

## TECHNICAL ADVANCES

# Improving the reliability of molecular sexing of birds using a W-specific marker

DAIZABURO SHIZUKA and BRUCE E. LYON

*Department of Ecology and Evolutionary Biology, Earth and Marine Sciences Building, University of California, Santa Cruz, CA 95064, USA*

## Abstract

Molecular techniques for identifying sex of birds utilize length differences between CHD-Z and CHD-W introns, but in some cases these methods can lead to sexing errors. Here we show that an additional W-specific primer can be used in conjunction with a pre-existing sexing primer pair to dramatically improve the reliability of molecular sexing methods. We illustrate the approach with American coots (*Fulica americana*), a species with CHD-Z polymorphism that could not be accurately sexed using traditional methods. We developed a reverse primer GWR2 designed to sit within the intron of the W chromosome and amplify a distinctively small DNA fragment that serves as a W-specific marker. Analysis of known-sex individuals indicates that this W-specific primer provides an efficient and reliable protocol to identify the sex of *F. americana*. The development of such sex-specific primers will likely increase the reliability of molecular sexing methods in other birds as well. Comparisons between CHD-Z alleles of coots and common moorhens (*Gallinula chloropus*) revealed that CHD-Z polymorphism evolved separately in these two closely related species. We discuss the implications of repeated evolution of CHD-Z polymorphisms among birds.

*Keywords:* CHD1, molecular sexing, Rallidae, sex chromosomes, sexing errors

*Received 29 April 2008; revision accepted 17 July 2008*

## Introduction

The ability to identify the sex of individuals is critical for investigating many subjects in evolution and ecology, including sex ratio evolution, sexual selection, parental care strategies, demography and conservation. However, morphological identification of sex in birds can often be difficult in monomorphic species and sexually immature individuals, and thus, the innovation of molecular sexing techniques has opened new doors for researchers in many of these fields (Ellegren & Sheldon 1997; Griffiths *et al.* 1998; Kahn *et al.* 1998; Fridolfsson & Ellegren 1999). For example, the development of molecular sexing techniques has revolutionized the study of sex allocation and sex ratios (Ellegren & Sheldon 1997). Molecular sexing has also enabled researchers to determine the sex of rare monomorphic birds such as the Po'ouli (*Melamprosops placosoma*) and the Spix macaw (*Cyanopsitta spixii*), a species that is now extinct

in the wild, thus aiding conservation efforts in these species (Griffiths & Tiwari 1995; Groombridge *et al.* 2004). More generally, chromosome-helicase-DNA binding protein 1 (or CHD1), the sex-linked gene locus used for these molecular sexing techniques, has contributed to a greater understanding of sex chromosome evolution (Ellegren & Fridolfsson 1997).

The most widely used molecular methods for sexing birds exploit the size differences between introns of the CHD gene on the Z and W sex chromosomes (Griffiths *et al.* 1998; Kahn *et al.* 1998; Fridolfsson & Ellegren 1999). Heterogametic (ZW) females are expected to have two different-sized introns while homogametic (ZZ) males should only have one intron size. However, recent studies have revealed the existence of polymorphisms in the introns of CHD-Z in several species, which complicates the identification of sex because heterozygote (e.g. ZZ') males will also have two different-sized introns (Dawson *et al.* 2001; Lee *et al.* 2002; Robertson & Gemmill 2006). We do not yet know how many species possess polymorphism in the CHD-Z gene, nor do we know how these polymorphisms arise and are

Correspondence: Daizaburo Shizuka, Fax: 831-459-5353;

