LETTER TO THE EDITOR

Sources of inversion variation in the small single copy (SSC) region of chloroplast genomes[1]

Joseph F. Walker[2], Robert K. Jansen[3,4], Michael J. Zanis[5], and Nancy C. Emery[6,7]

Modern sequencing technology has led to a proliferation of whole-genome sequences of chloroplasts in a growing number of plant lineages, bringing opportunities for comparisons that provide insights into the evolutionary history of the plastomes and their host plants (Jansen et al., 2007; Doorduin et al., 2011). Amid the emerging literature in this area is a hypothesis that the small single copy (SSC) region is a "hotspot" for inversion events (sensu Liu et al., 2013) because different orientations of the region have been reported in relatively high frequencies among closely related taxa (Liu et al., 2013; Walker et al., 2014). We would like to draw attention to a study by Palmer (1983) that bears heavily on this discussion, yet has been overlooked by several authors of publications investigating whole-chloroplast genome sequence order, including one study by some of the authors of this letter (Walker et al., 2014). Using restriction enzyme analyses, Palmer (1983) demonstrated that chloroplast DNA within individual plants exhibits a form of heteroplasmy in which the plastome exists in two equimolar states (i.e., inversion isomers) that differ in the relative orientation of the small single copy (SSC) region. Since Palmer (1983) originally documented this phenomenon in *Phaseolus vulgaris*, it has been confirmed in a wide variety of plant species (Palmer, 1985), including algae (Aldrich et al., 1985; Bourne et al., 1992; Linne von Berg and Kowallik, 1992; Cattolico et al., 2008) and ferns (Stein et al., 1986), and is now a well-accepted feature of chloroplast genomes (e.g., Heinhorst and Cannon, 1993; Doyle and Doyle, 1999). Nonetheless, this phenomenon has been overlooked in several recent analyses that have evaluated the orientation of the SSC region a phylogenetic context (e.g., Ibrahim et al., 2006; Yang et al., 2010; Liu et al., 2013; Walker et al., 2014; Zhang et al., 2014; Wang et al., 2015). These analyses compare the SSC orientation among lineages using a single plastome to represent each lineage and thus have missed the within-individual variation that exists in this region. Currently, whole-chloroplast genomes are published in GenBank without preference for the orientation of the SSC region, leading to apparent variation in the orientation of the SSC region *among* individuals that is actually due to chloroplast heteroplasmy *within* individuals (Wolfe and Randle, 2004), as originally described by Palmer (1983). For example, two sequences of *Lactuca sativa* that have been independently published (NC_007578 and DQ_383816) were entered with different orientations of the SSC region, which could be interpreted as a major inversion existing within the species if the investigators are not aware that two isomers naturally exist (e.g., Walker et al., 2014). This misinterpretation has now occurred in several studies (e.g., Ibrahim et al., 2006; Yang et al., 2010; Liu et al., 2013; Walker et al., 2014; Zhang et al., 2014; Wang et al., 2015), leading to the hypothesis that the SSC region is an inversion "hotspot" (Liu et al., 2013). Experiments that attempted to use PCR to allegedly confirm a single orientation of the SSC region within samples have likely perpetuated this misconception (e.g., Nie et al., 2012; Liu et al., 2013). Specifically, the reverse complementarity of the IR regions within chloroplast sequences, and the SSC regions between different isomers, inhibit the PCR approach (which relies on sequence orientation) from detecting the two different isomers, giving the impression that individual plant lineages have only one isomer or the other (e.g., Liu et al., 2013).

To date, it is not entirely clear how the two SSC orientations are maintained within cells (Maréchal and Brisson, 2010). Originally, it was proposed that the two states were the result of intramolecular recombination between the two inverted repeat (IR) regions that flank the SSC in the circular plastome (Palmer, 1983). More recent data indicate that recombination-dependent DNA replication of the chloroplast genome in its linear (as opposed to circular) conformation may provide the mechanism underlying flip–flop recombination (Oldenburg and Bendich, 2004). Regardless of the mechanism, it is misleading to refer to the SSC region as a "hotspot" for inversions because the two orientations of the SSC region occur regularly during the course of chloroplast DNA replication within individual plant

[1] Manuscript received 23 June 2015; revision accepted 29 September 2015

[2] Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan 48109 USA;

[3] Department of Integrative Biology, University of Texas at Austin, Austin, Texas 78712 USA;

[4] Biotechnology Research Group, Department of Biological Sciences, King Abdulaziz University, Jeddah 21589 Saudi Arabia;

[5] Department of Biology, Seattle University, Seattle, Washington 98122 USA; and

[6] Department of Ecology & Evolutionary Biology, Campus Box 334, University of Colorado Boulder, Boulder, Colorado 80309-0334 USA

[7] Author for correspondence (e-mail: nancy.emery@colorado.edu), phone: (303) 735-7548, fax: (303) 492-8699

doi:10.3732/ajb.1500299

cells, rather than the relatively rare inversion events that can be used to distinguish distantly related lineages (e.g., Doyle et al., 1992, 1996).

LITERATURE CITED

Aldrich, J., B. Cherney, E. Merlin, C. Williams, and L. Mets. 1985. Recombination within the inverted repeat sequences of the *Chlamydomonas reinhardii* chloroplast genome produces 2 orientation isomers. *Current Genetics* 9: 233–238.

