High-throughput sequencing-based waterbird diet analysis: application to wintering herbivorous geese

Yuzhan Yang[1], Aibin Zhan[2], Lei Cao[2], Fanjuan Meng[2], Wenbin Xu[3]

[1]School of Life Sciences, University of Science and Technology of China, Hefei, Anhui, China

[2]Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, China

[3]Anhui Shengjin Lake National Nature Reserve Administration, Chizhou, Anhui, China

Corresponding Author:

Lei Cao[2]

18 Shuangqing Road, Haidian District, Beijing, China, 100085

Email address: caolei@ustc.edu.cn

Deleted: s'

Abstract

Food availability and diet selection are important factors influencing the abundance and distribution of wild waterbirds. In order to better understand changes in waterbird populations, it is essential to figure out what they feed on. However, analyzing diet could be difficult and inefficient using traditional methods, such as microhistologic observation. Here, we addressed this gap of knowledge by investigating diet of greater white-fronted goose *Anser albifrons* and bean goose *Anser fabalis*, which are obligate herbivores wintering, in China, mostly in Middle and Lower Yangtze River Floodplain. We firstly prepared a local plant reference library by selecting an optimal marker gene (P6 loop of chloroplast *trn*L intron) and amplifying the most common plants that these geese would consume. Then, utilizing DNA metabarcoding, we discovered 15 food items in total from feces of these birds. Of the 15 unique dietary sequences, 10 could be identified at species-level. As for greater white-fronted goose, 73% of sequences belonged to *Poaceae* spp., and 26% belonged to *Carex* spp. In contrast, almost all sequences of bean goose belonged to *Carex* spp. (99%). Using the same samples, microhistology provided consistent food composition with metabarcoding results for greater white-fronted goose, while 13% of *Poaceae* was recovered for bean goose. In addition, two other taxa were discovered only through microhistologic analysis. Although most of the identified taxa matched relatively well between the two methods, DNA metabarcoding gave taxonomically more detailed information. Discrepancies were likely due to biased PCR amplification in metabarcoding, low discriminating power of current marker genes for monocots, and biases in microhistologic analysis. The diet differences between two geese species might indicate deeper ecological significance beyond the

Deleted: better about

Deleted: of

Deleted: .10

Deleted: .52

Comment [1]: Spurious precision given the fecal pellets sampled (although a far greater number of sequences were generated). The reduced precision is easier to interpret and does not detract from presentation.

Deleted: .72

Deleted: Besides

Deleted: The discrepancy were was

Deleted: discrepancy

Deleted: ecology

scope of this study. We concluded that DNA metabarcoding provided new perspectives for studies of herbivorous waterbird diets and inter-specific interactions, as well as new possibilities to investigate interactions between herbivores and plants. In addition, microhistologic analysis should be used together with metabarcoding methods to integrate these information.

Deleted: s

Deleted: bring integrative

Introduction

Wetlands are one of the most important ecosystems in nature, and they harbor a variety of ecosystem services such as protection against floods, water purification, climate regulation and recreational opportunities (Brander, Flora & Vermaat, 2006). Waterbirds are typically wetland-dependent animals upon which they could get abundant food and suitable habitats (Ma *et al.*, 2010). Waterbird abundance and distribution could reflect the status of wetland structure and functions, making them important bio-indicators for wetland health (Fox *et al.*, 2011). Among all factors affecting waterbird community dynamics, food availability is frequently considered to play one of the most important roles (Wang *et al.*, 2013). However, recently suitable food resources have tended to decrease or even disappear due to deterioration and loss of natural wetlands (Fox *et al.*, 2011). As a result, waterbirds are forced to discard previous habitats and sometimes even feed in agricultural lands (Zhang *et al.*, 2011). In addition, migratory waterbirds may aid the dispersal of aquatic plants or invertebrates by carrying and transporting them between water bodies at various spatial scales (Reynolds, Miranda & Cumming, 2015). Consequently, long-time monitoring and systematic studies of waterbird diets are essential to understand population dynamics of waterbirds, as well as to establish effective management programs for them (Wang *et al.*, 2012).

Traditional methods for waterbird diet analysis, were direct observation in the field (Swennen & Yu, 2005) or microhistologic analysis of remnants in feces and/or gut contents (James & Burney, 1997; Fox *et al.*, 2007). While these approaches have been proved useful in some cases, they are relatively labor-intensive and greatly skill-dependent (Fox *et al.*, 2007;

Deleted: s'

Deleted: As

Deleted: traditional method was

Samelius & Alisauskas, 1999; Symondson, 2002). Applications of other methods for analyzing gut contents or feces were also restricted due to inherent limitations, as reviewed by Pompanon *et al.* (Pompanon *et al.*, 2012). Recently, metabarcoding methods, based on high-throughput sequencing, have provided new perspectives for diet analysis and biodiversity assessment (Taberlet *et al.*, 2007; Creer *et al.*, 2010). These methods provide higher taxonomic resolution and enormous sequence output simultaneously from large-scale environmental samples, such as soil, water and feces (Shokralla, Spall & Gibson, 2012; Bohmann *et al.*, 2014). Owing to these advantages, metabarcoding has been widely employed in diet analysis of herbivores (Taberlet *et al.*, 2012; Ando *et al.*, 2013; Hibert *et al.*, 2013), carnivores (Deagle, Kirkwood & Jarman, 2009; Shehzad *et al.*, 2012) and omnivores (De Barba *et al.*, 2014). But pitfalls of metabarcoding should not be ignored when choosing suitable techniques for new studies. For instance, many researches have shown that it is difficult to obtain quantitative data using metabarcoding (Sun *et al.*, 2015). This drawback might result from both technical issues of this method and relevant biological features of samples (Pompanon *et al.*, 2012).

