This is the author's final, peer-reviewed manuscript as accepted for publication. The publisher-formatted version may be available through the publisher's web site or your institution's library.

Identification of 12 EST-derived SSR markers in *Lumbricus rubellus*

Zachary T. Lounsberry, Mark C. Ungerer, Bruce A. Snyder

How to cite this manuscript

If you make reference to this version of the manuscript, use the following information:

Lounsberry, Z. T., Ungerer, M. C., & Snyder, B. A. (2013). Identification of 12 EST-derived SSR markers in *Lumbricus rubellus*. Retrieved from http://krex.ksu.edu

Published Version Information

Citation: Lounsberry, Z. T., Ungerer, M. C., & Snyder, B. A. (2013). Identification of 12 EST-derived SSR markers in Lumbricus rubellus. Pedobiologia - International Journal of Soil Biology, 56(4-6), 191-193.

Copyright: © 2013 Elsevier GmbH

Digital Object Identifier (DOI): doi:10.1016/j.pedobi.2013.07.002

Publisher's Link: http://www.sciencedirect.com/science/article/pii/S0031405613000383

This item was retrieved from the K-State Research Exchange (K-REx), the institutional repository of Kansas State University. K-REx is available at http://krex.ksu.edu

Title:

Identification of 12 EST-derived SSR markers in *Lumbricus rubellus*

Authors:

Zachary T. Lounsberry [a,b], Mark C. Ungerer [a], Bruce A. Snyder [a*]

Affiliations:

[a] Division of Biology, Kansas State University, Manhattan, KS, USA

[b] Veterinary Genetics Laboratory, University of California, Davis, CA, USA

Running Title:

EST-derived SSR markers in *L. rubellus*

[*] Corresponding author at: Kansas State University, Division of Biology, 116 Ackert Hall, Manhattan, KS 66506-4901. Tel.: +1 785 532 2430; fax: +1 785 532 6653. E-mail address: bruceasnyder@gmail.com (B.A. Snyder).

Summary

While many species of earthworms are globally distributed, very little is known about the genetic population dynamics of this diverse group. We present the characterization of novel simple sequence repeat (SSR) markers, including primer information, number of alleles, repeat motif, and approximate size ranges, to be used in population genetic analyses of the earthworm *Lumbricus rubellus* Hoffmeister 1843. Specifically, we designed and characterized 12 novel, polymorphic markers derived from published expressed sequence tags (EST) for amplification in *L. rubellus*. The mean number of alleles per locus was 6.25 ± 1.91, indicating these markers will be sufficiently polymorphic for population genetic studies of this species.

Keywords: earthworm; invasion ecology; population genetics; simple sequence repeats

Earthworm introductions are a widespread phenomenon, resulting in a global distribution of many earthworm species (Hendrix et al. 2008). Ecological effects of a few earthworms are well known but their population-level genetic characteristics (e.g., migration and dispersal patterns, relatedness among and within populations) are poorly understood (Bohlen et al. 2004; Frelich et al. 2006; Loss et al. 2012). To date, few studies have focused on population genetic markers in earthworms (Dupont 2009; Velavan et al. 2009; Novo et al. 2010). Here, we present a characterization of 12 polymorphic simple sequence repeat (SSR) markers developed from published expressed sequence tag (EST) libraries and amplified in individuals from three isolated sites of *Lumbricus rubellus* Hoffmeister 1843 in Kansas.

Native to Europe, *L. rubellus* is a widespread earthworm species with a global geographic distribution (Gates 1972). As a diploid species that is incapable of self-fertilization or parthenogenesis, *L. rubellus* is a good model for understanding otherwise complex earthworm population genetics. Despite a pervasive presence in the literature, only a single study has examined SSR markers in *L. rubellus*, having employed a genome-wide SSR enrichment protocol (Hammond et al. 1998, Harper et al. 2006). The goal of the current study was to develop EST-SSRs for this species to expand upon existing resources for understanding population dynamics in *L. rubellus*. EST-SSRs are located in coding regions and thus completely linked with genes potentially serving important functions in nature (Li et al. 2004). In addition, EST-SSRs typically are more conserved at the sequence level and thus more likely to amplify across genetically divergent populations or species (Arnold et al. 2002).

Citizen scientists collected earthworm specimens as part of a state-wide earthworm sampling initiative, Earthworms Across Kansas, which aimed to collect earthworm distribution information while engaging students in scientific research (Snyder and Callaham, unpublished results). We sent primary and secondary school teachers collecting kits, including 70% ethanol in vials, and instructed them to collect earthworms with their students in suburban habitats using methods including digging and hand sorting. Collectors shipped specimens with collection information to Kansas State

University where we took tissue samples from adult specimens, preserved them in 95% ethanol, stored them at -20°C, and transferred them to long-term storage at -80°C. We preserved the remainder of each specimen for species determination by author BAS using Sims and Gerard (1999). While this initiative produced over 3000 specimens in total, 19 adult *L. rubellus* were obtained from three sites: near a pond on the campus of Louisburg Middle School in Miami Co. (Site 1, n=3), on the campus of Olathe North High School in Johnson Co. (Site 2, n=3), and on the campus of Blue Valley West High School in Johnson Co. (Site 3, n=13). Mean distance between these sites is 20.6 km. These specimens are the only *L. rubellus* known from Kansas, and 14 specimens (Supplemental Figure 1) yielded sufficient quality DNA for molecular analyses (i.e., samples amplified at eight or more loci). While sample sizes for two of our three sites were small, sampling multiple geographic locations and habitat types allowed us to test these EST-SSR markers on a larger pool of genetic diversity.

