

Research Article

OPEN ACCESS

# A new R package and web application for detecting bilateral asymmetry in parasitic infections

Matthew T. Wayland<sup>1</sup> and James C. Chubb<sup>2</sup>

<sup>1</sup>Department of Zoology, University of Cambridge, Cambridge, UK;

<sup>2</sup>Department of Evolution, Ecology and Behaviour, Institute of Integrative Biology, University of Liverpool, Liverpool, UK

**Abstract:** When parasites invade paired structures of their host non-randomly, the resulting asymmetry may have both pathological and ecological significance. To facilitate the detection and visualisation of asymmetric infections we have developed a free software tool, Analysis of Symmetry of Parasitic Infections (ASPI). This tool has been implemented as an R package (<https://cran.r-project.org/package=aspi>) and a web application (<https://wayland.shinyapps.io/aspi>). ASPI can detect both consistent bias towards one side, and inconsistent bias in which the left side is favoured in some hosts and the right in others. Application of ASPI is demonstrated using previously unpublished data on the distribution of metacercariae of species of *Diplostomum* von Nordmann, 1832 in the eyes of ruffe *Gymnocephalus cernua* (Linnaeus). Invasion of the lenses appeared to be random, with the proportion of metacercariae in the left and right lenses showing the pattern expected by chance. However, analysis of counts of metacercariae from the humors, choroid and retina revealed asymmetry between eyes in 38% of host fish.

**Keywords:** bilateral asymmetry, Trematoda, *Diplostomum*, fish, *Gymnocephalus cernua*, ruffe, software, statistical analysis

This article contains supporting information (Tables S1–S3) online at <http://folia.paru.cas.cz/suppl/2016-63-039.pdf>

Bilateral asymmetry arises when parasites infect paired structures non-randomly. Two types of asymmetry may be observed. In the first, one side is consistently favoured in all hosts and examples include: echinostome trematodes preferentially invading the right kidney of their larval amphibian intermediate host (Thiemann and Wassersug 2000, Johnson et al. 2014); and the monogenean *Entobdella soleae* (van Benenden et Hesse, 1864) migrating to the lower surface (biological left) of its flatfish host, *Solea solea* (Linnaeus) (see Kearn 1984). The second is an inconsistent bias, with the left side favoured in some hosts and the right in others, as observed in the distribution of metacercariae of species of *Diplostomum* von Nordmann, 1832 (Trematoda) in the eyes of fish (Rau et al. 1979, Dwyer and Smith 1989, Graczyk 1991, Rintamäki-Kinnunen et al. 2004, Karvonen and Seppala 2008) and larvae of the dipteran *Protocalliphora avium* Shannon et Dobrosky in the ears of the hawk, *Buteo lineatus* (Gmelin) (see King et al. 2010).

The mechanisms responsible for bilateral asymmetry are poorly understood, but may involve host morphology, host physiology, host mortality, parasite mortality and cue following (see Johnson et al. 2014). Bilateral asymmetry may have adaptive significance; through localising the