E-mail: Shizuka@biology.ucsc.edu

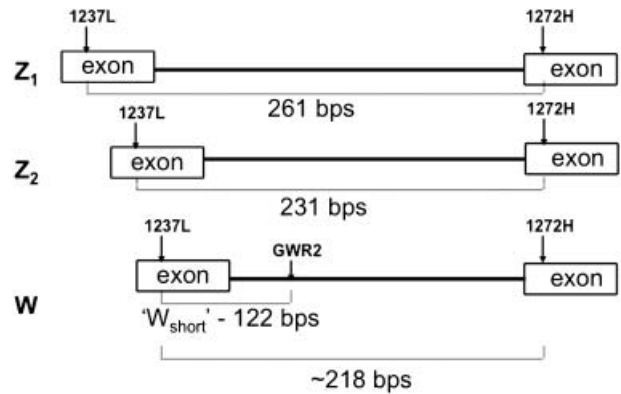
maintained within clades. Polymorphisms may be difficult to detect without adequate sampling of known-sexed individuals before conducting a full study, which can often be cumbersome or even impossible.

One way to improve the reliability of molecular sexing techniques in birds is to amplify an additional, W chromosome-specific DNA fragment of a different size than those produced by the pre-existing sexing primer pairs to independently confirm the existence of the W chromosome. Here, we describe a case of CHD-Z polymorphism in the American coot (*Fulica americana*), and offer a simple polymerase chain reaction (PCR)-based method using a W-specific marker that allows reliable identification of sex in this species. We expect that the general methods are applicable to a wide range of avian species. Furthermore, we compare the genetic structure of the polymorphism found in coots with a similar polymorphism found previously in the common moorhen (*Gallinula chloropus*, Lee *et al.* 2002) to determine how the CHD-Z polymorphism has arisen in these two closely related species.

## Methods

We analysed DNA from 12 known-sex American coots (nine males and three females) captured in the field and from museum specimens. We obtained purified DNA from five blood samples from known-sex adult American coots from our field sites near Williams Lake, British Columbia in May 2006. The sex of four colour-banded individuals were identified by vocalizations during subsequent behavioural observations. The sex of one individual found dead at the site was identified by dissection. We also obtained six tissue samples from known-sexed individuals from the California Academy of Sciences (San Francisco, CA) and the Museum of Vertebrate Zoology at UC Berkeley (CAS: nos 90816, 90580; MVZ: nos 179206, 179796, and two uncatalogued specimens). An additional tissue sample was collected from a dead specimen collected by a wild-animal care facility in Marin County, California, in August 2006 and dissected for sex identification. We used standard 24:1 chloroform-isoamyl extraction method to isolate DNA from the tissue and blood samples.

We first determined the efficacy of the traditional PCR-based sexing method using the primer pair 1237 L/1272H developed by Kahn *et al.* (1998). This primer pair sits on exons that flank an intron section of the CHD1 gene on both Z and W chromosomes (Fig. 1). Each 25  $\mu$ L reaction consisted of 9.3  $\mu$ L of H<sub>2</sub>O, 2  $\mu$ L of 2 mM dNTPs, 0.2  $\mu$ L of 5 U/ $\mu$ L *Taq*-polymerase, 2.5  $\mu$ L 10 $\times$  buffer, 1  $\mu$ L of each 20  $\mu$ M primer and 8  $\mu$ L of purified DNA diluted to approximately 20 mM. PCR amplification was conducted on an Applied Biosystems GeneAmp 9700 Thermocycler under the following conditions: 92  $^{\circ}$ C for 2 min, followed by 30 cycles of 92  $^{\circ}$ C for 30 s, 57  $^{\circ}$ C for 45 s, 72  $^{\circ}$ C for 45 s



**Fig. 1** A visual representation of the section of the CHD1 gene used for sex identification. Arrows indicate the approximate binding site for each primer, and dotted lines indicate the section of the gene that is amplified by each primer pair. Fragment size of the whole CHD-W intron is based on the alignment of 1237L/1272H primer pair to published CHD-W sequence of common moorhens, and is therefore indicated as approximate size. All other fragment sizes are based on direct sequencing.

with a final extension of 72  $^{\circ}$ C for 5 min. PCR products were separated on a 2–3% NuSieve agarose gel and stained with ethidium bromide or SYBR Gold.

