Low-copy nuclear markers in Isoëtes (Isoëtaceae) identified with transcriptomes

Peter W. Schafran1,2,3*, Gabriel Johnson2, W. Carl Taylor2, Elizabeth A. Zimmer2, and Lytton J. Musselman1

Isoëtes L. (Isoëtaceae, Lycopodiophyta) is a cosmopolitan genus of ca. 250 recognized species. These heterosporous lycophytes consist of a 2–3-lobed rootstock that bears linear, quill-like, microphyllous leaves or sporophylls. All microphylls have the potential to develop into sporophylls (Foster and Gifford, 1974). Mega- and microsporangia are produced at the base of sporophylls, in some species covered by a layer of tissue called a velum. Traditionally, spore ornamentation and velum coverage have been considered taxonomically important. Although species inhabit a variety of ecological niches, from obligate aquatic to ephemeral terrestrial habitats, their morphology is extremely conserved. Phylogenetic studies in closely related clades of Isoëtes have been limited by a dearth of morphological features and molecular markers. Hoot and Taylor (2001) identified the nuclear ribosomal gene internal transcribed spacer (ITS), a LEAFY homolog nuclear gene intron (LFY), and the plastid atpB-rbcL spacer region as informative markers in Isoëtes. However, although these markers and the plastid rbcL gene show utility in large-scale, global phylogenies, they generally lose resolution at the regional level (Rydin and Wikström, 2002; Hoot et al., 2006; Larsén and Rydin, 2016). LFY is more variable than the other three markers and is fairly informative in recently diverged species groups (Taylor et al., 2004; Hoot et al., 2004). With only a single informative nuclear marker within groups such as the eastern North American clade, it is difficult to fully test phylogenetic hypotheses of reticulate evolution and incomplete lineage sorting.

Transcriptomes provide a valuable tool for marker selection and PCR primer design in the absence of a sequenced genome, as is the case in Isoëtes. Databases such as the 1000 Plants project (http://www.onekp.com; Matasci et al., 2014) contain transcriptomes across all major lineages of land plants, allowing identification of unique marker regions for a group of interest. Here we describe use of transcriptome data to develop PCR primers for phylogenetically informative low-copy nuclear markers in Isoëtes.

**METHODS AND RESULTS**

Markers of interest were selected based on a literature search of reportedly low-copy nuclear markers in ferns and mosses (Table 1; Szövényi et al., 2006; Schuettpelz et al., 2008; Rothfels et al., 2013). Nucleotide sequences for these markers were obtained from the National Center for Biotechnology Information's (NCBI) GenBank (http://www.ncbi.nlm.nih.gov/genbank/; Clark et al., 2016) or TreeBASE (http://www.treebase.org; Sanderson et al., 1994) databases. Transcriptomes for three Isoëtes taxa were provided by other sources (I. echinospora Durieu from S. Hetherington, University of Oxford, Oxford, United Kingdom; and I. tegetiformans Rury and an unnamed Isoëtes species from the 1000 Plants project [http://www.oneKP.com]). Using the BLAST+ 2.4 software package (Camacho et al., 2009), local BLAST databases were constructed from each Isoëtes transcriptome. The sequences of selected fern
Markers were selected for Sanger sequencing based on their producing a single band across all samples and for a maximum size of ~1000 bp. PCR products were treated with ExoSAP-IT PCR cleanup enzyme mix (Affymetrix Inc., Santa Clara, California, USA) before cycle sequencing with BigDye Terminator v3.1 (Thermo Fisher Scientific Inc.). The labeled sequencing fragments were read on an ABI 3130xl Genetic Analyzer (Thermo Fisher Scientific Inc.), and the resulting chromatograms were edited and analyzed using Geneious (Kearse et al., 2012).

Initial screening of primers showed that all amplifying in at least some of the eastern North American taxa. Gel electrophoresis revealed that IBR3_1 and Transducin_2 are too long (~2000 bp) and Transducin_1 has both short and long copies in some individuals (~500 bp and ~1000 bp), making these poor candidates for a Sanger sequencing approach without needing molecular cloning or gel extraction. Although gapC_short readily amplified, it is contained within gapC_long, making sequencing of the shorter fragment redundant. pgIC, IBR3_2 (hereafter IBR3), and gapC_long (hereafter gapC) were selected for PCR and sequencing of the full taxa list (Appendices 2, 3).

