CSIRO PUBLISHING

Australian Journal of Zoology, 2012, **60**, 388–391
http://dx.doi.org/10.1071/ZO13017

Short communication

Isolation and characterisation of polymorphic microsatellite loci for the neriid fly *Telostylinus angusticollis* using MiSeq sequencing

Anna M. Kopps[A,B], *Russell Bonduriansky*[A], *Anthony S. Gilchrist*[A] *and Angela J. Crean*[A]

[A]Evolution and Ecology Research Centre and School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW 2052, Australia.
[B]Corresponding author. Email: anna.kopps@gmx.com

Abstract. The neriid fly *Telostylinus angusticollis* is being developed as a model organism for experimental research on developmental plasticity, nongenetic inheritance, ageing and sexual selection. No genetic information is currently available for *T. angusticollis*, so to facilitate paternity analyses for experimentation we characterised 17 polymorphic microsatellites for this species based on MiSeq sequences. The loci had, on average, 4.1 alleles and the observed heterozygosity ranged from 0.250 to 0.889. We show that MiSeq can be used successfully to develop microsatellite markers.

Additional keywords: Neriidae, paternal effect.

Received 11 February 2013, accepted 22 April 2013, published online 16 May 2013

Neriidae is a small family of true flies (Diptera) notable for the spectacular sexual dimorphism and male–male sexual competition seen in many species (de Meijere 1911; Eberhard 1998; Preston-Mafham 2001; Bonduriansky 2006). Two neriid species are known endemics in Australia: the tropical *Telostylinus lineolatus*, which breeds on rotting fruit and also occurs on several Pacific islands, and the temperate *T. angusticollis*, which breeds in rotting tree bark and forms mating aggregations on damaged trunks of *Acacia longifolia* and other trees (Bonduriansky 2006, 2007). *T. angusticollis* displays a large range of natural variation in body size and shape, which can be experimentally replicated in the laboratory by manipulating larval diet (Bonduriansky 2007, 2009). Interestingly, despite no obvious forms of paternal investment or care, environmental variation in male body size is inherited by offspring via a non-genetic paternal effect (Bonduriansky and Head 2007). This species is being developed as a model for research on developmental plasticity, parental effects, ageing and sexual coevolution (Eberhard 1998; Bonduriansky 2006, 2007, 2009; Bonduriansky and Head 2007; Kawasaki *et al*. 2008; Adler and Bonduriansky 2011, 2013; Bath *et al*. 2012; Adler *et al*. 2013; Sentinella *et al*. 2013). To facilitate research on this species, we sequenced fragments of the genome and developed microsatellite markers for use in paternity analyses and other genetic analyses.

Neriid flies (*T. angusticollis*) were collected from aggregations on the trunks of *A. longifolia* trees in Fred Hollows Reserve, Sydney, Australia (33.912°S, 151.248°E) and maintained as a large outbred population in the laboratory at the University of New South Wales with *ad libitum* access to food and water. The outbred laboratory population was supplemented annually with new wild-collected flies so that the gene pool of the laboratory population should resemble that of the wild population. In addition, inbred lines were established from single mating pairs in August 2010, and maintained in the laboratory through sibling matings. DNA from two males of the 16th generation of one inbred line was used to develop microsatellites. Flies were frozen at −20°C until extraction. DNA was extracted from whole flies with a Gentra PureGene DNA extraction kit (Qiagen) following manufacturer's instructions, and 1 μg of this extracted DNA was used as a template for sequencing. A library was prepared with a TruSeq DNA sample preparation kit (Illumina). Microsatellites were then identified from data produced on a MiSeq Next-Generation Sequencer (Illumina) targeting 550 bp fragments using the 2x250 bp read length configuration. MiSeq data were assembled using ABySS 1.3.0 (Simpson *et al*. 2009). QDD ver. 2.1 Beta (Meglécz *et al*. 2010) was used to detect microsatellite repeat motifs and primers were designed with Primer 3 (Rozen and Skaletsky 2000). We added a M13 tail to all forward primers. Fluorescently labelled M13 universal primers have lower annealing temperatures than the sequence-specific primers and anneal to the M13 tails when the annealing temperature is decreased during PCR (Schuelke 2000). We used four M13 universal primers, each labelled with a different fluorescent dye, in order to be able to multiplex four loci when screening for polymorphisms. The primers used were the M13 primer detailed in Schuelke (2000) and the three universal primers described by the UC Davis Veterinary Genetics Laboratory (http://www.vgl.ucdavis.edu/dogset/help.jsp#m13). We selected 20 loci with pure microsatellite repeats and screened 20 flies from the outbred laboratory population for polymorphisms at these loci. The 5 μL PCR volume contained 0.5 μL multiplex primer mix (1.5 μM reverse primer, 0.25 μM forward primer and 0.375–1.125 μM M13

