [297]
GOSHAWK tcc act cat aat tgg cgc ccc cga cat agc ctt ccc acg cat aaa caa cat **aag ctt ctg act act ccc tcc atc ttt** cct cct cct act agc ctc ctc [297]

SAKER cac agt aga agc ggg agt tgg aac agg atg aac cgt ata ccc ccc ctt agc agg caa cct agc **cca tgc cgg tgc ttc agt aga cct cgc cat** ttt ctc [396]
GYR cac agt aga agc ggg agt tgg aac agg atg aac cgt ata ccc ccc ctt agc agg caa cct agc **cca tgc cgg tgc ttc agt aga cct cgc cat** ttt ctc [396]
PEREGRINE cac agt aga agc tgg agt tgg aac agg atg aac cgt ata ccc ccc ctt agc agg caa cct agc cca cgc cgg tgc ttc agt aga cct ggc tat ctt ctc [396]
MERLIN cac agt aga agc tgg agt cgg aac agg atg aac cgt ata ccc ccc ctt agc agg caa cct agc cca tgc cgg tgc ctc agt aga cct ggc cat ctt ctc [396]
GOLDENEAGLE aac agt aga agc agg ggc tgg cac cgg gtg gac tgt cta ccc ccc act agc cgg caa cat agc cca cgc cgg cgc ttc ggt aga cct agc tat ctt ttc [396]
GOSHAWK aac cgt aga agc agg agc cgg aac tgg atg aac tgt tta ccc tcc att agc cgg caa cat agc tca tgc cgg agc ctc agt aga cct agc cat ctt ctc [396]