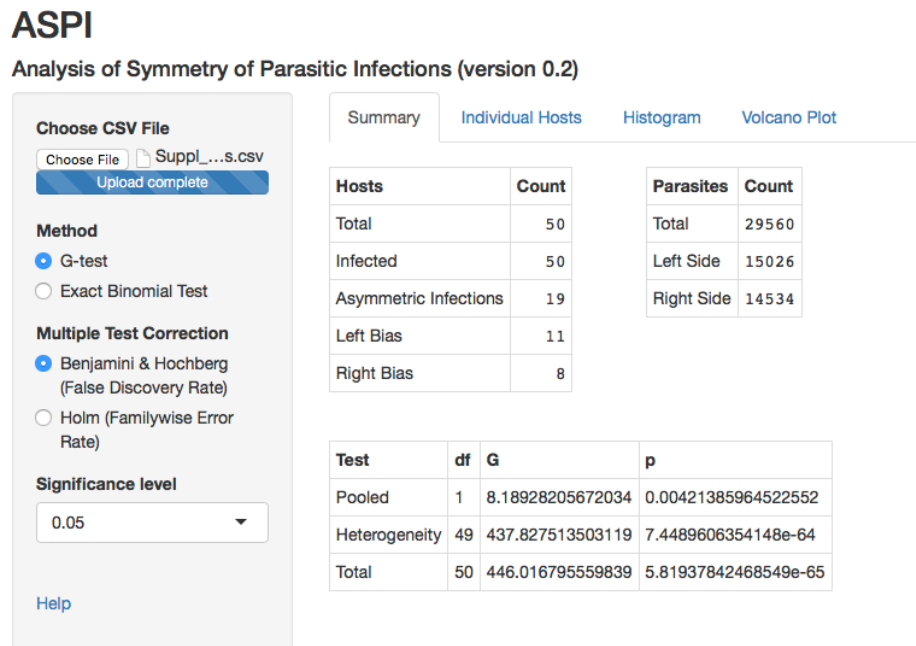
effects of infection it may minimise the pathogenicity of the parasite, with benefits for the fitness of both host and parasite (Johnson et al. 2014). Despite the potential ecological and pathological significance of deviations from symmetry, this aspect of the host-parasite relationship is not routinely investigated in field surveys. To facilitate study of this phenomenon we have developed a new tool, Analysis of Symmetry of Parasitic Infections (ASPI). ASPI is demonstrated here using observations on the distribution of metacercariae of *Diplostomum* spp. in the eyes of ruffe, *Gymnocephalus cernua* (Linnaeus) from Llyn Tegid, a mesotrophic lake in north Wales.

## MATERIALS AND METHODS

### Data collection

Fifty ruffe (45 females and 5 males) were captured between October 1994 and February 1995 using gill nets. Each eye was individually opened in a 10 cm petri dish. The lens were transferred to a separate dish, sliced open and examined under the dissecting microscope for metacercariae. The remainder of the eye was irrigated with water and the retinal and choroidal layers scraped to expose metacercariae.

Address for correspondence: M.T. Wayland, Department of Zoology, University of Cambridge, Downing Street, Cambridge, CB2 3EJ, UK.  
Phone: (+44) (0)1223 769282; Fax: (+44) (0)1223 336676; E-mail: [mw283@cam.ac.uk](mailto:mw283@cam.ac.uk)



**Fig. 1.** Screenshot of the web application. The panel on the left contains the controls for the application, including file upload, selection of test, choice of multiplicity correction and significance threshold. Results are displayed on the four tabbed pages of the main panel: summary, individual hosts, histogram and volcano plot.

### Statistical procedures

Replicated G-tests of goodness-of-fit (Sokal and Rohlf 1995) provide a comprehensive analysis of deviations from bilateral symmetry. This procedure computes four different G-statistics, each related to a different null hypothesis. The total G-test examines the null hypothesis that overall the parasite distributions in all individual hosts do not depart from symmetry. As the name implies, the pooled G-test evaluates the second null hypothesis, that the number of parasites on the left and right sides, summed across all infected hosts, fits a 1 : 1 ratio. The heterogeneity G-test evaluates the third null hypothesis, that the left : right ratios are the same in all individual hosts. The final step is to identify which of the hosts have asymmetrical parasite distributions using the individual G-test.

G-tests involve logarithmic transformation and so cannot be applied to counts of zero. Furthermore, G-tests are not recommended when counts are below ten (McDonald 2014). In this situation, exact binomial tests can be used instead, but these are limited to testing two null hypotheses: (i) the ratio of the total number of parasites on each side does not differ from 1 : 1 (equivalent to the pooled G-test); (ii) the distribution of parasites in an individual host is symmetrical.