We developed EST-SSR loci by downloading *L. rubellus* EST sequences available on Genbank and submitting them to an online tandem repeat finder (http://sgdp.iop.kcl.ac.uk/nikammar/repeatfinder.html). We tested a total of 22 candidate SSR markers in *L. rubellus* (n = 14), two congeneric individuals of *Lumbricus terrestris*, and two confamilial individuals of *Bimastos parvus*. Loci were amplified via the polymerase chain reaction (PCR) according to the following chemistry: 10-50 ng genomic DNA, 2 µL 1 x reaction buffer (PROMEGA), 2mM $MgCl_2$, 0.2mM of each dNTP, 0.5 µM of reverse primer (Table 1), 0.3 µM of M13-tagged forward primer, 0.5 µM of M-13 universal primers labeled with a fluorescent dye attached to the 5' end (Schuelke 2000), and 0.1 units of GoTaq polymerase (PROMEGA). We used the program Genemarker (v1.95; Soft Genetics LLC, State College, PA) to visualize amplicons. PCR thermal cycling conditions consisted of an initial denaturing step of 4 min at 94°C, followed by 30 cycles of 30 s at 94°C, 30 s at Tm°C (see Table 1), and 30 s at 72°C, followed by 10 cycles of 30 s at 94°C, 30 s at 54°C, and 30 s at 72°C, with a final extension of 10 min at 72°C. Of 22 markers initially screened, 12 showed consistent amplification and polymorphism. Characteristics of these 12

loci are presented in Table 1. None of the loci tested in this study amplified reliably in *L. terrestris* or *B. parvus*; PCR amplicons in these tests of cross species utility were absent or weak and frequently dropped out in replicate PCRs.

Principal coordinate analysis (PCoA) implemented in Genalex 9 (Peakall and Smouse 2012) of genetic distance between individuals revealed no apparent clustering of individuals by sampling site (Supplemental Figure 1). Consistent with this result, there was no evidence for site-specific alleles at any of our three geographically isolated sites. For the entire sample, we tested departures from Hardy Weinberg equilibrium for each locus using a 1,000,000-step Markov Chain Monte Carlo simulation in Program Arlequin v3.0 (Excoffier et al. 2005). We did not conduct these tests for individual sites because of extremely small sample sizes at two of three sites ($n \leq 3$) and given the lack of unique, site-specific genetic signatures (Supplemental Figure 1). Several loci were found to deviate significantly from Hardy Weinberg equilibrium (Table 1); this result should be viewed with caution, however, given the small sample size and pooling of individuals from multiple sites. Additional tests are warranted utilizing a larger sample of individuals from a single population.

All 12 markers presented in this study are polymorphic and show promise of being useful for population-level genetic analyses of *L. rubellus*. Logistic constraints of the present study, including the limited number of available *L. rubellus* samples in Kansas, curtailed further population-level analyses of population structure or migration. However, the novel markers presented here, coupled with the markers from Harper et al. (2006), provide a framework for more rigorous population genetic analyses of *L. rubellus.* Moreover, because these markers were derived from expressed sequence tag libraries and thus show complete linkage with functional genes, it is possible they will be useful in studies focusing on the underlying genetics of habitat adaptability. Lastly, this study emphasizes the efficacy of mining published EST libraries as an extremely cost-effective means of SSR development.

TABLE 1: Characterization of twelve novel SSR markers derived from published EST library. Loci marked with * indicate deviation from Hardy-Weinberg equilibrium after sequential Bonferroni correction. n: number of earthworm samples that amplified for the locus; Tm: melting temperature of the primer pair; N_A: number of total alleles at that locus