**pgIC**

This primer pair is rooted in exons 14 and 16, and amplifies across introns 14, 15, and exon 15 of this locus (Rothfels et al., 2013). The region amplified easily across all taxa of *Isoëtes* and *Lycopodium clavatum* L., and generated consistently high-quality sequence data. All sequences aligned well, with a total alignment length of 466 bp and pairwise identity of 83%. Excluding *L. clavatum*, alignment length decreases to 357 bp and pairwise identity increases to 89%. Sequence length between these species of *Isoëtes* ranges from 310 to 347 bp, with a mean of 324 bp (Table 2). This is approximately half the length of the same region in ferns tested by Rothfels et al. (2013).

**gapC**

The gapC gene encodes cytosolic glyceraldehyde-3-phosphate and is part of the GAPDH gene family (Strand et al., 1997; Wall, 2002; Szövényi et al., 2006). Primers designed by Szövényi et al. (2006) are rooted in exons 5 and 9 and amplify all exons and introns in between. However, given concern that the resulting marker in *Isoëtes* may be too long for Sanger sequencing, the primers designed for this study were rooted in exons 5 and 8, amplifying introns 5, 6, 7, and exons 6 and 7. This marker showed the least ability to routinely generate high-quality sequence data. Although not detected in any of the transcriptomes available, it is possible this results from off-target amplification of other members of the GAPDH gene family (i.e., gapCp or an unnamed gapC/gapCp relative) (Schuettelpelz et al., 2013).

---

**TABLE 1.** Primers designed for low-copy markers identified in *Isoëtes* transcriptomes.

| Marker ID | Primer names | Primer sequences (5’–3’) | T<sub>a</sub> (°C) |
|-----------|--------------|--------------------------|------------------|
| pgIC      | pgIC<sub>1</sub>115F | F: GGTCTCTAATGTGCCTGAAATG  | 55               |
|           | pgIC<sub>1</sub>1900R | R: GTCTCCTAAATCTCTTTTCCT  |                  |
| IBR3<sub>1</sub>  | IBR3<sub>1</sub>2F | F: CTAACCATGGCTCAGAATTT  | 60               |
|           | IBR3<sub>1</sub>6R | R: AGCTCCCCACTCAACACAGC  |                  |
| IBR3<sub>2</sub>  | IBR3<sub>1</sub>3F | F: CAATGACGAAAGGCGAATTTG | 60               |
|           | IBR3<sub>1</sub>6R | R: GACCAAAGGTCTAATCCGAG  |                  |
| Transducin<sub>1</sub> | Transducin<sub>1</sub>1F | F: GATGTTGTTGCGATCTTG  | 55               |
|          | Transducin<sub>1</sub>1R | R: CACCTTATGGAATCTCAG  |                  |
| gapC<sub>-</sub>short | gapC<sub>-</sub>5F | F: GAACTCATGCTGTCCTCAC  | 55               |
|          | gapC<sub>-</sub>7R | R: TCTGGTTATATATCCATGGCG  |                  |
| gapC<sub>-</sub>long | gapC<sub>-</sub>5F | F: GAACTTCTAGGGTCCTCAC  | 55               |
|          | gapC<sub>-</sub>9R | R: ATGGCTCAATAGCCTTACG  |                  |

Note: F: forward; R: reverse; T<sub>a</sub>: annealing temperature.

**TABLE 2.** Alignment statistics for all sequences with quality scores >85%.

| Marker | Isoëtes | Lycopodium + Isoëtes |
|--------|---------|----------------------|
|        | Amplicon length range, bp (Mean) | Alignment length, bp | Pairwise % identity | No. of identical sites (%) | No. of PIS (%) | Amplicon length range, bp (Mean) | Alignment length, bp | Pairwise % identity | No. of identical sites (%) | No. of PIS (%) |
| pgIC   | 310–347 (324) | 557 | 89 | 240 (67) | 80 (22) | 310–458 (331) | 466 | 83 | 192 (41) | 82 (18) |
| IBR3   | 587–682 (659) | 700 | 87 | 415 (59) | 111 (16) | — | — | — | — | — |
| gapC   | 443–543 (507) | 561 | 85 | 304 (54) | 95 (17) | — | — | — | — | — |

Note: PIS = parsimony informative sites.
2008; Rothfels et al., 2013). The Isoëtes-only alignment is 561 bp and has a pairwise identity of 85% (Table 2).

