DNA Fingerprinting

and its Application to Aviculture: Nice to Know, For the Birds, or Critical?

by Benny J. Gallaway, Ph.D.
AFA Director of Conservation/ President, LGL Genetics, Inc., Bryan, TX

Introduction

The article is about aviculture and genetics.....Wait! Please don't go to the next article just yet. I promise to skip all the introductory esoteric, coma-inducing mumbo-jumbo about DNA that we scientists feel compelled to describe so you can properly understand 1) how DNA works, and 2) how it can be assayed to provide an immutable mark or fingerprint for individual birds. These fingerprints can be used to determine the relatedness of pairs or potential pairs, to improve breeding success, and to establish genetically-viable aviary collections for the long-term. While we don't all need to be geneticists to use the results, I think most would agree that we do need to have the fingerprint technology in order to achieve long-term success in aviculture. Also, the availability of DNA fingerprints will remove the foremost impediment (no foolproof marking system) presently restricting free, or nearly-free, international trade in legitimate captive-bred birds.

The DNA technology for achieving the stated avicultural goals, in fact, already exists for all birds, but the process is slow, labor intensive, and therefore, expensive. The present high cost and the fact that there are so few laboratories which offer this service has resulted in the approach being impractical, albeit technically feasible. Recognition of the need for DNA fingerprinting in aviculture, led me as 1) Director of Conservation for the American Federation of Aviculture, and 2) as President of a private-sector research genetics company (LGL Ecological Genetics, Inc.) to apply to the National Science Foundation for a Small Business Innovative Research (SBIR) grant to develop the needed technology.

We were awarded a Phase I grant (Award DMI-9461111, effective 1 January 1995) that allowed us six (6) months to conduct preliminary research towards development of the needed technology. Depending on a successful outcome of a scientific peer-review of our Phase I research results, we would be able to compete on a national basis for the limited funds available from the National Science Foundation to support larger Phase II awards. We completed our Phase I research on schedule, and based upon the results of scientific peer reviews of our Phase I report and Phase II research proposal, we were awarded a 2-yr Phase II grant on 15 July 1996 (Award DMI-9529743).

In our Phase I study we developed the technology enabling automated DNA profiling of the Hyacinth Macaw (*Anodorhynchus hyacinthinus*) and demonstrated that the method could identify individuals and their progeny with certainty, as well as establish the degree of relatedness among individuals. We also demonstrated that the method developed for this single species of macaw appeared to extend to other species of macaws, as well as to other groups of psittacines, at least in part. In our Phase II research we will develop the technology for all the major groups of psittacine birds involved in international trade and domestic aviculture.

Below, I provide a background on DNA profiling (fingerprinting), outline the approach we are using, summarize the results of our Phase I research, and describe the research plan for extending our technology to other groups of exotic birds. It should also be noted that we (Gallaway et al. 1995) have previously used the same approach to develop DNA fingerprinting protocols for the Emu (*Dromaius novaehollandiae*). In that study, we were able to demonstrate that the degree of relatedness was correlated with production attributes, and showed that genetic-based pairings of captive stock could markedly improve production and stabilize the captive gene pool.

Background

All you have to know for now is that DNA contains "genes" which are located at specific places on the DNA strands which are found in the nucleus of every cell in the body. These places are referred to as loci (plural for "locus," a specific address on the DNA molecule). At each locus, a gene has two elements (called alleles); one of which was donated by mom and one from pop. These alleles often have many forms in the overall population, and may govern things such as eye color, hair color, length of the big toe, etc. While there can be many alleles in the population as a whole, an individual is restricted to only two, one from mom and one from pop.

Some of the alleles, say for brown eyes, dominate others, say for blue eyes. If pop gave you his allele for brown eyes and mom gave you her allele for blue eyes, you are going to have brown eyes. Period. You, however, carry (and can pass along to your offspring) the allele for blue eyes. You are said to be "heterozygous" (with two different alleles) at that gene (locus). If both your eye color alleles had been for brown eyes you would be said to be "homozygous" (the two alleles are the same) at that gene.

For you to have the recessive trait of blue eyes both of your parents would have to have carried the allele for blue eyes. Thus, don't worry too much if your parents both have brown eyes and you ended up having blue eyes. They were both heterozygous at the eye color locus and each donated you the allele for blue eyes (you are homozygous). However, if both your parents have blue eyes and you have brown eyes, it might be time for a family chat.