To increase the reliability of sex identification, we also developed a primer designed to specifically amplify a fragment only from the CHD-W intron. If intron sequences between the W- and Z-chromosomes are distinct enough, a primer that sits within the intron of the CHD-W gene will not be able to bind to the CHD-Z gene and would therefore be W-specific. We used the published sequence of CHD-W intron of the common moorhen (Lee *et al.* 2002, GenBank Accession no. AY149278) to develop an additional reverse primer, GWR2 (5'-CCTGTAAAACCACCCAACC-3'), designed to sit within the intron sequence of CHD-W in coots (Fig. 1) and work under similar PCR conditions as the 1237L/1272H primer pair. Primer design was conducted using the web-based program Primer 3 (<http://frodo.wi.mit.edu/primer3/input.htm>). We used a combination of the 1237L/1272H/GWR2 primers to amplify the genes of the same known-sexed individuals under the same PCR conditions as above. This was essentially a multiplex reaction of two primer pairs, the 1237L/1272H pair that amplified the entire CHD1 intron of both Z and W chromosomes, and an additional 1237L/GWR2 pair that amplified a short fragment only when the W chromosome was present (Fig. 1).

To be certain that we amplified the targeted regions of the genome, we sequenced the two types of CHD-Z intron fragments amplified by the 1237L/1272H primer pair as well as the extra DNA fragment produced with the addition of the GWR2 primer. We sequenced each fragment using both forward and reverse primers on an ABI 3100 sequencer

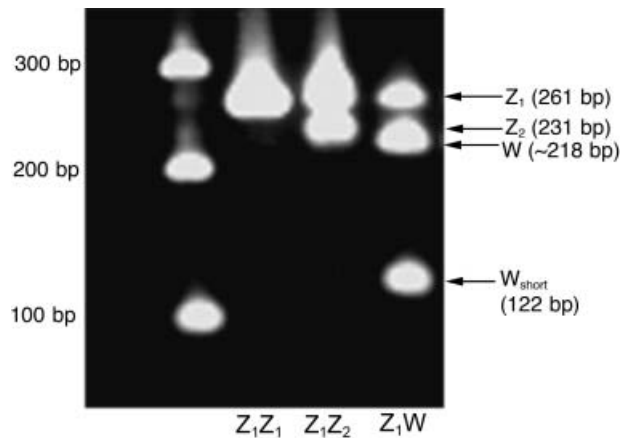
and produced consensus sequence using the program Sequencher. We then aligned and compared the sequences generated by the GWR2 primer with CHD-Z fragments of American coots and the published sequence of CHD-W from the common moorhen to confirm that this was indeed associated with CHD-W of the coot.

Finally, to investigate the pattern of sex chromosome evolution between coots and moorhens, we aligned and compared our consensus sequences of American coot CHD-Z introns with the published sequences of the same loci for common moorhens (GeneBank Accession nos AY149276 and AY149277). Alignments were conducted using the web-based program ClustalW2 (<http://www.ebi.ac.uk/Tools/clustalw2/index.html>).

**Results**

*CHD-Z polymorphism*

Amplification of the CHD1 intron with the primer pair 1237L/1272H revealed a CHD-Z polymorphism: a longer, more common Z<sub>1</sub> allele (261 bp) and a less frequent Z<sub>2</sub> allele (231 bp). Because the CHD-Z<sub>2</sub> intron is close in size to the CHD-W intron (~218 bp, based on alignment of primers to published moorhen sequence), the traditional use of the PCR-based method of sex identification could have easily led to the misidentification of the common Z<sub>1</sub>Z<sub>2</sub> males as females on an agarose gel (Fig. 2). We use the notation Z<sub>1</sub> and Z<sub>2</sub> to explicitly distinguish these alleles from the already known Z and Z' alleles of the closely related common moorhen (Lee *et al.* 2002). As we show below, different