One paramount prerequisite of metabarcoding methods is to select robust genetic markers and corresponding primers (Zhan *et al.*, 2014; Zhan & MacIsaac, 2015). For diet study of herbivores, at least eight chloroplast genes and two nuclear genes are used as potential markers for land plants (Hollingsworth, Graham & Little, 2001). Although mitochondrial cytochrome *c* oxidase I (COI) is extensively recommended as a standard barcode for animals, its relatively low rate of evolution in botanical genomes precludes it being an optimum for plants (Wolfe, Li & Sharp, 1987; Fazekas *et al.*, 2008). The internal transcribed spacer (ITS) is excluded due to

Deleted: metabarcoding methods

Deleted: taxonomy

Comment [2]: higher detectability? higher resolution of diverse items

Deleted: (s)

divergence discrepancies of individuals and low reproducibility (Álvarez & Wendel, 2003). A variety of combinations and comparisons have been performed for the eight candidate genes, however, none proved equally powerful for all cases (Fazekas *et al.*, 2008). Consequently, it is more effective to choose barcodes for a circumscribed set of species occurring in a regional community (Kress *et al.*, 2009). Another equally important aspect of metabarcoding applications is the construction of reference libraries which assist taxonomic assignment (Rayé *et al.*, 2011; Xu *et al.*, 2015). It is difficult to accurately interpret sequence reads without a reliable reference library (Elliott & Jonathan Davies, 2014).

Diet analysis is a central issue in waterbird research, both for deciphering waterfowl population dynamics and interpreting inter- or intra-specific interactions of cohabitating species (Zhao *et al.*, 2015). For instance, more than 60% of bean goose *Anser fabalis* and almost 40% of greater white-fronted goose *Anser albifrons* populations along the East Asian – Australian Flyway Route winter at the Shengjin Lake National Nature Reserve (Zhao *et al.*, 2015). Previous studies based on microhistologic observation illustrated that the dominant composition of their diets were monocotyledons, such as *Carex* spp. (Zhao *et al.*, 2012), *Poaceae* (Zhang *et al.*, 2011), and a relatively small proportion of non-monocots (referred to as dicotyledons in study of "Zhao, Cao & Fox, 2013"). However, few food items could be identified to species-level, mainly owing to variable tissue structures within plants, similar morphology between relative species, and a high level of degradation after digestion (Zhang *et al.*, 2011; Zhao *et al.*, 2012; Zhao, Cao & Fox, 2013). Ambiguous identification has hindered understanding of waterbird population dynamics and potential to establish effective conservation plans for them.

Comment [3]: the hyperbole does not add meaning to the statement

Deleted: ing

Deleted: can dramatically

Deleted:

Deleted: well-recovered

Comment [4]: This seems a dubious statement – I suggest that rather than "the"central issue, diet may be "a" central issue, along with others such as habitat modification and disturbance

Deleted: the

Deleted: there are

Deleted: population

Deleted: ing

Deleted: s

Deleted: abilities

Deleted: in the future

In this study, we aimed to improve this situation using a metabarcoding method to analyze diets of these species (see flowchart in Fig. 1). By examining the efficiency of eight candidate genes (*rbc*L, *rpo*C1, *rpo*B, *mat*K, *trn*H-*psb*A, *trn*L (UAA), *atp*F-*atp*H, and *psb*K-*psb*I), we selected robust genes and corresponding primers for reference library construction and high-throughput sequencing. Subsequently, we used the metabarcoding method to investigate diet composition of these two species based on feces collected from Shengjin Lake. Finally, we discussed and compared results from microhistology and DNA metabarcoding using the same samples to assess the utility and efficiency of these two methods.

Materials and Methods

Ethics Statement

Our research work did not involve capture or any direct manipulation or disturbances of animals. We collected samples of plants and feces for molecular analyses. We got access to the reserve under the permission of Shengjin Lake National Nature Reserve Administration (Chizhou, Anhui, China), which is responsible for the management of the protected area and wildlife. We were forbidden to capture or disturb geese in the field.

Study Area

Shengjin Lake (116º55′ - 117º15′ E, 30º15′ -30º30′ N) was established as National Nature Reserve in 1997, aiming to protect diverse waterbirds including geese, cranes and storks. The water level fluctuates greatly in this lake, with maximal water level of 17 m during summer (flood season) but only 10 m during winter (dry season). Due to this fluctuation, receding waters expose two large *Carex* spp. meadows and provide suitable habitats for waterbirds. This makes

Deleted: tried

Deleted: constructing

Comment [5]: American spelling used above

Deleted: a

Deleted: the

Deleted: There is no need for an ethics statement, because o

Deleted: only

158 Shengjin Lake one of the most important wintering sites for migratory waterbirds (Zhao *et al.*,
159 2015). Greater white-fronted goose and bean goose are the dominant herbivores wintering (from
160 October to April) in this area, accounting for 40% and 60% of populations along the East Asian
161 – Australian Flyway Route, respectively (Zhao *et al.*, 2015).

162 **Field Sampling**

163 The most common plant species that these two geese may consume were collected in May 2014
164 and January 2015, especially species belonging to *Carex* and *Poaceae*. Fresh and intact leaves
165 were carefully picked, tin-packaged in the field and stored at -80 °C in the laboratory before
166 further treatment. Morphological identification was carried out with the assistance of two
167 botanists (Profs Zhenyu Li and Shuren Zhang from Institute of Botany, Chinese Academy of
168 Sciences). In total, 87 specimens were collected, belonging to 25 families, 53 genera and 70
169 species (Table S1).

170 All feces were collected at the reserve (Fig. 2) in January 2015. Based on previous studies
171 and the latest waterbird survey, sites with big flocks of geese (i.e. more than 200 individuals)
172 were chosen (Zhang *et al.*, 2011). As soon as geese finished feeding and feces were defecated,
173 fresh droppings were picked and stored in dry ice. Droppings of bean geese were generally
174 thicker than those of smaller greater white-fronted goose, to the degree that these could be
175 reliably distinguished in the field (Zhao *et al.*, 2015). Disposal gloves were changed for each
176 sample to avoid cross contamination. To avoid repeated sampling and make sure samples were
177 from different individuals, each sample was collected with a separation of more than two meters.
178 In total, 21 feces were collected, including 11 for greater white-fronted goose and 10 for bean

Deleted: next

Deleted: possibly

Comment [6]: The redundancy in this table serves little purpose as it is not linked to voucher or sample numbers. A simple species list with number of samples analysed would suffice (should be in the paper rather than a supplement)

goose. All samples were transported to laboratory in dry ice and then stored at -80 °C until further analysis.