Now, lets think about the population. There are brown eyes, black eyes, green eyes, blue eyes, hazel eyes and on and on. However, some colors are more prevalent than others; i.e., brown eyes are more frequently seen than blue eyes (I think), black eyes are more common than green eyes, etc. Sometimes even a very rare color can be observed (say red). If you have one of these rare alleles, it can provide a marker showing you to be (outside of your immediate family) genetically distinct from most of the population.

Now, imagine that you were assayed for your particular allele make-up over 10 or 12 locations on the DNA molecule where, at each location, there were up to as many as 50 different alleles that occur in the population as a whole. Under these conditions, you can be positively separated from all other humans (except an identical twin) for all practical purposes. The likelihood of there being another individual having the same compliment of alleles at these locations can often be shown to be on the order of 1 in a billion or so. If every gene was assayed, every individual would be found to be at least a little bit different. However, (except for identical twins) we don't have to do every gene to be sure, for all practical purposes, that tested individuals are unique.

Phase I Research Results

In our preliminary study, we isolated nine (9) locations (loci) on the Hyacinth Macaw genome containing microsatellite markers which exhibited variable lengths (different alleles) at that location. Each location (locus) thus identified was characterized by a minimum of two (2) alleles, ranging up to as many as 16 alleles (Table 1). As shown in Table 1, each allele at a locus was given a single character allele-designation (for example 0-9, or A-Z) to simplify comparisons of the multi-locus genotypes.

As we discussed earlier, the alleles occur as diploid genotypes; i.e., meaning one allele is donated to the offspring from each parent. To determine the uniqueness of an individual, one needs to know the frequency of each allele in the overall population. Our sample included but 22 individual birds, including only 11 which were not related by pedigree to other birds in the sample. Despite the small sample size upon which to estimate allele frequency of the population, the probability that a Hyacinth Macaw drawn at random from the overall population would not be genetically distinct from all other birds considering all nine loci was, following Fredholm and Winterø (1955), 1 in 10,650. Following the National Research Council (1992) method of calculation, the probability of another Hyacinth Macaw having the same genotype as one of our 22 test birds ranged from 1 in 1,048 to 1 in 40 billion. While two (2) siblings in our samples did share a common genotype over the nine (9) loci, it is clear that a combination of increased sample sizes and more loci would result in the ability to identify individuals and their progeny with certainty.

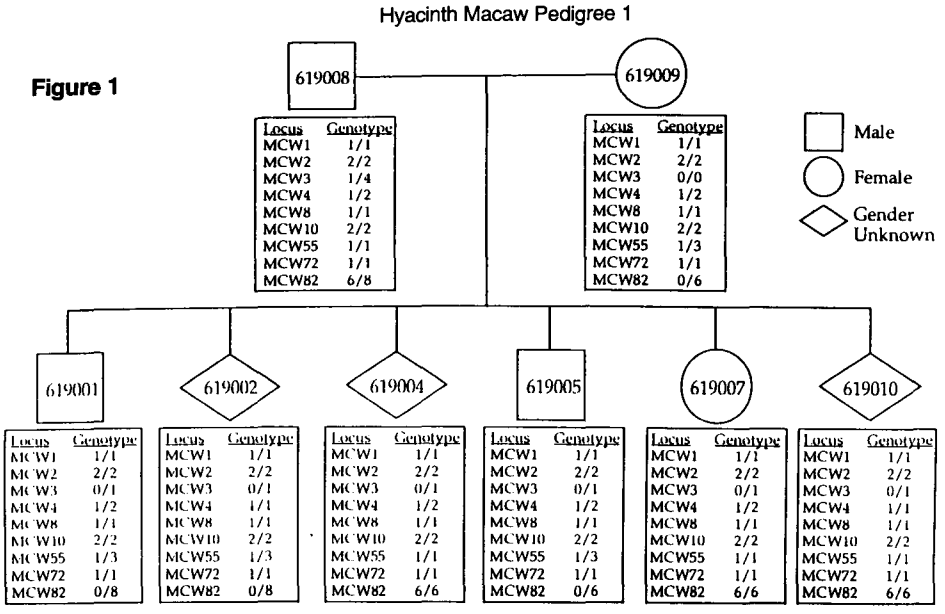
A Hyacinth Macaw pedigree investigated in Phase I is shown by Figure 1. The genotype for each of the parents is shown at the top of the figure, and the genotypes for each of six offspring (2 males, 1 female, and 3 of undetermined sex) are shown across the bottom of the figure. Note that at the first locus (MCW1), both parents had a 1/1 genotype, or were homozygous. Had any of the putative offspring had a genotype other than 1/1 at this locus, it would be conclusive that it was not an offspring of the claimed parental pair, since only allele 1 was represented in the parents. The same is true for loci MCW2 (2/2 in both parents), MCW8 (1/1), MCW10 (2/2), and MCW72 (1/1). Note that all the offspring have the proper homozygous genotype for these four (4) loci. In contrast, both parents were heterozygous at the MCW82 locus; i.e., two different alleles were present. The male had a MCW82 genotype of 6/8; the female was 0/6. If the 0 allele from the female was present in an offspring, the other allele present at MCW82 could only come from the male, and would have to be either allele 6 or 8. Using the same process of elimination, parentage can be documented or rejected. When a failure to reject