Table 1. Characteristics of 17 microsatellites isolated from *Telostylinus angusticollis*

N_A, no. of alleles; H_o, observed heterozygosity; H_e, expected heterozygosity; N, Number of samples; P_{HWE}, P value for deviance from Hardy–Weinberg equilibrium; MP, selected for multiplex

Locus	GenBank No.	Repeat motif	Dye		Primer sequence (5′–3′)[A]	*Telostylinus angusticollis*						MP
						Range (bp)	N_A	N	H_o	H_e	P_{HWE}	
Tangus 1	KC576888	$(ACAT)_{12}$	FAM	F:	TTTCCCAGTCACGACGTTGGCACTGCCAATATCCAAAGC	248–264	3	12	0.250	0.656	0.0065	
				R:	ACCAAATGCCAAATGCAAAT							
Tangus 2	KC576889	$(ACAT)_{11}$	NED	F:	TAAAACGACGGCCAGTGCGGCACGCAACAACACAACTA	342–370	5	18	0.889	0.716	0.1303	Yes
				R:	CAGGCACATACACATGCACA							
Tangus 4	KC576890	$(ACAT)_{12}$	VIC	F:	GCGGATAACAATTTCACACAGGTCGCTAATCGATCGTTGGTT	321–349	4	18	0.389	0.512	0.1916	
				R:	AGCAAGCGCTTACAAAGTCG							
Tangus 5	KC576891	$(ACAT)_{11}$	FAM	F:	TTTCCCAGTCACGACGTTGCATTAGACGCTCCATCCCTG	293–313	4	20	0.300	0.270	1	
				R:	CGACAATGTGAACGAATCCC							
Tangus 6	KC576892	$(ACAT)_{12}$	NED	F:	TAAAACGACGGCCAGTGCTCATACATCCCAGCTCCCAC	250–280	5	19	0.474	0.701	0.0431	Yes
				R:	TCGCATAGGGTGCATGTAGA							
Tangus 7	KC576893	$(ACAT)_{11}$	PET	F:	TGTAAAACGACGGCCAGTTTCCTGTCTTTGCTCTTTGCTT	226–248	2	8	0.250	0.219	1	
				R:	GCGATGCTGTTACGTGTTCA							
Tangus 8	KC576894	$(ACAT)_{12}$	VIC	F:	GCGGATAACAATTTCACACAGGTTCAATACAAGTTCACTTTCGCA	205–221	5	20	0.750	0.739	0.9938	Yes
				R:	CGGGTCGATTGCTGTGTATT							