Selection of Molecular Markers and Corresponding Primers

In this part, we aimed to select gene markers with adequate discriminating power for our study. We included eight chloroplast genes - *rbc*L, *rpo*C1, *rpo*B, *mat*K, *trn*H-*psb*A, *trn*L (UAA), *atp*F-*atp*H, and *psb*K-*psb*I for estimation. Although Shengjin Lake included an array of plant species, we focused mainly on the most likely food resources (Xue *et al.*, 2008; Zhao *et al.*, 2015) that geese would consume for candidate gene tests. These covered eleven genera and the family Poaceae (Table S2; Table 2). For tests of all candidate genes, we recovered sequences of representative species in the selected groups from GenBank (http://www.ncbi.nlm.nih.gov/nuccore). We calculated inter-specific divergence within every genus or family based on the Kimura 2-parameter model (K2P) using MEGA version 6 (Tamura *et al.*, 2013). We also constructed molecular trees based on UPGMA using MEGA and characterized the resolution of species by calculating the percentage of species recovered as monophyletic based on phylogenetic trees (Rf?). Secondly, primers selected out of eight of candidate genes were used to amplify all 87 specimens and to check their amplification efficiency and universality. Thirdly, we calculated inter-specific divergence based on sequences that we obtained from last step. Generally, a robust barcode gene is obtained when the minimal inter-specific distance exceeds the maximal intra-specific distance (e.g. existence of barcoding gaps). For reference database building, we calculated the rate of discrimination for the species in each family (Rf) by dividing the number of unique sequences per family by the number of

Deleted: consisted

Deleted: potential

Deleted: one family

Comment [7]: Needs clarification: were all sequences for a gene from different species, or were some distinct alleles within species?

Deleted: Kiruma

Comment [8]: In Table S2, give the number of species compared along with the number of sequences (so that the reader has some idea of variation in resolution of your resolvability measure)

Deleted: selected out of eight

Deleted: recommended

species resolved as monophyletic clades in each family. Finally, to allow the recognition of sequences after high-throughput sequencing, both of the forward and reverse primers of the selected marker gene were tagged specifically for each sample with 8nt nucleotide codes at the 5' end (Parameswaran *et al.*, 2007).

DNA Extraction, Amplification and Sequencing

Two hundred milligrams of leaf was used to extract the total DNA from each plant sample using a modified CTAB protocol (Cota-Sanchez, Remarchuk & Ubayasena, 2006). DNA extraction of feces was carried out using the same protocol with minor modification in incubation time (elongate to 12 h). Each fecal sample was crushed thoroughly and divided into four quarters. All quarters of DNA extracts were then pooled together. DNA extraction was carried out in a clean room used particularly for this study. For each batch of DNA extraction, negative controls (i.e. extraction without feces) were included to monitor possible contamination.

For plant DNA extracts, PCR amplifications were carried out in a volume of 25µl with ~100 ng total DNA as template, 1U of *Taq* Polymerase (Takara, Dalian, Liaoning Prov., China), 1× PCR buffer, 2 mM of Mg^{2+}, 0.25 mM of dNTPs, 0.1 µM of forward primer and 0.1 µM of reverse primer. After 4 min at 94 °C, the PCR cycles were as follows: 35 cycles of 30 s at 94 °C, 30 s at 56 °C and 45 s at 72 °C, and the final extension was 10 min at 72 °C. We applied the same PCR conditions for all primers. All the successful PCR products were sequenced with Genewiz (Suzhou, Jiangsu Prov., China).

For fecal DNA extracts, PCR mixtures (25µl) were prepared in six replicates for each sample to reduce biased amplification. Each replicate was subjected to the same amplification

Comment [9]: This is a very poor measure of discrimination and not comparable across groups because it depends on unspecified sampling details (how many sampled species per group, how many individuals sampled, phylogenetic diversity among sampled species) which vary across groups. Part of the issue is that it is reported as a percentage (ratio, where both numerator and denominator vary informatively)

Deleted: At last

Deleted: in

Comment [10]: duplication

Deleted: . Each replicate consisted of ~100 ng total DNA of DNA extract as template, 1U of *Taq* Polymerase (Takara, Dalian, Liaoning Prov., China), 1× buffer, 2 mM of $MgCl_2$, 0.25 mM of each dNTPs, 0.1 µM of foward primer and 0.1 µM of reverse primer. After 4 min of denaturation at 94 °C, the PCR cycles were as follows: 35 cycles of 30 s at 94 °C, 30 s at 56 °C, 45 s at 72 °C, and the final elongation was 10 min at 72 °C

procedure used for plant extracts. Each set of six replicates was pooled and purified using the Sangon PCR product purification kit (Sangon Biotech, Shanghai, China). Quantification was carried out to ensure equilibrium of contribution of each sample using the NanoDrop ND-2000 UV-Vis Spectrophotometer (NanoDrop Technologies, Delaware, United States of America). High-throughput sequencing was performed using Illumina MiSeq platform following manufacturer's instructions by BGI (Shenzhen, Guangdong Prov., China). Reads of high-throughput sequencing could be found at NCBI's Sequence Read Archive (Accession number: SRP070470).

Data Analysis for Estimating Diet Composition

After high-throughput sequencing, pair-ended reads were merged with using the UPARSE pipeline (http://drive5.com/usearch, Edgar, 2010). Reads were then split into independent files according to unique tags using RDP pipeline (http://rdp.cme.msu.edu/). We removed sequences i) that didn't perfectly match tags and primer sequences; ii) that contained ambiguous nucleotide (N's). Tags and primers were then trimmed using the RDP pipeline. Further quality filtering based on the UPARSE pipeline discarded sequences with i) quality score less than 30 (<Q30) and ii) shorter than 100 bp and longer than 200 bp. Unique sequences were clustered to operational taxonomy units (OTUs) at the similarity threshold of 98% (Edgar, 2013). All OTUs were assigned to unique taxonomy with local blast 2.2.30+ (Altschul *et al.*, 1990). We detected a plant within the reference library for each sequence with the threshold of length coverage > 98%, identity > 98% and e-value < 1.0 e^{-50}. If a query sequence matched two or more taxa, it was assigned to a higher taxonomic level which included all taxa.

Deleted: with

Comment [11]: The RDP pipeline (correct URL https://pyro.cme.msu.edu/) is a set of tools and workflows mainly for rDNA libraries (particularly 16S, not the genes used here – although for trimiming ends and compiling sequences this is not a problem); UPARSE is an algorithm for identification of OTUs (correct web site http://drive5.com/uparse). Specific details of workflows and filters need to be stated for reproducibility)

Microhistology analysis

We used the method described by Zhang *et al.* (2011) to perform microhistologic examination of fecal samples. Each sample was first washed with pure water and filtered with a 25-μm filter. Subsequently, the suspension was examined under a light microscope at 10× magnification for quantification statistics and at 40× magnification for species identification. We compared photos of visible fragments with an epidermis database of plants from Shengjin Lake to identify food items (Source?...Zhang? Fox?).