It is important to note that when applying G-tests or exact binomial tests to individual hosts, the problem of multiplicity arises. Consider a sample of 50 infected hosts. If the null hypothesis of symmetry is true for all 50 hosts, at a significance level of 5% the probability of getting at least one Type I error (false rejection of the null hypothesis) is  $1-(1-0.05)^{50} = 0.923$ ; the expected number of significant results obtained purely by chance would be  $50 \times 0.05 = 2.5$ . ASPI provides two options for adjusting p-values for multiple individual tests. Holm's (1979) procedure is used to control the family-wise error rate, i.e. the probability of a single Type I error. However, guarding against a single Type I error will be unnecessarily conservative for many studies, especially those

surveying a large number of hosts. If a small proportion of false positives among the set of rejected null hypotheses can be tolerated, then it will be preferable to control the false discovery rate (FDR) using the procedure of Benjamini and Hochberg (1995).

### Software implementation

ASPI's algorithms have been implemented in R (R Core Team 2016) and the web application developed using the shiny framework (Chang et al. 2016). All source code is freely available under version 3 of the GNU General Public License from the following repository: <https://github.com/WaylandM/aspi>.

The web application (<https://wayland.shinyapps.io/aspi>) is intended to present an easy-to-use interface (Fig. 1). A user starts by uploading a comma-separated-value (csv) file containing three columns: host (unique host id), left (number of parasites on left) and right (number of parasites on right). The web application will default to an analysis using replicated G-tests of goodness-of-fit and a significance threshold of 5%, but the user can choose alternative settings from the control panel situated on the left. Full instructions for use are provided in the documentation: <https://github.com/WaylandM/aspi/blob/master/doc/gui.md>. Experienced users of R may prefer to use the ASPI package, which is available from the Comprehensive R Archive Network: <https://cran.r-project.org/package=aspi>. Documentation for this package is self-contained.

### RESULTS

Total length of fish ranged from 105–139 mm (mean  $\pm$  standard deviation,  $123.4 \pm 7.5$  mm) and weight from 13–39 g ( $25.2 \pm 4.6$  g). Using the key of Shigin (1986), trematodes from the lens were tentatively identified as *Diplostomum paraspathaceum* Shigin, 1965 and those from the remainder of the eye as *D. volvens* von Nordmann, 1832 and *Diplostomum gasterostei* Williams, 1966.

**Table 1.** Results of the G-test of heterogeneity applied to the distribution of metacercariae of species of *Diplostomum* von Nordmann, 1832 in the eyes (excluding lenses) of ruffe.

Test	df	G	p
Pooled	1	8.189	0.004
Heterogeneity	49	437.828	$7.449 \times 10^{-64}$
Total	50	446.017	$5.819 \times 10^{-65}$

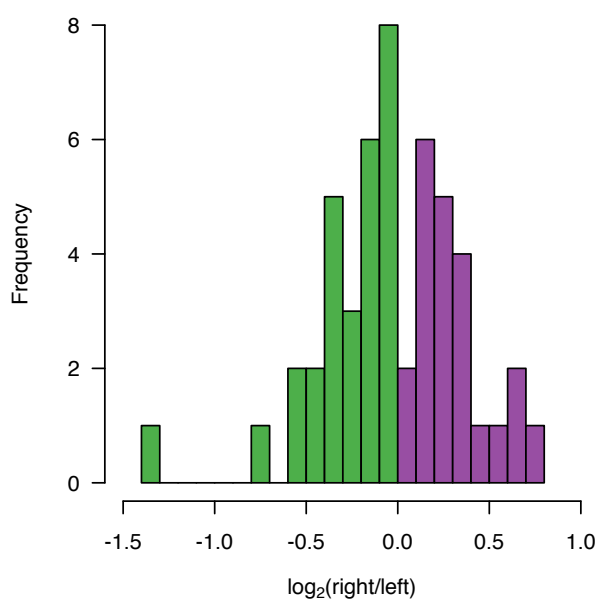
**Table 2.** Number of asymmetric infections detected using two different corrections for multiplicity: Holm's procedure for controlling the family-wise error rate (FWER) and Benjamini and Hochberg's method for controlling the false discovery rate (FDR).