Table 1

Microsatellite allele data for Hyacinth Macaws. The mean or range of the GENESCAN™ allele measurements and whole number allele size (number of nucleotides) are shown, as well as the single-character allele designation.

Locus	Allele Designation	Allele Size	Mean	SD	Max	Min	Range
MCW1	1	173	173.19	0.05	173.27	173.15	0.12
MCW1	2	175	175.18	0.00	175.18	175.18	0.00
MCW2	1	123	123.43	0.06	123.49	123.34	0.15
MCW2	2	125	125.37	0.07	125.48	125.20	0.28
MCW2	3	127	127.27	0.06	127.31	127.23	0.08
MCW3	0	183	182.89	0.06	183.01	182.84	0.17
MCW3	1	196	196.21	0.06	196.31	196.08	0.23
MCW3	2	198	198.13	0.09	198.20	198.01	0.19
MCW3	3	200	200.01	0.03	200.06	199.98	0.08
MCW3	4	202	201.99	0.03	202.04	201.96	0.08
MCW4	1	274	273.63	0.08	273.75	273.47	0.28
MCW4	2	276	275.67	0.10	275.85	275.49	0.36
MCW4	3	278	277.82	0.06	277.91	277.70	0.21
MCW8	1	269	268.53	0.07	268.66	268.40	0.26
MCW8	2	271	270.69	0.00	270.69	270.69	0.00
MCW10	1	200	200.56	0.00	200.56	200.56	0.00
MCW10	2	204	204.47	0.05	204.56	204.31	0.25
MCW55	1	340	340.25	0.13	340.52	340.07	0.45
MCW55	2	342	342.38	0.00	342.38	342.38	0.00
MCW55	3	344	344.36	0.09	344.50	344.22	0.28
MCW72	1	183	182.87	0.05	182.95	182.79	0.16
MCW72	2	185	184.77	0.02	184.78	184.75	0.03
MCW82	0	338	337.96	0.30	338.37	337.74	0.63
MCW82	1	339	339.09	0.16	339.20	338.98	0.22
MCW82	2	344	343.66	0.05	343.72	343.61	0.11
MCW82	3	346	345.78	0.00	345.78	345.78	0.00
MCW82	4	347	346.91	0.01	346.91	346.90	0.01
MCW82	5	350	349.90	0.24	350.07	349.73	0.34
MCW82	6	354	353.62	0.18	353.80	353.29	0.51
MCW82	7	355	354.61	0.00	354.61	354.61	0.00
MCW82	8	356	355.61	0.10	355.69	355.50	0.19
MCW82	9	359	358.73	0.00	358.73	358.73	0.00
MCW82	A	363	362.81	0.00	362.81	362.81	0.00
MCW82	B	366	366.24	0.00	366.24	366.24	0.00
MCW82	C	368	368.28	0.06	368.32	368.24	0.08
MCW82	D	384	384.14	0.02	384.15	384.12	0.03
MCW82	E	380	380.24	0.05	380.29	380.20	0.09
MCW82	F	376	376.26	0.00	376.26	376.26	0.00