**Fig. 2** PCR amplifications of three different genotypes using the combination of 1237L/1272H/GWR2 primers, run and imaged on a NuSieve agarose gel and stained with SYBR Gold. Note that without the presence of the short W-specific fragment produced by the 1237L/GWR2 pair, Z<sub>1</sub>Z<sub>2</sub> males and Z<sub>1</sub>W males are extremely difficult to differentiate, and that the short W fragment is diagnostic.

notation from that used for the moorhen is justified because of sequence differences between the two species. Sequencing revealed one single-nucleotide polymorphism (SNP) and a 30-bp insertion/deletion (indel) as the cause of this length polymorphism (Fig. 3). Z<sub>1</sub>Z<sub>2</sub> heterozygotes occurred in 22% of males (two of nine males). No Z<sub>2</sub>Z<sub>2</sub> homozygotes were identified in our sample. No length polymorphisms were detected in CHD-W.

F. am_Z1	GGTGTCTCTTGTTCGACTGGCTTCCACTTTTATGTTGCTGTTGGTTTA	50
F. am_Z2	-----G-----*****	
G. ch_Z1	-----A-----	
G. ch_Z2	-----C-----	
F. am_Z1	GTTTATTTCTTCGGGATTATTATTGGGTTTGGGTTGTGGTTTGTGGGN	100
F. am_Z2	*****C-----T	
G. ch_Z1	-----G-----T	
G. ch_Z2	-----G-----T	
F. am_Z1	TTTTTTC*TCAACACATGTTTTTGACAGACTAGGTAAACTTAGGTTTGT	150
F. am_Z2	-----*---A-----T-----	
G. ch_Z1	-----*---A-----T-----	
G. ch_Z2	-----C---G-----C-----	
F. am_Z1	TAATTGCACAGCTTTGAACTATTATTCTGAAGTTGCAGATCAGCTTTAA	200
F. am_Z2	T-----AG-----	
G. ch_Z1	C-----CA-----	
G. ch_Z2	T-----CA-----	
F. am_Z1	TGGAAGTGAAGGGAGGCGCAGC	222
F. am_Z2	-----	
G. Ch_Z1	-----	
G. ch_Z2	-----	

**Fig. 3** Alignment of CHD-Z intron sequences of American coots (*F. am*) and common moorhens (*G. ch*). Asterisks indicate gaps and dashes refer to nucleotide positions identical to the sequence above. The symbol 'N' (position 112 on *Fulica americana* allele Z<sub>1</sub>) denotes an ambiguous position where the nucleotide could not be resolved (G or T). Because the coot and moorhen sequences were recovered using different primers (P2/P8 and 1237L/1272H, respectively), we show only the overlapping region.

*Sexing using GWR2 primer*

The 1237 L/GWR2 primer set amplified a short fragment (122 bp) of the CHD-W chromosome in American coots, which we will refer to as CHD-W<sub>short</sub>. This fragment only appeared on known female samples (three of 12 total samples), and when sequenced, it aligned well (98%, with two SNPs) with the W chromosome of the common moorhen. W<sub>short</sub> did not align well with the coot Z<sub>1</sub> (55%) or Z<sub>2</sub> (60%) alleles. Therefore, we conclude that GWR2 indeed amplified a fragment of the CHD-W intron.

We were then able to successfully run the three primers together in one PCR. As a result, one PCR amplified a unique set of DNA fragments for each genotype that were accurate and unambiguous (Fig. 2). Female individuals could always be identified by the presence of the W<sub>short</sub> fragment, regardless of the similarity of the whole CHD-W intron to the CHD-Z<sub>2</sub> intron.