	Holm (FWER)		Benjamini and Hochberg (FDR)	
	p < 0.01	p < 0.05	p < 0.01	p < 0.05
Asymmetric infections	6	11	17	19
Left bias	3	6	9	11
Right bias	3	5	8	8

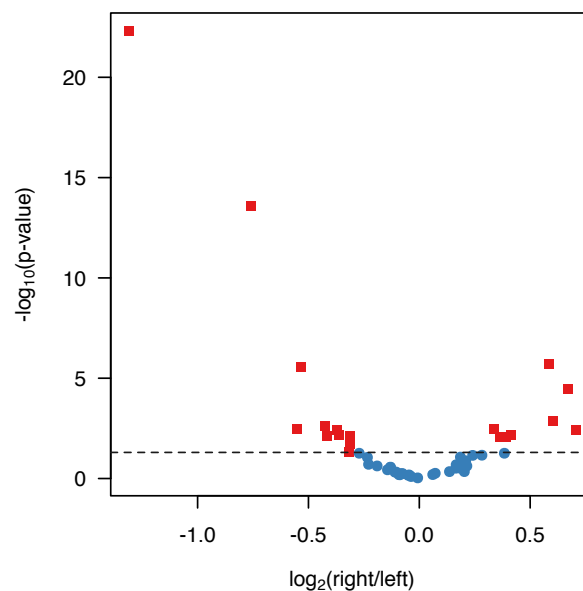
*D. gasterostei* has previously been reported from the ruffe of Llyn Tegid (Faulkner 1989) and *D. volvens* from ruffe in the river basins of north-east Russia (Dorovskikh 1997). However, all of these species determinations are subject to revision in the light of DNA sequence data (Georgieva et al. 2013, Locke et al. 2015). Data for host samples, eye infections and lens infections are provided as Tables S1–S3, respectively.

Infection with species of *Diplostomum* in the eye, excluding the lens, will be considered first. All fish were infected and parasites were over-dispersed in the host population (variance-to-mean ratio  $\sigma^2/\mu = 82.4$ ). The parasite burden of the left eyes ranged from 73–603 (mean  $\pm$  standard deviation,  $300.5 \pm 122.9$ ) and that of the right eyes from 99–586 ( $290.7 \pm 109.7$ ). For the current data set, the total G-test is highly significant (p-value =  $5.819 \times 10^{-65}$ ) and so the null hypothesis of overall symmetry can be rejected (Table 1). The other three G-statistics reveal the nature of the asymmetry.

A total of 15026 metacercariae were recovered from the left eyes (excluding lenses) and 14534 from the right. The pooled G-test (p-value = 0.004) determines this ratio of approximately 1.03 : 1 to be significantly different from 1 : 1. However, this slight left bias in the pooled data does not necessarily mean that there is a left bias in all or even most hosts. Moreover, a highly significant p-value ( $7.449 \times 10^{-64}$ ) for the heterogeneity G-test provides convincing evidence that the proportions of metacercariae in the left and right eyes varies from host to host. At the conventional significance level of 5%, Holm's (1979) procedure finds 11 of the 50 individual G-tests to be significant, i.e. 11 infections are asymmetric (Table 2). Of these, 6 show a left bias and 5 a right bias. Benjamini and Hochberg's (1995) method provides more power to detect asymmetric infections, identifying a total of 19 (11 showing a left bias and 8 a right bias), although one (5% of 19) of these is likely to be a false positive.



**Fig. 2.** Histogram showing distribution of fold differences in abundance of *Diplostomum* spp. metacercariae between left and right eyes (excluding lenses). For each host the number of parasites in the right eye was divided by the number of parasites in the left eye, and the result was binary log transformed. The  $\log_2$  ratio will be negative if there are more parasites in the left than right eye, and positive if there are more parasites in the right than left eye. A  $\log_2$  ratio of one corresponds to a one-fold difference, i.e. double the number of parasites. Perfect symmetry is a  $\log_2$  ratio of zero.