Tangus 9	KC576895	$(ACAT)_{10}$	FAM	F:	TTTCCCAGTCACGACGTTGTACCTGGACGAGGCATTCAT	212–232	4	20	0.250	0.308	0.2985	Yes
				R:	CGTGCAGCTGTTTAAATTTGG							
Tangus 10	KC576896	$(ACAT)_{10}$	NED	F:	TAAAACGACGGCCAGTGCTTGTCTTTGCAGCCACTTTG	194–212	5	20	0.550	0.720	0.0711	Yes
				R:	GGCATGGTGGGCATTAGATA							
Tangus 11	KC576897	$(ACAT)_{12}$	FAM	F:	TTTCCCAGTCACGACGTTGTGTGATGAAATGAAATGCGG	162–178	4	20	0.550	0.694	0.0628	
				R:	ATTCCTCCTTGGAGCGTTGT							
Tangus 12	KC576898	$(ACTC)_{10}$	NED	F:	TAAAACGACGGCCAGTGCTTGGTGCACTTCTTTGCTCA	165–185	3	20	0.550	0.489	0.6222	
				R:	CTTCAATTCATCCAGTCGCA							
Tangus 13	KC576899	$(ACAT)_{14}$	PET	F:	TGTAAAACGACGGCCAGTTGCATTGCCTTATGATGTATGTG	158–188	4	18	0.889	0.665	0.2143	
				R:	CTTACACGTAGCGCCCATTT							
Tangus 14	KC576900	$(ACAT)_{12}$	VIC	F:	GCGGATAACAATTTCACACAGGTGCCATACATGCACTCATACA	134–164	4	19	0.737	0.694	0.6084	
				R:	AGGAATTCGGTCAAGTGCTG							
Tangus 15	KC576901	$(ACAT)_{13}$	FAM	F:	TTTCCCAGTCACGACGTTGGGCGTCTTCGGTTTCATACA	158–170	4	17	0.765	0.618	0.0837	Yes
				R:	AGTCGACTTGACGTGCGAA							
Tangus 18	KC576902	$(ACAT)_{11}$	VIC	F:	GCGGATAACAATTTCACACAGGGCGGTAAAGCGACACAACA	102–114	3	8	0.750	0.656	0.0617	Yes
				R:	ATGTGCGTGTAGTGGAGGGT							
Tangus 19	KC576903	$(AC)_{25}$	VIC	F:	GCGGATAACAATTTCACACAGGCGGTGTGGAAGGTGGTTAGA	293–337	5	20	0.850	0.791	0.4396	
				R:	CTGAGCCACACACAAACTCG							
Tangus 20	KC576904	$(AC)_{13}$	PET	F:	TGTAAAACGACGGCCAGTCAAACGAATCGCCTCACTG	151–177	6	20	0.800	0.766	0.4419	Yes
				R:	AACGTCGAGCGAATTGATGT							