Bourne, C. M., J. D. Palmer, and E. F. Stoermer. 1992. Organization of the chloroplast genome of the freshwater centric diatom *Cyclotella meneghiniana. Journal of Phycology* 28: 347–355.

Cattolico, R. A., M. A. Jacobs, Y. Zhou, J. Chang, M. Duplessis, T. Lybrand, J. McKay, et al. 2008. Chloroplast genome sequencing analysis of *Heterosigma akashiwo* CCMP452 (West Atlantic) and NIES293 (West Pacific) strains. *BMC Genomics* 9: 211.

Doorduin, L., B. Gravendeel, Y. Lammers, Y. Ariyurek, T. Chin-A-Woeng, and K. Vrieling. 2011. The complete chloroplast genome of 17 individuals of pest species *Jacobaea vulgaris*: SNPs, microsatellites and barcoding markers for population and phylogenetic studies. *DNA Research* 18: 93–105.

Doyle, J. J., J. I. Davis, R. J. Soreng, D. Garvin, and M. J. Anderson. 1992. Chloroplast DNA inversions and the origin of the grass family (Poaceae). *Proceedings of the National Academy of Sciences, USA* 89: 7722–7726.

Doyle, J. J., and J. L. Doyle. 1999. Nuclear protein-coding genes in phylogeny reconstruction and homology assessment. *In* P. M. Hollingsworth, R. M. Bateman, and R. J. Gornall [eds.], Molecular Systematics and plant evolution, 229–254. CRC Press, Boca Raton, Florida, USA. doi:10.1201/9781439833278.ch12.

Doyle, J. J., J. L. Doyle, J. A. Ballenger, and J. D. Palmer. 1996. The distribution and phylogenetic significance of a 50-kb chloroplast DNA inversion in the flowering plant family leguminosae. *Molecular Phylogenetics and Evolution* 5: 429–438.

Heinhorst, S., and G. C. Cannon. 1993. DNA replication in chloroplasts. *Journal of Cell Science* 104: 1–9.

Ibrahim, R. I. H., J.-I. Azuma, and M. Sakamoto. 2006. Complete nucleotide sequence of the cotton (*Gossypium barbadense* L.) chloroplast genome with a comparative analysis of sequences among 9 dicot plants. *Genes & Genetic Systems* 81: 311–321.

Jansen, R. K., Z. Cai, L. A. Raubeson, H. Daniell, C. W. dePamphilis, J. Leebens-Mack, K. F. Mueller, et al. 2007. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proceedings of the National Academy of Sciences, USA* 104: 19369–19374.

Linne von Berg, K. H., and K. V. Kowallik. 1992. Structural organization of the chloroplast genome of the chromophytic alga *Vaucheria bursata. Plant Molecular Biology* 18: 83–95.

Liu, Y., N. Huo, L. Dong, Y. Wang, S. Zhang, H. A. Young, X. Feng, and Y. Q. Gu. 2013. Complete chloroplast genome sequences of Mongolia medicine *Artemisia frigida* and phylogenetic relationships with other plants. *PLoS One* 8: e57533.

Maréchal, A., and N. Brisson. 2010. Recombination and the maintenance of plant organelle genome stability. *New Phytologist* 186: 299–317.

Nie, X., S. Lv, Y. Zhang, X. Du, L. Wang, S. S. Biradar, X. Tan, et al. 2012. Complete chloroplast genome sequence of a major invasive species, crofton weed (*Ageratina adenophora*). *PLoS One* 7: e36869.

Oldenburg, D. J., and A. J. Bendich. 2004. Most chloroplast DNA of maize seedlings in linear molecules with defined ends and branched forms. *Journal of Molecular Biology* 335: 953–970.

Palmer, J. D. 1983. Chloroplast DNA exists in two orientations. *Nature* 301: 92–93.

Palmer, J. D. 1985. Comparative organization of chloroplast genomes. *Annual Review of Genetics* 19: 325–354.

Stein, D. B., J. D. Palmer, and W. F. Thompson. 1986. Structural evolution and flip–flop recombination of chloroplast DNA in the fern genus *Osmunda. Current Genetics* 10: 835–841.

Walker, J. F., M. J. Zanis, and N. C. Emery. 2014. Comparative analysis of complete chloroplast genome sequence and inversion variation in *Lasthenia burkei* (Madieae, Asteraceae). *American Journal of Botany* 101: 722–729.

Wang, M., L. Cui, K. Feng, P. Deng, X. Du, F. Wan, S. Weining, and X. Nie. 2015. Comparative analysis of Asteraceae chloroplast genomes: Structural organization, RNA editing and evolution. *Plant Molecular Biology Reporter* 33: 1526–1538.

Wolfe, A. D., and C. P. Randle. 2004. Recombination, heteroplasmy, haplotype polymorphism, and paralogy in plastic genes: Implications for plant molecular systematics. *Systematic Botany* 29: 1011–1020.

Yang, M., X. Zhang, G. Liu, Y. Yin, K. Chen, Q. Yun, D. Zhao, et al. 2010. The complete chloroplast genome sequence of date palm (*Phoenix dactylifera* L.). *PLoS One* 5: e12762.

Zhang, Y., L. Li, T. L. Yan, and Q. Liu. 2014. Complete chloroplast genome sequences of *Praxelis* (*Eupatorium catarium* Veldkamp), an important invasive species. *Gene* 549: 58–69.