**IBR3**

Unlike pgIC and gapC, this gene does not have an extensive history of use as a phylogenetic marker. The IBR3 gene is thought to encode an indole-3-butyric acid–specific peroxisomal enzyme related to acyl-CoA dehydrogenases (Zolman et al., 2007). Rothfels et al. (2013) showed it to be single-copy throughout selected fern lineages, and this also appears to be the case in Isoëtes. Primers for the IBR3 marker amplify most species of Isoëtes easily, with the exception of two members of the Mediterranean clade (*I. histrix* Bory & Durieu and *I. nuttallii* A. Braun ex Engelm.). Alignment of Isoëtes sequences is 700 bp long with 87% pairwise identity (Table 2).

**CONCLUSIONS**

Transcriptome mining is shown to be a useful tool for identification of putative low-copy markers for primer design. Despite having access to transcriptomes of just three species of Isoëtes in the North American clade, primers could be designed for regions that show phylogenetic signal across widely divergent clades in the genus, and potentially across all Lycopodiophyta. Although techniques such as target enrichment allow for generation of data sets orders of magnitude larger (Mandel et al., 2014), design of primers for Sanger sequencing is still more time- and cost-efficient in taxonomic groups for which just a few markers may be needed to infer well-resolved phylogenies.

**ACKNOWLEDGMENTS**

The authors wish to thank Sandy Hetherington of the University of Oxford (Oxford, United Kingdom) and the 1000 Plants project (http://www.oneKP.com) for providing Isoëtes transcriptomes used for primer design. Funding was provided by the Mary Payne Hogan endowment at Old Dominion University and the Smithsonian Institution Fellowship Program. All or portions of the laboratory and/or computer work were conducted in and with the support of the Laboratories of Analytical Biology of the National Museum of Natural History or its partner labs.

**LITERATURE CITED**

Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, and T. L. Madden. 2009. BLAST+: Architecture and applications. *BMC Bioinformatics* 10: 421.

Clark, K., I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and E. W. Sayers. 2016. GenBank. *Nucleic Acids Research* 44(D1): D67–D72.

Foster, A. S., and E. M. Gifford Jr. 1974. Comparative morphology of vascular plants. W. H. Freeman and Company, San Francisco, California, USA.

Hoot, S. B., and W. C. Taylor. 2001. The utility of nuclear ITS, a *LEAFY* homolog intron, and chloroplast *atpB-rbcL* spacer region data in phylogenetic analyses and species delimitation in Isoëtes. *American Fern Journal* 91(3): 166–177.

Hoot, S. B., N. S. Napier, and W. C. Taylor. 2004. Revealing unknown or extinct lineages within Isoëtes (Isoëtaceae) using DNA sequences from hybrids. *American Journal of Botany* 91(6): 899–904.

Hoot, S. B., W. C. Taylor, and N. S. Napier. 2006. Phylogeny and biogeography of Isoëtes (Isoëtaceae) based on nuclear and chloroplast DNA sequence data. *Systematic Botany* 31(3): 449–460.

Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, et al. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28(12): 1647–1649.

Larsén, E., and C. Rydin. 2016. Disentangling the phylogeny of Isoëtes (Isoëtaceae), using nuclear and plastid data. *International Journal of Plant Sciences* 177(2): 157–174.

Mandel, J. R., R. B. Dikow, V. A. Funk, R. R. Masalia, S. E. Staton, A. Kozik, R. W. Michelmore, et al. 2014. A target enrichment method for gathering phylogenetic information from hundreds of loci: An example from the Compositae. *Applications in Plant Sciences* 2(2): 1300085.