[A]M13 tag sequences underlined.

primers), 2.5 µL Qiagen Multiplex Kit™ (Qiagen), 1 µL water and 1 µL DNA template. The following touchdown PCR conditions were used: 15 min initial activation at 95°C, followed by cycles of 30 s denaturation at 94°C, 90 s annealing at 62° (5 cycles)/58° (5 cycles)/55° (5 cycles)/50°C (20 cycles) and 90 s extension at 72°C, followed by a final extension step of 10 min at 72°C. Samples were run on an ABI 3730 DNA Analyser and analysed with GeneMapper ver. 3.7 software (both Applied Biosystems) using GS-500 (Liz) as a size standard. The number of alleles per locus and heterozygosity were calculated in GENALEX (Peakall and Smouse 2006) and statistical tests were performed using GENEPOP (Rousset 2008).

The MiSeq run resulted in 16.27 Mega reads that passed the Illumina chastity filter. These reads were assembled into 95 632 contigs. Because QDD ver. 2.1 Beta could not process 95 632 contigs at the same time, we divided the dataset into several subsets. Of 59 122 contigs that were longer than 90 bp, 398 contigs contained at least one microsatellite. From the 20 pure microsatellite loci selected for screening for polymorphism, 17 loci amplified and were polymorphic (Table 1). Loci Tangus 5 and Tangus 6 as well as Tangus 8 and Tangus 19 showed evidence for linkage; all other locus pairs showed no evidence for linkage disequilibrium after sequential Bonferroni correction (Rice 1989). Without Bonferroni correction, the following locus pairs appeared to be linked: Tangus 4 and Tangus 19, Tangus 9 and Tangus 10, Tangus 10 and Tangus 12, Tangus 5 and Tangus 12, Tangus 2 and Tangus 15, Tangus 8 and Tangus 15, Tangus 12 and Tangus 19. None of the loci showed deviation from Hardy–Weinberg equilibrium after sequential Bonferroni correction (Rice 1989) (Table 1). The number of alleles and observed heterozygosity ranged from 2 to 6 and 0.250 to 0.889 respectively. We designed a multiplex containing eight unlinked, polymorphic loci of complementary lengths (Table 1). Testing of the multiplex on five mothers and 16 of their offspring revealed that Tangus 6 was not suitable for the study population because mothers and offspring appeared to be homozygous for different alleles on locus Tangus 6, which indicates the presence of null alleles. The remaining seven loci in the multiplex and the other polymorphic loci identified (Table 1) are likely to shed light on relatedness in future studies.

MiSeq data proved very useful, and was time and cost efficient for identifying microsatellites, as has been shown for other next-generation sequencing platforms (Castoe *et al*. 2012). Of the tested microsatellite loci, 85% amplified and were polymorphic. Our success rate was higher than the 60% success rate (average of 11 species) from 454 shotgun sequencing reported in Gardner *et al*. (2011).

The assembly requirement of MiSeq data may appear as a disadvantage of our approach but it is offset by two main advantages: due to the assembler's requirement for 5-fold coverage of all contigs, the assembly removes most sequencing error, improving the reliability of primer design. In our case, the assembly of one lane of MiSeq data took <24 h (8-core server with 90 GB RAM) and gave quite long contigs (N50 ~17 kb). This was most likely a consequence of the *T. angusticollis* genome being relatively small (estimated ~170 Mbp, comparable to that of *Drosophila melanogaster*). The length of contigs will decrease for larger genomes, yet even for medium-sized genomes a single assembly will yield a large number of 2 kb contigs. The only disadvantage of the assembly approach is the requirement for computing facilities, which are, however, becoming increasingly widely available, including cloud-based options.

Acknowledgements

We thank Jason Koval from the Ramaciotti Centre for Gene Function Analysis (University of New South Wales, Sydney, Australia) for performing the MiSeq run, and Paul Worden from Macquarie University (Sydney, Australia) for running our samples on the ABI 3730. We are also grateful to the editor and an anonymous referee for helpful comments. Funding was provided by the Australian Research Council through a Discovery Early Career Researcher Award to AJC and a Discovery Grant to RB.

References

Adler, M., and Bonduriansky, R. (2011). The dissimilar costs of love and war: age-specific mortality as a function of the operational sex ratio. *Journal of Evolutionary Biology* **24**, 1169–1177. doi:10.1111/j.1420-9101.2011.02250.x

Adler, M., and Bonduriansky, R. (2013). Paternal effects on offspring fitness reflect father's social environment. *Evolutionary Biology*, In press.

Adler, M. I., Cassidy, E. J., Fricke, C., and Bonduriansky, R. (2013). The lifespan–reproduction trade-off under dietary restriction is sex-specific and context-dependent. *Experimental Gerontology*, In press. doi:10.1016/j.exger.2013.03.007

Bath, E., Tatarnic, N., and Bonduriansky, R. (2012). Asymmetric reproductive isolation and interference in neriid flies: the roles of genital morphology and behaviour. *Animal Behaviour* **84**, 1331–1339. doi:10.1016/j.anbehav.2012.08.025

Bonduriansky, R. (2006). Convergent evolution of sexual shape dimorphism in Diptera. *Journal of Morphology* **267**, 602–611. doi:10.1002/jmor.10426