**Fig. 3.** Volcano plot: statistical significance ( $-\log_{10}(\text{p-value})$ ) of G-tests for asymmetry of parasite distributions in individual hosts plotted against fold difference ( $\log_2$  ratio) in parasite abundance between right and left eyes. Each point in the scatterplot represents the parasite distribution between eyes (excluding lenses) in an individual host. The G-test p-values were adjusted using the Benjamini and Hochberg (1995) procedure before  $-\log_{10}$  transformation. The dashed horizontal line represents the p-value threshold for a 5% false discovery rate. Parasite distributions deviating significantly from symmetry are shown as red squares, whereas those not differing significantly from a 1 : 1 ratio are represented by blue circles.

For visualisation the proportion of parasites in each eye can be expressed as a binary log transformed ratio. These ratios can be plotted as a histogram (Fig. 2) or combined with corresponding p-values from individual G-tests in a volcano plot (Fig. 3). Both plots reveal the presence of an outlier; a binary log ratio of -1.3, corresponding to a host with more than twice as many metacercariae in its left eye than in its right (402 vs 162). If this host is excluded from the data-set and the analysis re-run, the total and heterogeneity G-tests remain highly significant (p-values of  $2.783 \times 10^{-45}$  and  $2.672 \times 10^{-45}$ , respectively), but the pooled G-test does not ( $p = 0.139$ ). This result demonstrates the sensitivity of the pooled test to outliers and the importance of examining parasite distribution in individual hosts.

Infection of the lens with *Diplostomum paraspathaceum* was present in only 31 of the 50 fish sampled and the lens metacercariae were overdispersed in the host population ( $\sigma^2/\mu = 6.9$ ). The parasite burden of infected ruffe ranged from 1 to 24 ( $7.48 \pm 5.48$ ), with the left lens containing 0–15 ( $3.81 \pm 3.65$ ) metacercariae and the right lens 0–9 ( $3.68 \pm 2.61$ ). Since this dataset contained zero counts, replicated G-tests of goodness-of-fit were not applicable and exact binomial tests were used instead. No evidence of asymmetry in lens infections was found in either the pooled test ( $p = 0.84$ ) or the tests of individual host infections (smallest FDR corrected  $p = 0.484$ ).

## DISCUSSION

Few studies have investigated deviations from symmetry in parasitic infections. However, evidence is starting to emerge that asymmetry in infections of paired organs may be beneficial to host fitness, and consequently parasite fitness (Johnson et al. 2014). For example, in infections with echinostomes (Trematoda) of amphibian kidneys, host condition is positively correlated with the level of parasite bias towards the right kidney (Johnson et al. 2014). Therefore, when investigating the impact of parasites on hosts it may be useful to supplement conventional metrics such as intensity of infection, with measures of asymmetry of infection (e.g. individual G-test statistic).

ASPI can detect both consistent bias towards one side, and inconsistent bias in which the left side is favoured in some hosts and the right in others. The latter, less obvious form of asymmetry is sometimes overlooked when lateralisation of infection is investigated (e.g. Buchmann 1988, Rodrigues and Saraiva 1996, Dzika 1999, Stumbo and Poulin 2016).