Results

Selection of Genes and Corresponding Primers and Reference Library Constructing

A total of 3,296 representative sequences were recovered from GenBank, ranging from 0 to 345 sequences per gene, per genus (Table S2). Among the eight candidate genes, *trn*L, *trn*H-*psb*A, *mat*K and *rbc*L showed largest inter-specific divergence in seven, three, one and one taxonomic groups, respectively. These four genes also displayed relatively high resolution of species (Table S2). For example, with *mat*K gene, 77% of *Carex* could be identified to species-level. However, our results indicated that none of these eight genes could simultaneously differentiate all 12 genera or families to species-level (Table S2). Considering the inter-specific divergence and resolution of species, we chose the most commonly used chloroplast genes *rbc*L, *mat*K, *trn*H-*psb*A and *trn*L for further tests.

Primers for these four genes (Table 1) were used to amplify the plants that we collected in the field. The selected primers for *trn*L and *rbc*L successfully amplified 100% and 91% of all species, respectively, while primers for *trn*H-*psb*A and *mat*K amplified only 71% and 43%,

Deleted: of the

Deleted: and the sequence numbers per gene ranged

Deleted: family

Deleted: re

Comment [12]: There seems to be a discrepancy here – the data in S2 do not support this statement (please clarify if I am misinterpreting this).
Highest mean divergence in S2:
*trn*H-*psb*A = 4 (Artemisia, Elymus, Juncus, Polygonum);
*atp*F-*atp*H = 4 (Carex, Potamogeton, Ranunculus, Rumex);
*trn*L = 3 (Echinochloa, Eleocharis, Trapa);
*psb*K-*psb*I =1 (Poaceae).

Based on maximum divergence in S2:
*trn*L = 6 (Elymus, Carex, Potamogeton, Rumex, Eleocharis, Trapa);
*trn*H-*psb*A = 4 (Artemisia, Juncus, Polygonum, Ranunculus);
*psb*K-*psb*I 1 (Poaceae);
*rbc*L =1 (Echinochloa)

Additionally there is wide variation in the number of species compared for each gene, and many cases where no comparison was possible (Eleocharis and Trapa are not fair tests of resolvability among genes). I suggest that these data are retained as is in S2 but that the mean divergence among species within groups is combined across taxa for comparing genes. Other important considerations are the abundance of database sequences for comparison and consistent amplification – an argument that supports the chosen used genes. On any other argument rbcL and matK would be a poor choice.

Deleted: of

respectively. Therefore, we chose *trn*L and *rbc*L to test their discriminating power in our target plants.

We calculated the inter-specific divergence within genera and families with at least two species to compare their discriminating power. Maximal, minimal and mean inter-specific distances were calculated for seven dominant genera and six dominant families (Table 2). Neither gene could differentiate species of *Vallisneria Linn.* (mean=0.000±0.000%) or *Artemisia Linn.* (mean=0.000±0.000%). But *trnL* showed a larger divergence range for the other six genera and five families. Hence, we chose *trn*L as the barcoding gene for reference library constructing and high-throughput sequencing for our study. The discriminating power of *trn*L was strong for most species (Table 3). However, some species could only be identified at genus-level or family-level with *trn*L. For instance, five species of *Potamogeton* shared the same sequences and this made them to be identified at genus-level. Species could be identified easily to genus and family, except for three grasses (Poaceae) *Beckmannia syzigachne, Phalaris arundinacea* and *Polypogon fugax*, which shared identical sequences.

Data Processing for Estimating Diet Composition

In total, 0.21 and 0.18 million reads were generated for greater white-fronted goose (GWFG) and bean goose (BG), respectively (Table 4). The number of recovered OTUs ranged from 8 to 123 for GWFG and BG samples. We used local BLAST to compare these sequences with the Shengjin Lake reference database. Finally, with DNA metabarocoding, 12 items were discovered in the feces of GWFG, including one at family-level, three at genus-level and eight at species-level (Table 5). Four items were discovered in the feces of BG, including one at

Deleted: Both

Deleted: s

Deleted: not

Deleted: and

Deleted: *aceae*

Comment [13]: These are all species of Potamogeton and therefore identified to genus level

Deleted: family

Deleted: for genera

Deleted: or

Deleted: containing one species,

Deleted: *L.*

Deleted:

Deleted: a

genus-level and three at species-level. In total, this method identified 15 taxa in feces of these geese.

However, the sequence percentage of each food item varied greatly (Table 5). For GWFG, the majority of sequences (96.36%) were composed of only five items - *Poaceae* spp. (47.98%), *Poa annua* (21.86%), *Carex heterolepis* (17.51%), *Carex* spp. (9.01%), and *Alopecurus aequalis* (3.21%). For BG, almost all the sequences belonged to *Carex heterolepis* (99.49%). Other items only occupied a relatively small proportion of sequences. In addition, the presence of each item per sample was also unequal (Table S3). For example in GWFG, *Carex heterolepis*, *Carex* spp., *Poa annua* and *Potentilla supina* were present in almost all the samples, while *Stellaria media*, *Asteraceae* sp. and *Lapsana apogonoldes* occurred in only about one third of samples.

When microhistologic examination were performed using the same samples, eight items were found in the feces of greater white-fronted goose, including one at family-level, four at genus-level and three at species-level (Table 5). Dominant items were *Poaceae* spp. (45.68%), *Alopecurus Linn.* (30.93%) and *Carex heterolepis* (16.39%). Seven items were found in the feces of bean goose, including four at genus-level and three at species-level (Table 5). Dominant items were *Carex heterolepis* (62.85%), *Asteraceae* sp. (14.55%), and *Alopecurus Linn.* (13.18%).

Discussion

Marker Selection and Reference Library Constructing for Diet Analysis

With greatly reduced cost, extremely high throughput and information content, metabarcoding has revolutionized the exploration and quantification of dietary analysis from noninvasive samples containing degraded DNA (Fonseca *et al.*, 2010; Shokralla *et al.*, 2014). Despite

Deleted: T

Comment [14]: Which of the sampled species from Shengjin could not be distinguished. Does Poaceae spp. Mean NOT Poa annua, Carex spp. = NOT c. heterolepsis? In which case, what other species could be excluded (or could it rather be stated as a group of Carex).