*Sequence divergence from common moorhen*

The source of the CHD-Z polymorphism in American coots was different from the source of CHD-Z polymorphism in the common moorhen. Published intron sequences from CHD-Z and CHD-Z' alleles of common moorhens shows 4–6 single nucleotide polymorphisms (SNPs) and a single base pair insertion/deletion (indel) as the source of this polymorphism (Lee *et al.* 2002), while the CHD-Z polymorphism in coots were due to a single SNP and a 30-bp indel (Fig. 3). Of the four alleles (moorhen Z/Z' and coot Z1/Z2), the coot Z2 allele is the only one containing the 30-bp deletion, suggesting that this deletion occurred after the evolutionary divergence of the coot and moorhen clades. Conversely, three of the four SNPs observed in moorhens do not occur in coots, which suggests that these mutations occurred in the moorhen clade after divergence from coots.

**Discussion**

While molecular sexing methods of birds can often be straightforward, there are several pitfalls that can lead to misidentification of sex. Our results show that American coots carry a polymorphism in the CHD-Z locus, and this length variation can make the application of the standard PCR-based methods difficult. In particular, the similarity in length of the CHD-Z<sub>2</sub> and CHD-W introns would result in misidentification of heterozygous Z<sub>1</sub>Z<sub>2</sub> males as females, using traditional molecular methods. Polymorphism in the CHD-Z intron has been documented in several species, and could be a source of sexing errors in other species that have not yet been tested (Dawson *et al.* 2001; Lee *et al.* 2002). We combined the traditional molecular sexing technique with an additional W-specific primer to develop a PCR-based protocol that can determine the sex of American

coots reliably and efficiently. We were able to design a W-specific primer that worked under the same PCR conditions as an existing primer pair for sex identification (1237L/1272H), which permitted a highly efficient multiplex design. Using this protocol, PCR amplification produces unambiguously unique DNA fragments for each genotype that can be resolved on agarose gels.

For many avian species, these modifications of the traditional PCR-based sexing methods may not be needed. Researchers now have access to three primer pairs (P2/P8, Griffiths *et al.* 1998; 1237L/1272H, Kahn *et al.* 1998; 2550F/2718R, Fridolfsson & Ellegren 1999) that amplify two different introns of the CHD1 gene (P2/P8 and 1237L/1272H target overlapping regions of the same region of the gene). Researchers will often be able to produce unambiguous results using one of these primer pairs. However, the existence of Z chromosome polymorphisms means that the standard methods will have to be verified for each new species, using DNA samples from known sex individuals. Our method of using a W-specific primer will be useful when these preliminary steps fail to produce reliable results. The presence of a female-specific DNA fragment can produce unambiguous DNA fragment combinations even if intron lengths between CHD-Z and CHD-W introns are too similar to interpret on an agarose gel (Griffiths *et al.* 1998), or when CHD-Z polymorphism leads to two different intron fragment lengths in males.

W-specific primers will have to be developed for each taxon in order to apply this method, but this task can be accomplished relatively easily. Because the rate of molecular evolution of the W-chromosome is among the lowest found among vertebrate chromosomes, the CHD-W intron sequence is often conserved between related species (Ellegren & Fridolfsson 1997). We were able to develop the GWR2 primer without sequencing the entire CHD-W intron sequence from American coots, but instead by simply using the previously published sequence of the common moorhen. To develop an appropriate W-specific primer for a given species, we suggest that researchers start by looking for the CHD-W sequence from the closest related taxa available on GenBank. Using a primer-design software, one can identify potential W-specific forward or reverse primers that sit within the CHD-W intron. The most appropriate primers would be those that work under the same PCR conditions as the pre-existing primers that amplify that intron.