## REFERENCES

- BENJAMINI Y., HOCHBERG Y. 1995: Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. B* 57: 289–300.
- BUCHMANN K. 1988: Spatial distribution of *Pseudodactylogyrus anguillae* and *P. bini* (Monogenea) on the gills of the European eel, *Anguilla anguilla*. *J. Fish Biol.* 32: 801–802.
- CHANG W., CHENG J., ALLAIRE J.J., XIE Y., MCPHERSON J. 2016: shiny: Web Application Framework for R. R package version 0.13.2. <https://CRAN.R-project.org/package=shiny>
- DOROVSKIKH G.N. 1997: [Results of the study of fishes' parasites in river basins of the north-east of the European part of Russia. Trematodes (Trematoda).] *Parazitologiya* 21: 551–564. (In Russian.)
- DWYER W.P., SMITH C.E. 1989: Metacercariae of *Diplostomum spathaceum* in the eyes of fishes from Yellowstone Lake, Wyoming. *J. Wildlife Dis.* 25: 126–129.
- DZIKA E. 1999: Microhabitats of *Pseudodactylogyrus anguillae* and *P. bini* (Monogenea: Dactylogyridae) on the gills of large-

Another important feature of ASPI is its capability to detect differences in parasite distribution between hosts. Some studies have restricted their analysis to pooled data (i.e. counts of the number of parasites on the left and right sides, summed across all infected hosts), which obfuscates the patterns in individual hosts and may lead to erroneous conclusions about symmetry or lack thereof. For example, the symmetry of infections of *Pseudodactylogyrus anguillae* Yin et Sproston, 1948 and *P. bini* Kikuchi, 1929 (Monogenea) on the gills of European eel *Anguilla anguilla* (Linnaeus) was investigated using pooled data alone, in three separate studies (Buchmann 1988, Rodrigues and Saraiva 1996, Dzika 1999). In the case of *P. anguillae*, Buchmann (1988) reported that the parasites show a preference for the left, Rodrigues and Saraiva (1996) determined a bias to the right and Dzika (1999) found no evidence of lateralisation. For *P. bini* infections, both Buchmann (1988) and Rodrigues and Saraiva (1996) reported right bias, but Dzika (1999) was unable to refute the null hypothesis of symmetry. These apparently contradictory findings might reflect real differences between the different host and/or parasite populations sampled. Alternatively the lack of reproducibility may be an artifact of the pooled test procedure. The current study has demonstrated that the presence of just one outlier can make the difference between accepting or rejecting the null hypothesis of a pooled test.

Parallels can be seen between deviations from symmetry in parasitic infections and the concept of fluctuating asymmetry (see Graham et al. 2010 for a recent review). However, there are also important distinctions. Studies of fluctuating asymmetry are primarily concerned with deriving a measure of developmental stability from small, random deviations from phenotypic symmetry. In contrast, the current study of parasite distribution is concerned with the detection of non-random, sometimes large, deviations from symmetry. Investigation of the potential relationship between symmetry of parasitic infection and symmetry of host phenotype is an obvious, but interesting topic for future research.

In summary, ASPI provides a comprehensive statistical and graphical analysis of the distribution of parasites between paired host organs, overcoming the limitations of some approaches used in earlier studies. The authors welcome feedback from users to improve the functionality and usability of the software.