Comment [15]: The attribution of authors to species is inconsistent and unnecessary

Deleted: *(L.) Cyr.*

Deleted: of greater white-fronted goose

Deleted:

Deleted: explosive

Deleted: with

enormous potential to boost data acquisition, successful application of this technology relies greatly on the power and efficiency of genetic markers and corresponding primers (Bik *et al.*, 2012; Zhan *et al.*, 2014). In order to select the most appropriate marker gene for our study, we compared the performance of eight commonly used chloroplast genes, *rbc*L, *rpo*B, *rpo*C1, *mat*K, *trn*L, *trn*H-*psb*A, *atp*F-*atp*H, and *psb*K-*psb*I and their corresponding primers. Although a higher level of discriminating power was shown in several studies, *atp*F-*atp*H, *psb*K-*psb*I, *rpo*B and *rpo*C1 were not as commonly used as other barcoding genes (Hollingsworth, Graham & Little, 2001). As one of the most rapidly evolving coding genes of plastid genomes, *mat*K was considered as the closest plant analogue to the animal barcode *COI* (Hilu & Liang, 1997). However, *mat*K was difficult to amplify using available primer sets, with only 43% of successful amplification in this study. In spite of the higher species discrimination success of *trn*H-*psb*A than *rbc*L+*mat*K in some groups, the presence of duplicated loci, microinversions and premature termination of reads by mononucleotide repeats lead to considerable proportion (30% in this study) of low-quality sequences and over-estimation of genetic difference when using *trn*H-*psb*A (Graham *et al.*, 2000; Whitlock Hale & Groff, 2010). In contrast, the barcode region of *rbc*L is easy to amplify, sequence, and align in most plants and was recommended as the standard barcode for land plants (Chase *et al.*, 2007). The relatively modest discriminating power (compared to *trn*L) precludes its application for our study aiming to recover high resolution of food items. Consequently, *trn*L was selected out of eight candidate markers, with 100% amplification success, more than 90% of high quality sequences, and relatively large inter-specific divergence.

One of the biggest obstacles in biodiversity assessment and dietary analysis is the lack of a comprehensive reference library, without which it is impossible to accurately interpret and assign sequences generated from high-throughput sequencing (Valentini, Pompanon & Taberlet 2009; Barco *et al.*, 2015). In this study, we constructed a local reference library by amplifying the most common species (70 morpho-species in total) during the wintering period with the *trn*L gene. Although not all of them could be identified at species-level with *trn*L due to relatively low inter-specific divergence, many species could be separated with distinctive sequences. Previous studies have recommended group-specific barcodes to differentiate closely related plants at the species level (Li *et al.*, 2015). For instance, *mat*K has been proved to be more efficient for the discrimination of *Carex* spp. (Starr, Naczi & Chouinard, 2009). However, the primer set of *mat*K failed to amplify species of *Carex* spp. in our study, suggesting the universality of selected primer pairs should be tested in each study (Zhan *et al.*, 2014).

Applications of Metabarcoding for Geese Diet Analysis

A variety of recent studies have demonstrated the great potential of metabarcoding for dietary analysis, mainly owing to the high throughput, high discriminating power, and the ability to process large-scale samples simultaneously (Creer *et al.*, 2010; Taberlet *et al.*, 2012; Shehzad *et al.*, 2012). In this study, we applied this method to recover diets of herbivorous geese and provided standard protocols for dietary analysis of these two ecologically important waterbirds. Our results further proved the more objective, less experience-dependent and more time-efficient character of DNA metabarcoding. However, not all the species in the reference library could be identified at species-level, owing to low inter-specific divergence. We suggest that multiple

Deleted: the enormous

Deleted: ed

group-specific markers to be incorporated in the future, as in De Barba *et al.*(2014). Two species, *Carex thunbergii* and *Fabaceae* sp., were only discovered via microhistologic analysis rather than metabarcoding. This failure might reflect the biased fragment amplification of current technology, of which dominant templates could act as inhibitors of less dominant species (Piñol *et al.*, 2015). However, three species of *Poaceae* were only discovered using metabarcoding. In total, more taxa and higher resolution were attained using metabarcoding. But microhistology still proved a powerful supplementary. Previous studies using metabarcoding usually detected dozens of food items, even as many as more than one hundred species. For instance, 18 taxa prey were identified for leopard cat (Shehzad *et al.*, 2012); 44 plant taxa were recovered in feces of red-headed wood pigeon (Ando *et al.*, 2013); while more than 100 taxa were found in diet studies of brown bear (De Barba *et al.*, 2014). The relatively narrow diet spectrum of herbivorous geese may lead to misunderstanding that this result of our study is merely an artefact due to small sampling effort. However, this result is credible since these two geese species only feed on *Carex* meadow, where the dominant vegetation is *Carex* spp., with other species such as *Poaceae* and dicots (Zhao *et al.*, 2015). Even though other wetland plants exist, they usually composed only a small proportion of the geese diets.

Quantification of food composition is another key concern in dietary analysis. Although the relative percentage of sequences were not truly a quantitative estimate of diet, taxa of the majority sequences in this study were in accord with microhistologic observations, which was considered an efficient way to provide quantitative results (Wang *et al.*, 2013, Discrepancies might come from the semi-quantitative nature of metabarcoding methods (Sun *et al.*, 2015). This

Deleted: such as

Deleted: studies of

Deleted: Besides

Deleted: proliferation

Deleted: a

Deleted: granted as

Deleted:), while discrepancy also existed

Deleted: This

is likely derived from PCR amplification, which always entails biases caused by universal primer-template mismatches, annealing temperature or number of PCR cycles (Zhan *et al.*, 2014; Piñol *et al.*, 2015). Other methods such as shot-gun sequencing or metagenomic sequencing could be incorporated in the future to give information on abundance of food items (Srivathsan *et al.*, 2015).