Our comparisons of CHD-Z sequences from American coots and common moorhens show that these two closely related species independently evolved and retained a polymorphism in this intron at relatively high frequencies. The common moorhen is the only species to date in which a clear fitness effect of CHD-Z polymorphism has been demonstrated. In this species, the CHD-Z alleles show a rare pattern of genetic underdominance, where heterozygous

male juveniles suffer extremely high levels of mortality (Lee *et al.* 2002). Given the close evolutionary relationship between coots and moorhens (Olson 1973; Livezey 1998; Fain *et al.* 2007), a similar effect might be expected in American coots. However, given that the sequence comparisons suggest the independent origin of CHD-Z polymorphism in the two species, it is unclear whether the genetic effects of the CHD-Z intron would be the same. Regardless, we do not expect that all instances of CHD-Z polymorphisms are driven by patterns similar to those seen in common moorhens. Therefore, the infrequent but widespread existence of CHD-Z polymorphisms across bird species begs the question, why do CHD-Z polymorphisms arise and remain in some species but not in others? A thorough investigation of the evolutionary origins and maintenance of such polymorphisms would contribute to a greater understanding of the evolution of sex chromosomes, as well as to further development of molecular sexing techniques.

### Acknowledgements

We thank J. Clarke, J. Mehlhaff, C. Nelson, and G. Peters for field help, and J. Herrick for access to field sites on his property. We also thank California Academy of Sciences (San Francisco, CA), Museum of Vertebrate Zoology (Berkeley, CA), and WildCare (San Rafael, CA) for kindly providing us with specimens. We are grateful to A. Corl, K. Deiner, K. Dlugosh, E. Ferree, G. Pogson and K. Rich for help with development of ideas and technical assistance. Field samples were collected under permits from the Canadian Wildlife Service and the University of California, Santa Cruz IACUC. This study was funded by NSF grant IOS 0443807.

### References

- Dawson DA, Darby S, Hunter FM, Krupa AP, Jones IL, Burke T (2001) A critique of avian CHD-based molecular sexing protocols illustrated by a Z-chromosome polymorphism detected in auklets. *Molecular Ecology Notes*, **1**, 201–204.
- Ellegren H, Sheldon BC (1997) New tools for sex identification and the study of sex allocation in birds. *Trends in Ecology & Evolution*, **12**, 255–259.
- Ellegren H, Fridolfsson AK (1997) Male-driven evolution of DNA sequence in birds. *Nature Genetics*, **17**, 182–184.
- Fain MG, Krajewski C, Houde P (2007) Phylogeny of 'core Gruiformes' (Aves: Grues) and resolution of the Limpkin–Sungrebe problem. *Molecular Phylogenetics and Evolution*, **43**, 515–529.
- Fridolfsson AK, Ellegren H (1999) A simple and universal method for molecular sexing of non-ratite birds. *Journal of Avian Biology*, **30**, 116–121.
- Griffiths R, Double MC, Orr K, Dawson RJG (1998) A DNA test to sex most birds. *Molecular Ecology*, **7**, 1071–1075.
- Griffiths R, Tiwari B (1995) Sex of the last wild Spix's macaw. *Nature*, **375**, 454–454.
- Groombridge JJ, Massey JG, Bruch JC *et al.* (2004) An attempt to recover the Po'ouli by translocation and an appraisal of recovery strategy for bird species of extreme rarity. *Biological Conservation*, **118**, 365–375.
- Kahn NW, St John J, Quinn TW (1998) Chromosome-specific intron size differences in the avian CHD gene provide an efficient method for sex identification in birds. *Auk*, **115**, 1074–1078.
- Lee PLM, Brain PF, Forman DW, Bradbury RB, Griffiths R (2002) Sex and death: CHD1Z associated with high mortality in moorhens. *Evolution*, **56**, 2548–2553.
- Livezey BC (1998) A phylogenetic analysis of the Gruiformes (Aves) based on morphological characters, with an emphasis on the rails (Rallidae). *Philosophical Transactions of the Royal Society Series B: Biological Sciences*, **353**, 2077–2151.
- Olson SL (1973) Classification of Rallidae. *Wilson Bulletin*, **85**, 381–416.
- Robertson BC, Gemmill NJ (2006) PCR-based sexing in conservation biology: wrong answers from an accurate methodology? *Conservation Genetics*, **7**, 267–271.