Figure 1



Locus	619008	619009	619001	619002	619004	619005	619007	619010
MCW1	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1
MCW2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2
MCW3	1/4	0/0	0/1	0/1	0/1	0/1	0/1	0/1
MCW4	1/2	1/2	1/2	1/1	1/2	1/2	1/2	1/1
MCW8	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1
MCW10	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2
MCW55	1/1	1/3	1/3	1/3	1/1	1/3	1/1	1/1
MCW72	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1
MCW82	6/8	0/6	0/8	0/8	6/6	0/6	6/6	6/6

Pedigree of a Hyacinth Macaw family based on microsatellite alleles represented at nine loci.

claimed parentage is achieved, the odds of the chick possibly being from any other pair are so astronomically low that the parentage claim can be accepted.

Now, how related are the parental birds in Figure 1? The answer is obtained by means of a similarity index calculated as the number of alleles held in common by the pair, divided by the total number of alleles, taking into account the frequency in the overall population of the alleles held in common. For our pair in Figure 1, they share 14 of the 18 alleles represented or 28 of the 36 total alleles that might have been possible. The relatedness index on this basis is 0.78. However, the shared alleles are common in the overall population, whereas the differing alleles are rare or less common. When the frequency distribution of the alleles is taken into account, the relatedness index decreases to 0.49. The birds in this breeding pair of Hyacinth Macaws are highly related, possibly even a sibling pair since, in diploid animals, brothers and sisters share 50% of their genes on average. This pair of birds were wild-caught, imported birds that were paired based on the assumption that they were unrelated. Obviously, this assumption was incorrect.

This particular pair of birds typically produce two (2) clutches of two (2) fertile eggs per year, but only 50% of the eggs develop to hatchlings. The cause of this low hatch rate is likely attributable to the high degree of relatedness of the parents. Low hatch rates seem characteristic for many Hyacinth Macaws in U.S. aviculture. Genetic assays are needed to determine the status of the captive population of Hyacinth Macaws with regard to inbreeding potential and to define an overall genetic management plan.

In Table 2, we show a matrix of relatedness indices for 7 male and 10 female Hyacinth Macaws included in our study. Note that male 3 and female 6 were the pair used in our example above. Had the owner of this pair been a member of a captive breeding cooperative which maintained genetic profiles of all birds available for pairing, female 6 could have been paired with male 7 (Relatedness Index = 0.145) and male 3 could have been paired with either of females 4 or 10 (Relatedness Index = 0.293) in order to preserve a higher level of genetic diversity in the captive population. Only about half (47%) of the possible pairings shown in Table 2 would be "good" genetic matches. Thus, if you pair Hyacinth Macaws at random, about half of your pairs are likely going to exhibit genetic disorders. Such matrices can be used to define trade requirements on a genetic basis; and the profile information necessary to generate this information provides an immutable mark that will ensure that any "laundering" of birds could be detected if attempted.

Phase II Research Plan

In Phase I of our SBIR project, we developed nine DNA markers for Hyacinth Macaws. These markers varied in size among individuals, were inherited in the expected fashion, and the results were repeatable. They are useful for individual identity, as only 2 of 22 Hyacinth Macaws shared the same nine-locus genotype. The two macaws with the same genotypes were siblings from a highly-related pair (themselves probably siblings), and several of the other macaws we analyzed were also from sibling groups. We are confident that these microsatellites, and others we will develop in Phase II, will provide very high probabilities of identity for macaws. The nine microsatellite loci developed in Hyacinth Macaws were successfully amplified in several other macaw species (Scarlet, Blue and Gold, Illiger's, Military, Green-winged, and Buffon's). Three or four of the macaw microsatellite loci (depending on the species) were also amplified in cocka-

Table 2

A matrix of relatedness for 7 male and 10 female Hyacinth Macaws.

	Males						
	1	2	3	4	5	6	7
Females							
1	0.898	0.757	0.508	0.240	0.269	0.139	0.130
2	0.387	0.393	0.444	0.229	0.357	0.235	0.423
3	0.612	0.757	0.677	0.291	0.256	0.296	0.225
4	0.249	0.253	0.293	0.215	0.792	0.312	0.369
5	0.612	0.757	0.677	0.291	0.256	0.296	0.225
6	0.755	0.889	0.490	0.236	0.284	0.242	0.145
7	0.240	0.244	0.438	0.842	0.181	0.297	0.163
8	0.333	0.338	0.542	0.723	0.263	0.302	0.232
9	0.333	0.446	0.493	0.192	0.311	0.198	0.377
10	0.408	0.414	0.293	0.317	0.558	0.219	0.318

toos, Old World parrots, and New World parrots. For these other species, levels of variation need to be assessed for utility in individual identity. Additional loci are also needed for these other species for reliable individual identity, as not all of the loci developed for macaws will work in the other groups.