Implications for Waterbird Conservation and Wetland Management

For long-distance migratory waterbirds, such as the wild geese in this study, their abundance and distribution are greatly influenced by diet availability and habitat use (Wang *et al.*, 2013). For example, waterbirds may be restricted at (forced to leave) certain areas due to favoring (loss) of particular food (Wang *et al.*, 2013), while the recovery of such food may contribute to return of bird populations (Noordhuis *et al.*, 2002). Results of both metabarcoding and microhistologic analysis in this study revealed that *Carex* and *Poaceae* were dominant food components which is in accordance with previous studies. The increasing number of these two geese wintering at the Shengjin Lake may be attributed to the expansion of *Carex* meadow, which offers access to abundant food resources (Zhao *et al.*, 2015). Considering the long-distance migratory character of these birds, it is important to maintain energy balances and good body conditions in wintering areas because this might further influence their departure dates and reproductive success after arriving at breeding areas (Prop, Black & Shimmings, 2003). Based on this, it is important for wetland managers to maintain suitable habitats and food resources for sustainable conservation of waterbirds, which highlights the significance of diet information. Our study also indicated that overlap and dissimilarity existed between the diets of these two geese. As we all know, animals

Deleted: bring

Deleted: integrative

Deleted: were

Deleted: s

Deleted: accordant

Deleted: Besides, considering

Deleted: the

Deleted: to perform

Deleted: and all of

Deleted: highlighted

foraging in the same habitats may compete for limited food resources (Madsen & Mortensen, 1987). This discrepancy of food composition may arise from the avoidance of inter-specific competition (Zhao *et al.*, 2015). However, with the increase of these two species in Shengjin Lake, further research is needed to investigate the mechanisms of food resource partitioning and spatial distribution.

Shengjin Lake is one of the most important wintering sites for tens of thousands of migratory watebirds, while annual life cycles of these birds depend on the whole migratory route, including breeding sites, stop-over sites and wintering sites (Kear, 2006). Thus, a molecular reference library covering all the potential food items along the whole migratory route will be useful both for understanding of wetland connections and waterbird conservation. Besides, the ability of DNA metabarcoding to process lots of samples simultaneously enables rapid analyses and makes this method helpful for waterbird studies.

Acknowledgements

We are very grateful to the stuff of the Shengjin Lake National Nature Reserve for their excellent assistance during the field work. Great thanks to Zhujun Wang and An An for feces collection in the field. We thank Song Yang for collecting plants in the Shengjin Lake Reserve. We also thank Profs. Zhenyu Li and Shuren Zhang for plant identification. Special thanks to Drs. Meijuan Zhao, Xin Wang, Fanjuan Meng and Peihao Cong for preparing the epidermis database and guiding microhistologic analysis.

Deleted: es are

Deleted: s

Deleted: the quick

Deleted: treatment

Deleted: a

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology* **215:**403-410 DOI 10.1006/jmbi.1990.9999.

Álvarez I, Wendel JF. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* **29:**417-434 DOI 10.1016/S1055-7903(03)00208-2.

Ando H, Setsuko S, Horikoshi K, Suzuki H, Umehara S, Inoue-Murayama M, Isagi Y. 2013. Diet analysis by next‐generation sequencing indicates the frequent consumption of introduced plants by the critically endangered red‐headed wood pigeon (*Columba janthina nitens*) in oceanic island habitats. *Ecology and evolution* **3:**4057-4069 DOI 10.1002/ece3.773.

Barco A, Raupach MJ, Laakmann S, Neumann H, Knebelsberger T. 2015. Identification of North Sea molluscs with DNA barcoding. *Molecular Ecology Resources* **16:**288-297 DOI 10.1111/1755-0998.12440.

Bik HM, Porazinska DL, Creer S, Caporaso JG, Knight R, Thomas WK. 2012. Sequencing our way towards understanding global eukaryotic biodiversity. *Trends in Ecology and Evolution* **27:**233-243 DOI 10.1016/j.tree.2011.11.010.

Bohmann K, Evans A, Gilbert MTP, Carvalho GR, Creer S, Knapp M, Yu WD, de Bruyn M. 2014. Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology and Evolution* 29:358-367 DOI 10.1016/j.tree.2014.04.003.

Brander LM, Florax, RJ, Vermaat JE. 2006. The empirics of wetland valuation: a comprehensive summary and a meta-analysis of the literature. *Environmental and Resource Economics* **33:**223-250. DOI 10.1007/s10640-005-3104-4.

Chase MW, Cowan RS, Hollingsworth PM, van den Berg C, Madriñán S, Petersen G, Seberg O, Jorgsensen T, Cameron KM, Carine M, Pedersen N, Hedderson TAJ, Conrad F, Salazar GA, Richardson JE, Hollingsworth M, Barraclough TG, Kelly L, Wilkinson M. 2007. A proposal for a standardised protocol to barcode all land plants. *Taxon* **56:**295-299.

Cota-Sanchez JH, Remarchuk K, Ubayasena K. 2006. Ready-to-use DNA extracted with a CTAB method adapted for herbarium specimens and mucilaginous plant tissue. Plant *Molecular Biology Reporter* **24:**161-167 DOI 10.1007/BF02914055.

Creer S, Fonseca VG, Porazinska DL, Giblin-Davis RM, Sung W, Power DM, Packer M, Carvalho GR, Blaxter ML, Lambshead PJD, Thomas WK. 2010. Ultrasequencing of the meiofaunal biosphere: practice, pitfalls and promises. *Molecular Ecology* **19:**4-20 DOI 10.1111/j.1365-294X.2009.04473.x.

De Barba M, Miquel C, Boyer F, Mercier C, Rioux D, Coissac E, Taberlet P. 2014. DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: application to omnivorous diet. *Molecular Ecology Resources* **14:**306-323 DOI 10.1111/1755-0998.12188.

Deagle BE, Kirkwood R, Jarman SN. 2009. Analysis of Australian fur seal diet by pyrosequencing prey DNA in faeces. *Molecular Ecology* **18:**2022-2038 DOI 10.1111/j.1365-294X.2009.04158.x.

Dunning LT, Savolainen V. 2010. Broad-scale amplification of matK for DNA barcoding plants, a technical note. *Botanical Journal of the Linnean Society*, **164:**1-9.

Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26:**2460-2461 DOI 10.1093/bioinformatics/btq461.

Edgar RC. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods*

10:996-998 DOI 10.1038/NMETH.2604.

Elliott TL, Jonathan Davies T. 2014. Challenges to barcoding an entire flora. *Molecular Ecology Resources* **14:**883-891 DOI 10.1111/1755-0998.12277.

Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, Newmaster SG, Husband BC, Percy DM, Hajibabaei M, Barrett SC. 2008. Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS One* **3:**e2802 DOI 10.1371/journal.pone.0002802.

Fonseca VG, Carvalho GR, Sung W, Johnson HF, Power DM, Neill SP, Packer M, Blaxter ML, Labmshead PJD, Thomas WK, Creer S. 2010. Second-generation environmental sequencing unmasks marine metazoan biodiversity. *Nature communications* **1:**98 DOI 10.1038/ncomms1095.

Ford CS, Ayres KL, Haider N, Toomey N, van-Alpen-Stohl J. 2009. Selection of candidate DNA barcoding regions for use on land plants. *Botanical Journal of the Linnean Society* **159:**1-11 DOI 10.1111/j.1095-8339.2008.00938.x.

Fox AD, Bergersen E, Tombre IM, Madsen J. 2007. Minimal intra-seasonal dietary overlap of barnacle and pink-footed geese on their breeding grounds in Svalbard. *Polar Biology* **30:**759-768 DOI 10.1007/s00300-006-0235-1.

Fox AD, Cao L, Zhang Y, Barter M, Zhao M, Meng F, Wang S. 2011. Declines in the tuber−feeding waterbird guild at Shengjin Lake National Nature Reserve, China–a barometer of submerged macrophyte collapse. *Aquatic Conservation: Marine and Freshwater Ecosystems* **21:**82-91 DOI 10.1002/aqc.1154.

Graham SW, Reeves PA, Burns AC, Olmstead RG. 2000. Microstructural changes in noncoding chloroplast DNA: interpretation, evolution, and utility of indels and inversions in basal angiosperm phylogenetic inference. *International Journal of Plant Sciences* **161:**S83-S96 DOI 10.1086/317583.

Hibert F, Taberlet P, Chave J, Scotti-Saintagne C, Sabatier D, Richard-Hansen C. 2013. Unveiling the diet of elusive rainforest herbivores in next generation sequencing era? The tapir as a case study. *PloS One* **8:**e60799 DOI 10.1371/journal.pone.0060799.

Hilu KW, Liang H. 1997. The *mat*K gene: sequence variation and application in plant systematics. *American Journal of Botany* **84:**830-839 DOI 10.2307/2445819.

Hollingsworth PM, Graham SW, Little DP. 2011. Choosing and using a plant DNA barcode. *PloS One* **6:**e19254 DOI 10.1371/journal.pone.0019254.

James HF, Burney DA. 1997. The diet and ecology of Hawaii's extinct flightless waterfowl: evidence from coprolites. *Biological Journal of the Linnean Society* **62:**279-297 DOI 10.1111/j.1095-8312.1997.tb01627.x.

Kear J. 2005. Ducks, geese and swans. Oxford: Oxford University Press.

Kress WJ, Erickson DL, Jones FA, Swenson NG, Perez R, Sanjur O, Bermingham E. 2009. Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. *Proceedings of the National Academy of Sciences of the United States of America* **106:**18621-18626 DOI 10.1073/pnas.0909820106.

Li X, Yang Y, Henry RJ, Rossetto M, Wang Y, Chen S. 2015. Plant DNA barcoding: from gene to genome. *Biological Reviews* **90:**157-166 DOI 10.1111/brv.12104.

Ma Z, Cai Y, Li B, Chen J. 2010. Managing wetland habitats for waterbirds: an international perspective. *Wetlands* **30:**15-27 DOI 10.1007/s13157-009-0001-6.

Madsen J, Mortensen CE. 1987. Habitat exploitation and interspecific competition of moulting geese in East

561 Greenland. *Ibis* **129:**25-44 DOI 10.1111/j.1474-919X.1987.tb03157.x.
562 **Noordhuis R, van der Molen DT, van den Berg MS. 2002.** Response of herbivorous water-birds to the return
563 of *Chara* in Lake Veluwemeer, The Netherlands. *Aquatic Botany* **72:**349-367 DOI
564 10.1016/S0304-3770(01)00210-8.
565 **Parameswaran P, Jalili R, Tao L, Shokralla S, Gharizadeh B, Ronaghi M, Fire AZ. 2007.** A
566 pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample
567 multiplexing. *Nucleic Acids Research* **35:**e130 DOI 10.1093/nar/gkm760.
568 **Piñol J, Mir G, Gomez-Polo P, Agustí N. 2015.** Universal and blocking primer mismatches limit the use of
569 high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Molecular Ecology*
570 *Resources* **15:**819-830 DOI 10.1111/1755-0998.12355.
571 **Pompanon F, Deagle BE, Symondson WO, Brown DS, Jarman SN, Taberlet P. 2012.** Who is eating what:
572 diet assessment using next generation sequencing. *Molecular Ecology* **21:**1931-1950 DOI
573 10.1111/j.1365-294X.2011.05403.x.
574 **Prop J, Black JM, Shimmings P. 2003.** Travel schedules to the high arctic: barnacle geese trade‐off the
575 timing of migration with accumulation of fat deposits. *Oikos* **103:**403-414 DOI
576 10.1034/j.1600-0706.2003.12042.x.
577 **Rayé G, Miquel C, Coissac E, Redjadj C, Loison A, Taberlet P. 2011.** New insights on diet variability
578 revealed by DNA barcoding and high-throughput sequencing: chamois diet in autumn as a case study.
579 *Ecological Research* **26:**265-276 DOI 10.1007/s11284-010-0780-5.
580 **Reynolds C, Miranda NA, Cumming GS. 2015.** The role of waterbirds in the dispersal of aquatic alien and
581 invasive species. *Diversity and Distribution* 21:744-754 DOI 10.1111/ddi.12334.
582 **Samelius G, Alisauskas RT. 1999.** Diet and growth of glaucous gulls at a large Arctic goose colony.
583 *Canadian Journal of Zoology* **77:**1327-1331 DOI 10.1139/z99-091.
584 **Sang T, Crawford DJ, Stuessy TF. 1997.** Chloroplast DNA phylogeny, reticulate evolution, and
585 biogeography of *Paeonia* (*Paeoniaceae*). *American Journal of Botany* **84:**1120-1136.
586 **Shehzad W, Riaz T, Nawaz MA, Miquel C, Poillot C, Shah SA, Pompanon F, Coissac E, Taberlet P. 2012.**
587 Carnivore diet analysis based on next‐generation sequencing: application to the leopard cat (*Prionailurus*
588 *bengalensis*) in Pakistan. *Molecular Ecology* **21:**1951-1965 DOI 10.1111/j.1365-294X.2011.05424.x.
589 **Shokralla S, Spall JL, Gibson JF. 2012.** Next-generation sequencing technologies for environmental DNA
590 research. *Molecular Ecology* 21:1794-1805 DOI 10.1111/j.1365-294X.2012.05538.x.
591 **Srivathsan A, Sha J, Vogler AP, Meier R. 2015.** Comparing the effectiveness of metagenomics and
592 metabarcoding for diet analysis of a leaf-feeding monkey (*Pygathrix nemaeus*). *Molecular Ecology*
593 *Resources* **15:**250-261 DOI 10.1111/1755-0998.12302.
594 **Starr JR, Naczi RFC, Chouinard BN. 2009.** Plant DNA barcodes and species resolution in sedge (Carex,
595 *Cyperaceae*). *Molecular Ecology Resources* **9:**151-163 DOI 10.1111/j.1755-0998.2009.02640.x.
596 **Sun C, Zhao Y, Li H, Dong Y, MacIsaac HJ, Zhan A. 2015.** Unreliable quantification of species abundance
597 based on high-throughput sequencing data of zooplankton communities. Aquatic Biology **24:**9-15 DOI
598 10.3354/ab00629.
599 **Swennen C, Yu YT. 2005.** Food and feeding behavior of the Black-faced Spoonbill. *Waterbirds* **28:**19-27
600 DOI 10.1675/1524-4695(2005)028[0019:FAFBOT]2.0.CO;2.
601 **Symondson WOC. 2002.** Molecular identification of prey in predator diets. *Molecular Ecology* **11:**627-641