Locus name	Primer sequence (5'−3')	Genbank accession	Repeat motif	n	Tm (^{o}C)	N_A
Lr1*	F: M13-GGAGCACATGTCACTTCGAT R: TTGATTGCCACCATCGTC	EL517770	$(ACA)_{13}$	14	60	6
Lr4*	F: M13-TGTTGAGCAGGAGATGTCG R: CGATGTTTGCTGGTTTTCC	EH168054	$(TGT)_{9}$	13	60	8
Lr5	F: M13-TGGCTTGAGGGCTTAAGTG R: ACAGGCTGCTGACACAACA	EH166762	$(TTGA)_{9}$	10	60	8
Lr6	F: M13-GATGTGCCGAACCGTAATC R: AAAAGGGGCACAGAAGAGA	DR697298	$(TA)_{26}$	12	55	9
Lr10	F: M13-CGATGACCTTCGATGACCT R: ATGTTGCCGATCTGCTGTT	DR077447	$(CAA)_{8}$	13	60	4
Lr12*	F: M13-GGCAAGCTGATTACCGACT R: ATCGGACGAGAGATGCTTG	DR009495	$(AGC)_{13}$	12	55	5
Lr17	F: M13-ATCCGACGTTTCATCTGCT R: TGTATCCACGAAACCATGC	CV462559	$(GT)_{20}$	8	60	7
Lr19	F: M13-AGAACCCATCGACCAATCA R: TCTACAGCTGCTGCATTCG	CV462279	$(AAT)_{8}$	7	55	2
Lr21*	F: M13-CCGGTTGAAGATGCTGAAG R: ACTTCTTCTTTGCGCATCC	CO869573	$(AGC)_{10}$	11	60	7
Lr23*	F: M13-TGGCTCAGCAATGAGTCAC R: CGAGAGGCAACTTTCTGGT	CO408577	$(TG)_{12}$	14	60	7
Lr24*	F: M13-GCCACCACGTTCAGCTAAT R: TGCATCAACGTCAGTGAGAA	CO378033	$(TTA)_{13}$	8	60	6
Lr30*	F: M13-GTTGCGATTGTAGCTGCTG R: TGTGCATGTGCAGACAACA	CO046624	$(TGA)_{11}$	12	60	6

Acknowledgments

We would like to thank H. Tetreault for assistance with laboratory procedures. This project was funded by Kansas State University's Center for Engagement and Community Development.

References

Arnold, C., Rossetto, M., McNally, J., Henry, R.J., 2002. The application of SSRs characterized for grape (*Vitis vinifera*) to conservation studies in Vitaceae. American Journal of Botany 89, 22–28.

Bohlen, P.J., Scheu, S., Hale, C.M., McLean, M.A., Migge, S., Groffman, P.M., Parkinson, D., 2004. Non-native invasive earthworms as agents of change in northern temperate forests. Frontiers in Ecology and the Environment 2, 427–435.

Dupont L., 2009. Perspectives on the application of molecular genetics to earthworm ecology. Pedobiologia 52, 191–205.

Excoffier, L., Laval, G., Schneider, S., 2005 Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1, 47–50.

Frelich, L.E., Hale, C.M., Scheu, S., Holdsworth, A.R., Heneghan, L., Bohlen, P.J., Reich, P.B., 2006. Earthworm invasion into previously earthworms-free temperate and boreal forests. Biological Invasions 8, 1235–1245.

Gates, G.E., 1972. Burmese earthworms: an introduction to the systematics and biology of megadrile oligochaetes with special reference to Southeast Asia. Transactions of the American Philosophical Society New Series 62, 1–326.

Hammond, R.L., Saccheri, I.J., Coifi, C., Coote, T., Funk, S.M., McMillan, W.O., Bayes, M.K., Taylor, E., and Bruford, M.W., 1998. Isolation of microsatellite markers in animals. In: Karp, A., Isaac, P.G., and Ingram, D.S. (Eds.), Molecular Tools for Screening Biodiversity: Plants and Animals. Chapman & Hall, London pp. 195–201.

Harper, G.L., Cesarini, S., Casey, S.P., Morgan, A.J., Kille, P., Bruford, M.W., 2006. Microsatellite markers for the earthworm *Lumbricus rubellus*. Molecular Ecology Notes 6, 325–327.

Hendrix, P.F., Callaham Jr, M.A., Drake, J.M., Huang, C-Y., James, S.W., Snyder, B.A., Zhang, W., 2008. Pandora's box contained bait: the global problem of introduced earthworms. Annual Review of Ecology, Evolution, and Systematics 39, 593–613.

Li, Y., Korol, A.B., Fahima, T., Nevo E., 2004, Microsatellites within genes: structure, function, and evolution. Molecular Biology and Evolution 21, 991–1007.

Loss, S.R., Niemi, G.J., Blair, R.B., 2012. Invasions of non-native earthworms related to population declines of ground-nesting songbirds across a regional extent in northern hardwood forests of North America. Landscape Ecology 27, 683–696.

Novo, M., Almodóvar, A., Fernández, R.M., Gutiérrez, M., Cosín, D.J.D, 2010. Mate choice of an endogeic earthworm revealed by microsatellite markers. Pedobiologia 53, 375–379.

Peakall, R., Smouse, P.E., 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28, 2537–2539.

Schuelke, M., 2000. An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology 18, 233–234.

Sims, R.W., Gerard, B.M., 1999. Earthworms. Notes for the Identification of British Species. Field Studies Council, Shrewsbury.

Velavan, T.P., Weller, S., Schulenburg, H., Michiels, N.K., 2009. High genetic diversity and heterogeneous parasite load in the earthworm *Lumbricus terrestris* on a German meadow. Soil Biology & Biochemistry 41, 1591–1595

Supplemental Figure 1. Principal coordinate analysis explaining 46.81% of variation in genetic distance between individuals collected from three isolated sites in Kansas. Percentage of variation explained by each coordinate is given in parentheses on axes. Sample sizes are as follows: Site 1, n=2; Site 2, n=3; Site 3, n=9.