**Acknowledgements.** The authors would like to thank two anonymous referees for helpful comments.

- size European eel *Anguilla anguilla* from Lake Gaj, Poland. *Folia Parasitol.* 46: 33–36.
- FAULKNER M. 1989: The application of sodium dodecyl sulphate-polyacrylamide gel electrophoresis to the taxonomic identification of the total body protein band profiles of *Diplostomum* spp. metacercariae (Digenea), parasites of fish eyes. *Electrophoresis* 10: 260–264.
- GEORGIEVA S., SOLDÁNOVÁ M., PÉREZ-DEL-OLMO A., DANGEL D.R., SITKO J., SURES B., KOSTADINOVA A. 2013: Molecular prospecting for European *Diplostomum* (Digenea: Diplostomidae) reveals cryptic diversity. *Int. J. Parasitol.* 43: 57–72.
- GRACZYK T. 1991: Cases of bilateral asymmetry of *Diplostomum pseudospathaceum* Niewiadomska, 1984 metacercariae infections (Trematoda, Diplostomidae) in the eye lens of fish. *Acta Parasitol. Pol.* 36: 131–134.
- GRAHAM J.H., RAZ S., HEL-OR H., NEVO E. 2010: Fluctuating asymmetry: methods, theory and applications. *Symmetry* 2: 466–540.
- HOLM S. 1979: A simple sequentially rejective multiple test procedure. *Scand. J. Statistics* 6: 65–70.
- JOHNSON P.T.J., KOPRIVNIKAR J., ORLOFSKE S.A., MELBOURNE B.A., LAFONTE B.E. 2014: Making the right choice: testing the drivers of asymmetric infections within hosts and their consequences for pathology. *Oikos* 123: 875–885.
- KARVONEN A., SEPPALA O. 2008: Eye fluke infection and lens size reduction in fish: a quantitative analysis. *Dis. Aquat. Org.* 80: 21–26.
- KEARN G.C. 1984: The migration of the monogenean *Entobdella soleae* on the surface of its host, *Solea solea*. *Int. J. Parasitol.* 14: 63–69.
- KING J.C., DUBAY S.A., HUSPENI T.C., VANLANEN A.R., GERHOLD R.W. 2010: Parasite infections in nestling red-shouldered hawks (*Buteo lineatus*) in northeast Wisconsin. *J. Parasitol.* 96: 535–540.
- LOCKE S.A., AL-NASIRI F.S., CAFFARA M., DRAGO F., KALBE M., LAPIERRE A.R., MCLAUGHLIN J.D., NIE P., OVERSTREET R.M., SOUZA G.T., TAKEMOTO R.M., MARCOGLIESE D.J. 2015: Diversity, specificity and speciation in larval Diplostomidae (Platyhelminthes: Digenea) in the eyes of freshwater fish, as revealed by DNA barcodes. *Int. J. Parasitol.* 45: 841–855.
- MCDONALD J.H. 2014: *Handbook of Biological Statistics*. Third Edition. Sparky House Publishing, Baltimore, Maryland, 299 pp.
- R CORE TEAM 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, <https://www.R-project.org/>.
- RAU M.E., GORDON D.M., CURTIS M.A. 1979: Bilateral asymmetry of *Diplostomum* infections in the eyes of lake whitefish *Coregonus clupeaformis* (Mitchell) and a computer simulation of the observed metacercarial distribution. *J. Fish Dis.* 2: 291–297.
- RINTAMÄKI-KINNUNEN P.A., KARVONEN A., ANTTILA P., VALTONEN E.T. 2004: *Diplostomum spathaceum* metacercarial infection and colour change in salmonid fish. *Parasitol. Res.* 93: 51–55.
- RODRIGUES A.A., SARAIVA A. 1996: Spatial distribution & seasonality of *Pseudodactylogyrus anguillae* & *P. bini* (Monogenea: Pseudodactylogyridae) on the gills of the European eel *Anguilla anguilla* from Portugal. *Bull. Eur. Ass. Fish Pathol.* 16: 85–88.
- SHIGIN A. A. 1986: [Trematode Fauna of USSR. Genus *Diplostomum*, metacercariae.] Nauka, Moscow, 253 pp. (In Russian.)
- SOKAL R.R., ROHLF F.J. 1995: *Biometry: the Principles and Practice of Statistics in Biological Research*. Third Edition. Freeman, New York, 887 pp.
- STUMBO A.D., POULIN R. 2016: Possible mechanism of host manipulation resulting from a diel behaviour pattern of eye-dwelling parasites? *Parasitology* 143: 1261–1267.
- THIEMANN G.W., WASSERSUG R.J. 2000: Biased distribution of trematode metacercariae in the nephric system of *Rana* tadpoles. *J. Zool.* 252: 534–538.

Received 3 July 2016

Accepted 10 October 2016

Published online 10 November 2016

**Cite this article as:** Wayland M.T., Chubb J.C. 2016: A new R package and web application for detecting bilateral asymmetry in parasitic infections. *Folia Parasitol.* 63: 039.