Matasci, N., L. H. Hung, Z. Yan, E. J. Carpenter, N. J. Wickett, S. Mirarab, N. Nguyen, et al. 2014. Data access for the 1,000 Plants (1KP) project. *GigaScience* 3: 17. https://doi.org/10.1186/2047-217X-3-17.

Rothfels, C. J., A. Larsson, F. W. Li, E. M. Sigel, L. Huiet, D. O. Burge, M. Ruhsam, et al. 2013. Transcriptome-mining for single-copy nuclear markers in ferns. *PLoS One* 8(10): e76957.

Rydin, C., and N. Wikström. 2002. Phylogeny of Isoëtes (Lycopsida): Resolving basal relationships using *rbcL* sequences. *Taxon* 51(1): 83–89.

Sanderson, M. J., M. J. Donoghue, W. H. Piel, and T. Eriksson. 1994. TreeBASE: A prototype database of phylogenetic analyses and an interactive tool for browsing the phylogeny of life. *American Journal of Botany* 81(6): 183.

Schuettpelz, E., A. L. Grusz, M. D. Windham, and K. M. Pryer. 2008. The utility of nuclear gapCp in resolving polyplody fern origins. *Systematic Botany* 33(4): 621–629.

Strand, A. E., J. Leebans-Mack, and B. G. Milligan. 1997. Nuclear DNA-based markers for plant evolutionary biology. *Molecular Ecology* 6: 113–118.

Szövényi, P., Z. Hock, E. Urmí, and J. Schneller. 2006. New primers for amplifying the GapC gene in bryophytes and its utility in infraspecific phylogenies in the genus *Sphagnum*. *Lindbergia* 31: 78–84.

Taylor, W. C., A. R. Lekchias, Q. F. Wang, X. Liu, N. S. Napier, and S. B. Hoot. 2004. Phylogenetic relationships of Isoëtes (Isoëtaceae) in China as revealed by nucleotide sequences of the nuclear ribosomal ITS region and the second intron of a LEAFY homolog. *American Fern Journal* 94(4): 196–205.

Wall, D. P. 2002. Use of the nuclear gene glyceraldehyde 3-phosphate dehydrogenase for phylogeny reconstruction of recently diverged lineages in *Mitthyridium* (Musci: Calymperaceae). *Molecular Phylogenetics and Evolution* 25(1): 10–26.

Zolman, B. K., M. Nyberg, and B. Bartel. 2007. IBR3, a novel peroxisomal acyl-CoA dehydrogenase-like protein required for indole-3-butyric acid response. *Plant Molecular Biology* 64: 59–72.
**APPENDIX 1.** Collection locations, vouchers, and GenBank accessions for taxa included in this study.

| Taxon | Phylogenetic clade | Collection locality | Voucher (Herbarium) | GenBank accession no. |
|-------|--------------------|---------------------|---------------------|----------------------|
| *Isoëtes butleri* Engelm. | Clade E | Texas, USA | Schafran 47 (ODU) | KY243331 KY270816 KY270832 |
| *I. echinospora* Durieu | Clade E | New York, USA | Schafran NY-4 (ODU) | KY243333 KY270818 KY270835 |
| *I. engelmannii* A. Braun | Clade E | Tennessee, USA | Schafran 46 (ODU) | KY243334 KY270819 — |
| *I. flaccida Shuttlew. var. chapmani* Engelm. (= *I. flaccida*) | Clade E | Florida, USA | Bolin JB_FL_01 (ODU) | KY243332 KY270817 KY270833 |
| *I. flaccida* var. *flaccida* | Clade E | Florida, USA | Schafran FL-01 (ODU) | KY243335 KY270820 KY270836 |
| *I. histrix Bory & Durieu* | Clade E | Sicily, Italy | A. Troia s.n. | KY243347 — — |
| *I. lithophila* N. Pfeiff. | Clade E | Texas, USA | Schafran 61 (ODU) | KY243336 KY270822 KY270838 |
| *I. longissima* Bory | Clade B | Sicily, Italy | A. Troia s.n. | KY243348 KY270823 KY270839 |
| *I. melanospora* J. Gay & Durieu subsp. *melanopoda* | Clade E | Mississippi, USA | Taylor 6796 (US) | KY243338 KY270825 KY270841 |
| *I. melanospora* subsp. *silvatica* (D. F. Brunet & D. M. Britton) | Clade E | North Carolina, USA | Schafran NC-05 (ODU) | KY243342 KY270828 KY270845 |
| *I. melanopoda* Engelm. | Clade E | Georgia, USA | Schafran 12 (ODU) | KY243339 KY270826 KY270842 |
| *I. nuttallii* A. Braun ex Engelm. | Clade B | California, USA | Taylor 6734 (US) | KY243351 — — |
| *I. pietramontana* (N. Pfeiff.) C. F. Reed | — | Georgia, USA | Schafran 18 (ODU) | KY243341 KY270827 KY270844 |
| *I. storkii* T. C. Palmer | Clade E | Costa Rica | Taylor 6760 (US) | KY243352 KY270829 KY270846 |
| *I. teginiformans* Rury | — | Georgia, USA | Schafran 19 (ODU) | KY243343 KY270830 KY270847 |
| *I. valida* (Engelm.) Clute | Clade E | Pennsylvania, USA | Schafran 37 (ODU) | KY243344 KY270831 — |
| *Lycopodium clavatum* L. | — | New York, USA | Schafran s.n. | MG434746 — — |