Bonduriansky, R. (2007). The evolution of condition-dependent sexual dimorphism. *American Naturalist* **169**, 9–19. doi:10.1086/510214

Bonduriansky, R. (2009). Condition dependence of developmental stability in the sexually dimorphic fly *Telostylinus angusticollis* (Diptera: Neriidae). *Journal of Evolutionary Biology* **22**, 861–872. doi:10.1111/j.1420-9101.2009.01686.x

Bonduriansky, R., and Head, M. (2007). Maternal and paternal condition effects on offspring phenotype in *Telostylinus angusticollis* (Diptera: Neriidae). *Journal of Evolutionary Biology* **20**, 2379–2388. doi:10.1111/j.1420-9101.2007.01419.x

Castoe, T. A., Poole, A. W., de Koning, A. P. J., Jones, K. L., Tomback, D. F., Oyler-McCance, S. J., Fike, J. A., Lance, S. L., Streicher, J. W., Smith, E. N., and Pollock, D. D. (2012). Rapid microsatellite identification from illumina paired-end genomic sequencing in two birds and a snake. *PLoS ONE* **7**(2), e30953.

de Meijere, J. (1911). Studien über südostasiatische Dipteren VI. *Tijdschrift voor Entomologie* **54**, 258–432.

Eberhard, W. G. (1998). Reproductive behavior of *Glyphidops flavifrons* and *Nerius plurivitatus* (Diptera, Neriidae). *Journal of the Kansas Entomological Society* **71**, 89–107.

Gardner, M. G., Fitch, A. J., Bertozzi, T., and Lowe, A. J. (2011). Rise of the machines – recommendations for ecologists when using next generation sequencing for microsatellite development. *Molecular Ecology Resources* **11**, 1093–1101. doi:10.1111/j.1755-0998.2011.03037.x

Kawasaki, N., Brassil, C., Brooks, R., and Bonduriansky, R. (2008). Environmental effects on the expression of lifespan and aging: An extreme contrast between wild and captive cohorts of Telostylinus angusticollis (Diptera: Neriidae). *American Naturalist* **172**, 346–357. doi:10.1086/589519

Meglécz, E., Costedoat, C., Dubut, V., Gilles, A., Malausa, T., Pech, N., and Martin, J.-F. (2010). QDD: a user-friendly program to select microsatellite markers and design primers from large sequencing projects. *Bioinformatics* **26**, 403–404. doi:10.1093/bioinformatics/btp670

Peakall, R. O. D., and Smouse, P. E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**, 288–295. doi:10.1111/j.1471-8286.2005.01155.x

Preston-Mafham, K. (2001). Resource defence mating system in two flies from Sulawesi: *Gymnonerius fuscus* Wiedemann and *Telostylinus* sp. near *duplicatus* Wiedemann (Diptera: Neriidae). *Journal of Natural History* **35**, 149–156. doi:10.1080/002229301447916

Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution* **43**, 223–225. doi:10.2307/2409177

Rousset, F. (2008). Genepop'007: a complete re-implementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources* **8**, 103–106. doi:10.1111/j.1471-8286.2007.01931.x

Rozen, S. and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. In 'Bioinformatics Methods and Protocols: Methods in Molecular Biology'. (Eds S. Krawetz and S. Misener.) pp. 365–386. (Humana Press: Totowa, NJ.)

Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* **18**, 233–234. doi:10.1038/72708

Sentinella, A. T., Crean, A. J., and Bonduriansky, R. (2013). Dietary protein mediates a trade-off between larval survival and the development of male secondary sexual traits. *Functional Ecology*, In press.

Simpson, J. T., Wong, K., Jackman, S. D., Schein, J. E., Jones, S. J. M., and Birol, İ. (2009). ABySS: a parallel assembler for short read sequence data. *Genome Research* **19**, 1117–1123. doi:10.1101/gr.089532.108

Handling Editor: Steven Cooper