It has been observed in other taxa that only a proportion of markers developed for a given species are useful across related species (e.g., Ellegren 1992, Garza et al. 1995). The proportion of microsatellite loci which will amplify in different species varies among taxa. For example, Fredholm and Winterø (1995) found 16 of 20 markers developed in dogs also worked in foxes. Sixteen per cent of 153 markers developed in rats worked in mice, and 12 % of 166 mouse markers worked in rats (Kondo et al. 1993). About 30% of markers developed in cattle work in deer (J. C. Patton, unpublished data) and about 60% of markers developed in cattle or sheep work in both species (Moore et al. 1991). In contrast, a relatively high proportion (92%) of markers developed in chickens also worked in turkeys (Levin et al. 1995).

Table 3

Species for which microsatellite enriched libraries will be developed for the purpose of generating PCR based microsatellite repeat loci.

Common Name	Scientific Name	Geographic Range
1. Yellow-crowned Amazon	*Amazona ochrocephala*	Mexico, Central & South America
2. Sun Conure	*Aratinga solstitialis*	Brazil, Venezuela, and Guianas
3. African Grey Parrot	*Psittacus erithacus*	Africa
4. Eclectus Parrot	*Eclectus roratus*	N. Australia, New Guinea, Moluccas through the Soloman Islands
5. Rainbow Lory	*Trichoglossus haematodus*	Bali through Indonesia and New Guinea and Loyalty Islands, Australia and Tasmania
6. Umbrella Cockatoo	*Cacatua alba*	Moluccas and surrounding islands

Our Phase II technical research will focus on development of new markers for six species (Table 3). This will include selection of several (6-10) loci with adequate levels of polymorphism, appropriate allele sizes for simultaneous analysis in a single lane of an electrophoretic gel, and PCR™-amplification reaction conditions which allow multiplexing (running all the markers in a single test). These characteristics are crucial for developing markers which can be used for individual identity, and which can be analyzed economically. Minimizing the number of PCR™ reactions and gel-lanes needed for analysis of individual birds will result in the lowest cost per sample, and the most competitive system possible.

Literature Cited

Ellegren, H. 1992. Polymerase-Chain-Reaction (PCR™) analysis of microsatellites—a new approach to studies of genetic relationships of birds. Auk 109:886-895.

Fredholm, M., and A. K. Winterø. 1995. Variation of short tandem repeats within and between species belonging to the canidae family. Mammalian Genome 6:11-18.

Gallaway, B. J., J. C. Patton, K. Coldwell, and W. Sealey. 1995. Ratite genetics. P. 63-78 in C. Drenowatz [Ed.] Ratite Encyclopedia. Ratite Records Incorporated, San Antonio, Texas. 478 p.

Garza, J. C., M. Slatkin, and N. B. Freimer. 1995. Microsatellite allele frequencies in humans and chimpanzees, with implications for constraints on allele size. Molecular Biology and Evolution 12:594-603.

Kondo, Y., M. Mori, T. Kuramoto, J. Yamada, J. S. Beckmann, D. Simon-Chazottes, X. Montagutelli, J. L. Guenet, and T. Serikawa. 1993. DNA segments mapped by reciprocal use of microsatellite primers between mouse and rat. Mamm. Genome 4:571-576.

Levin, I., H. H. Cheng, C. Baxter-Jones, J. Hillel. 1995. Turkey microsatellite DNA loci amplified by chicken-specific primers. Animal Genetics 26:107-110.

Moore, S. S., L. L. Sargeant, T. J. King, J. S. Mattick, M. Georges, D. J. S. Hetzel. 1991. The conservation of dinucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species. Genomics 10:654-660.

National Research Council (U.S.). 1992. DNA technology in forensic science. Committee on DNA Technology in Forensic Science, Board on Biology, Commission on Life Sciences. ISBN 0-309-04587-8.