---

*aOne individual was sampled per taxon.*

*bPer Larsén and Rydin (2016).*

*cHerbaria are abbreviated according to Index Herbariorum (http://sweetgum.nybg.org/science/ih/).*

*dTissue samples provided by A. Troia (Università degli Studi di Palermo, Palermo, Italy); not deposited in a recognized herbarium.*

*eVoucher deposited in P. Schafran’s personal collection.*

**APPENDIX 2.** Amplification and sequence quality of markers across taxa.

| Taxon | Amplification | Sequencing |
|-------|---------------|------------|
| *Isoëtes butleri* | + + + + + | + + + |
| *I. echinospora* | + + + + + | + + + |
| *I. engelmannii* | + + + + + | + + + |
| *I. flaccida var. chapmani* | + + + + + | + + + |
| *I. flaccida var. flaccida* | + + + + + | + + + |
| *I. histrix* | — — — | NA NA |
| *I. lithophila* | + + + + + | + + + |
| *I. longissima* | + + + + + | + + + |
| *I. melanospora* subsp. *melanopoda* | + + + + + | + + + |
| *I. melanospora* subsp. *silvatica* | + + + + + | + + + |
| *I. nuttallii* | + — — + | — — |
| *I. pietramontana* | + + + + + | + + + |
| *I. storkii* | + + + + + | + + + |
| *I. teginiformans* | + + + + + | + + + |
| *I. valida* | + + + | + + |
| *Lycopodium clavatum* | + — + + | NA — |

**Note:** + = successful amplification or sequence quality >85%; — = no amplification or sequence quality <85%; NA = sequencing not attempted.
APPENDIX 3. Pairwise number of nucleotide differences between pgiC/IBR3/gapC sequences.