602 DOI 10.1046/j.1365-294X.2002.01471.x.
603 **Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E. 2012.** Towards next-generation
604 biodiversity assessment using DNA metabarcoding. *Molecular Ecology* **21:**2045-2050 DOI
605 10.1111/j.1365-294X.2012.05470.x.
606 **Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, Vermat T, Corthier G, Brocmann**
607 **C, Willerslev E. 2007.** Power and limitations of the chloroplast *trn*L (UAA) intron for plant DNA
608 barcoding. *Nucleic Acids Research* **35:**e14 DOI 10.1093/nar/gkl938.
609 **Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013.** MEGA6: molecular evolutionary genetics
610 analysis version 6.0. *Molecular Biology and Evolution* **30:**2725-2729 DOI 10.1093/molbev/mst197.
611 **Tate JA, Simpson BB. 2003.** Paraphyly of Tarasa (Malvaceae) and diverse origins of the polyploidy species.
612 *Systematic Botany* **28:**723-737.
613 **Valentini A, Pompanon F, Taberlet P. 2009.** DNA barcoding for ecologists. *Trends in Ecology and Evolution*
614 **24:**110-117 DOI 10.1016/j.tree.2008.09.011.
615 **Wang X, Fox AD, Cong P, Barter M, Cao L. 2012.** Changes in the distribution and abundance of wintering
616 Lesser White-fronted Geese *Anser erythropus* in eastern China. *Bird Conservation International*
617 **22:**128-134 DOI 10.1017/S095927091100030X.
618 **Wang X, Fox AD, Cong P, Cao L. 2013.** Food constraints explain the restricted distribution of wintering
619 Lesser White-fronted Geese *Anser erythropus* in China. *Ibis* **155:**576-592 DOI 10.1111/ibi.12039.
620 **Whitlock BA, Hale AM, Groff PA. 2010.** Intraspecific inversions pose a challenge for the *trn*H-*psb*A plant
621 DNA barcode. *PLoS One* **5:**e11533 DOI 10.1371/journal.pone.0011533.
622 **Wolfe KH, Li WH, Sharp PM. 1987.** Rates of nucleotide substitution vary greatly among plant mitochondrial,
623 chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences of the United States of*
624 *America* **84:**9054-9058 DOI 10.1073/pnas.84.24.9054.
625 **Xu C, Dong W, Shi S, Cheng T, Li C, Liu Y, Wu P, Wu H, Gao P, Zhou S. 2015.** Accelerating plant DNA
626 barcode reference library construction using herbarium specimens: improved experimental techniques.
627 *Molecular Ecology Resources* **15:**1366-1374 DOI 10.1111/1755-0998.12413.
628 **Xu L, Xu W, Sun Q, Zhou Z, Shen J, Zhao X. 2008.** Flora and vegetation in Shengjin Lake. *Journal of*
629 *Wuhan Botanical Research* **27:**264-270.
630 **Zhan A, Bailey SA, Heath DD, Macisaac HJ. 2014.** Performance comparison of genetic markers for
631 high-throughput sequencing-based biodiversity assessment in complex communities. *Molecular Ecology*
632 *Resources* **14:**1049-1059 DOI 10.1111/1755-0998.12254.
633 **Zhan A, MacIsaac HJ. 2015.** Rare biosphere exploration using high-throughput sequencing: research progress
634 and perspectives. *Conservation Genetics* **16:**513-522 DOI 10.1007/s10592-014-0678-9.
635 **Zhang Y, Cao L, Barter M, Fox AD, Zhao M, Meng F, Shi H. 2011.** Changing distribution and abundance
636 of Swan Goose *Anser cygnoides* in the Yangtze River floodplain: the likely loss of a very important
637 wintering site. *Bird Conservation International* **21:**36-48 DOI 10.1017/S0959270910000201.
638 **Zhao M, Cao L, Fox AD. 2013.** Distribution and diet of wintering Tundra Bean Geese *Anser fabalis*
639 *serrirostris* at Shengjin Lake, Yangtze River floodplain, China. *Wildfow* **60:**52-63.
640 **Zhao M, Cao L, Klaassen M, Zhang Y, Fox AD. 2015.** Avoiding competition? Site use, diet and foraging
641 behaviours in two similarly sized geese wintering in China. *Ardea* **103:**27-38 DOI
642 10.5253/arde.v103i1.a3.

643 **Zhao M, Cong P, Barter M, Fox AD, Cao L. 2012.** The changing abundance and distribution of Greater
644 White-fronted Geese *Anser albifrons* in the Yangtze River floodplain: impacts of recent hydrological
645 changes. *Bird Conservation International* **22:**135-143 DOI 10.1017/S0959270911000542.