| Species                  | 1. butleri | 1. echinospora | 1. engelmannii var. chapmanii | 1. flaccida var. flaccida | 1. histrix | 1. lithophila | 1. lonchisima | 1. melanopoda subsp. melanopoda | 1. melanopoda subsp. silvatica | 1. melanospora | 1. nuttallii | 1. piedmontana | 1. storkii | 1. tegetiformans s | 1. valida |
|--------------------------|------------|----------------|-----------------------------|--------------------------|-----------|-------------|-------------|------------------------------|--------------------------|--------------|-------------|----------------|---------|----------------|---------|
| I. echinospora           | 10/34/27   | 10/34/27       | 10/34/27                    | 10/34/27                 | 10/34/27  | 10/34/27    | 10/34/27    | 10/34/27                                   | 10/34/27                 | 10/34/27    | 10/34/27    | 10/34/27    | 10/34/27 | 10/34/27       | 10/34/27 |
| I. engelmannii var. chapmanii | 10/34/27 | 10/34/27       | 10/34/27                    | 10/34/27                 | 10/34/27  | 10/34/27    | 10/34/27    | 10/34/27                                   | 10/34/27                 | 10/34/27    | 10/34/27    | 10/34/27    | 10/34/27 | 10/34/27       | 10/34/27 |
| I. flaccida var. flaccida | 10/34/27   | 10/34/27       | 10/34/27                    | 10/34/27                 | 10/34/27  | 10/34/27    | 10/34/27    | 10/34/27                                   | 10/34/27                 | 10/34/27    | 10/34/27    | 10/34/27    | 10/34/27 | 10/34/27       | 10/34/27 |
| I. histrix               | 10/34/27   | 10/34/27       | 10/34/27                    | 10/34/27                 | 10/34/27  | 10/34/27    | 10/34/27    | 10/34/27                                   | 10/34/27                 | 10/34/27    | 10/34/27    | 10/34/27    | 10/34/27 | 10/34/27       | 10/34/27 |
| I. lithophila            | 10/34/27   | 10/34/27       | 10/34/27                    | 10/34/27                 | 10/34/27  | 10/34/27    | 10/34/27    | 10/34/27                                   | 10/34/27                 | 10/34/27    | 10/34/27    | 10/34/27    | 10/34/27 | 10/34/27       | 10/34/27 |
| I. lonchisima            | 10/34/27   | 10/34/27       | 10/34/27                    | 10/34/27                 | 10/34/27  | 10/34/27    | 10/34/27    | 10/34/27                                   | 10/34/27                 | 10/34/27    | 10/34/27    | 10/34/27    | 10/34/27 | 10/34/27       | 10/34/27 |
| I. melanopoda subsp. melanopoda | 10/34/27 | 10/34/27       | 10/34/27                    | 10/34/27                 | 10/34/27  | 10/34/27    | 10/34/27    | 10/34/27                                   | 10/34/27                 | 10/34/27    | 10/34/27    | 10/34/27    | 10/34/27 | 10/34/27       | 10/34/27 |
| I. melanopoda subsp. silvatica | 10/34/27 | 10/34/27       | 10/34/27                    | 10/34/27                 | 10/34/27  | 10/34/27    | 10/34/27    | 10/34/27                                   | 10/34/27                 | 10/34/27    | 10/34/27    | 10/34/27    | 10/34/27 | 10/34/27       | 10/34/27 |
| I. nuttallii             | 10/34/27   | 10/34/27       | 10/34/27                    | 10/34/27                 | 10/34/27  | 10/34/27    | 10/34/27    | 10/34/27                                   | 10/34/27                 | 10/34/27    | 10/34/27    | 10/34/27    | 10/34/27 | 10/34/27       | 10/34/27 |
| I. piedmontana           | 10/34/27   | 10/34/27       | 10/34/27                    | 10/34/27                 | 10/34/27  | 10/34/27    | 10/34/27    | 10/34/27                                   | 10/34/27                 | 10/34/27    | 10/34/27    | 10/34/27    | 10/34/27 | 10/34/27       | 10/34/27 |
| I. storkii               | 10/34/27   | 10/34/27       | 10/34/27                    | 10/34/27                 | 10/34/27  | 10/34/27    | 10/34/27    | 10/34/27                                   | 10/34/27                 | 10/34/27    | 10/34/27    | 10/34/27    | 10/34/27 | 10/34/27       | 10/34/27 |
| I. tegetiformans s       | 10/34/27   | 10/34/27       | 10/34/27                    | 10/34/27                 | 10/34/27  | 10/34/27    | 10/34/27    | 10/34/27                                   | 10/34/27                 | 10/34/27    | 10/34/27    | 10/34/27    | 10/34/27 | 10/34/27       | 10/34/27 |
| I. valida                | 10/34/27   | 10/34/27       | 10/34/27                    | 10/34/27                 | 10/34/27  | 10/34/27    | 10/34/27    | 10/34/27                                   | 10/34/27                 | 10/34/27    | 10/34/27    | 10/34/27    | 10/34/27 | 10/34/27       | 10/34/27 |
| Lycopodium clavatum      | 10/34/27   | 10/34/27       | 10/34/27                    | 10/34/27                 | 10/34/27  | 10/34/27    | 10/34/27    | 10/34/27                                   | 10/34/27                 | 10/34/27    | 10/34/27    | 10/34/27    | 10/34/27 | 10/34/27       | 10/34/27 |