Isolation and characterisation of the first microsatellite markers for the European stag beetle, *Lucanus cervus* (Coleoptera: Lucanidae)

Niall J. McKeown[1], Deborah J. Harvey[2], Amy J.E. Healey[1], Ilze Skujina[1], Karen Cox[3], Alan C. Gange[2] and Paul W. Shaw[1]

[1] Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth, UK; e-mails: njm2@aber.ac.uk, amh5@aber.ac.uk, ils3@aber.ac.uk, pws3@aber.ac.uk

[2] Royal Holloway University of London, Egham, Surrey, UK; e-mails: D.Harvey@rhul.ac.uk, a.gange@rhul.ac.uk

[3] Research Institute for Nature and Forest, Gaverstraat 4, B-9500 Geraardsbergen, Belgium; e-mail: karen.cox@inbo.be

Key words. Coleoptera, Lucanidae, genetic, simple sequence repeats, conservation, biodiversity, ecology, saproxylic

Abstract. The European stag beetle, *Lucanus cervus*, is recognised as a flagship species for biodiversity conservation. Although the species is widely distributed across Europe declines have led to it being granted protected or endangered status in a number of countries and regarded as "near threatened" by the IUCN. The integration of genetic approaches into conservation efforts is urgently needed but has been impeded to date by the lack of appropriate genetic markers. To provide such a resource the development of the first microsatellite loci for stag beetle is described. Loci were identified using two methods (i) enriched library cloning (ELC) and (ii) Restriction enzyme Associated DNA sequencing (RAD-Seq). Inefficient microsatellite detection using the ELC method suggests that RAD-Seq, or other Next Generation Sequencing based methods, may ultimately be more cost effective for obtaining informative suites of markers for this and other coleopteran species. 18 loci were characterised by genotyping 42 UK specimens collected as prey leftover/roadkill. All loci produced unambiguous genotypes and were polymorphic. Though preliminary, estimates of genetic variability suggest UK populations may be genetically depauperate. The microsatellite loci represent a suite of genetic markers that can be applied to non-invasive population monitoring and numerous other areas of *Lucanus* conservation and evolutionary research.

INTRODUCTION

Saproxylic invertebrates, defined as organisms that for at least some period of their life cycle are dependent on fungal decay of wood in living or dead trees, or on other saproxylic organisms (Alexander, 2008), are recognised as one of the most threatened components of European fauna (Nieto & Alexander, 2010). The European stag beetle, *Lucanus cervus*, is undoubtedly the most well-known saproxylic beetle and as such is recognised as a focal species for saproxylic habitat conservation and more broadly as a flagship species for European biodiversity conservation (Thomaes et al., 2008; Chiari et al., 2014). Although stag beetles are widely distributed across Europe (Harvey et al., 2011a), declining numbers (Harvey et al., 2011b) mean that the species is of conservation concern and has been granted protected or endangered status in a number of countries and listed as "near threatened" in the IUCN red list of threatened species (Nieto & Alexander, 2010). A major hindrance to the efficacy of conservation efforts for the stag beetle is the established difficulty of inferring population demographics from available ecological methods (Camparano et al., 2011; Harvey et al., 2011b; Chiari et al., 2013, 2014). This highlights the need to integrate conservation genetic approaches. Genetic based studies have already provided considerable insight into species relationships within Lucanidae (Cox et al., 2013; Solano et al., 2016). However, the genetic markers used in those studies lack the necessary resolution to elucidate finer-scale intraspecific population processes that may significantly influence population viability. In this regard microsatellites remain popular markers for both cost effective and informative intraspecific conservation genetic inference. This research therefore reports on the development of the first microsatellite loci for *L. cervus* as a resource for future conservation efforts. The characterisation of these markers also provides a preliminary insight into nuclear genetic variability among UK samples.

MATERIALS AND METHODS

DNA was extracted from a number of stag beetle specimens (n = 42), collected around Surrey (UK), using a standard phenol chloroform method. In all cases individual samples were in the form of parts of a beetle collected as prey leftover, or road kill. Microsatellite isolation was initially performed by traditional means of microsatellite enriched genomic library cloning (ELC) following McKeown & Shaw (2008). Briefly, genomic DNA was digested with the restriction enzyme *Rsa*I (New England Biolabs, Ipswich, MA, USA) and the blunt ended fragments ligated to double-stranded SuperSNX linkers. Enrichment was then performed by selective hybridization of biotin-labelled repeat motif oligonucleotide probes [$(TG)_{12}$, $(GA)_{12}$, $(AAAT)_8$, $(AACT)_8$, $(AAGT)_8$, $(ACAT)_8$, $(AGAT)_8$] with hybridised complexes captured using

streptavidin-coated magnetic beads (DYNAL, Invitrogen, Carlsbad, CA, USA) and unbound DNA removed by a series of washes. DNA fragments were then eluted from the magnetic beads and amplified by polymerase chain reaction (PCR) using the Super-SNX24F oligonucleotide. The PCR products were cloned using the TOPO-TA cloning kit (Invitrogen) and recombinant colonies identified by disruption of *β*-galactosidase activity. Cloned inserts were individually PCR amplified and sequenced using the M13 forward and reverse primers. As this method yielded a low number of usable microsatellites (described later) a Restriction enzyme Associated DNA sequencing (RAD-Seq) library was constructed for 16 randomly chosen specimens following Baird et al. (2008) to identify microsatellite containing sequences. Sequencing adaptors and individual barcodes were ligated to *Sbf* I digested genomic DNA and fragments from individuals jointly sequenced on an Illumina MiSeq. After sequencing, reads were de-multiplexed, required as other species were also present in the RAD-Seq library, and barcodes/adaptors removed using CLC genomics workbench (Qiagen, Hilden, Germany) resulting in sequence fragments of 70–100 base pairs.

Sequences obtained from both methods (ELC and RAD-Seq) were scanned for microsatellite motifs using MSATCOM-MANDER (Faircloth, 2008). PCR primers were designed using PRIMER 3 (Rozen & Skaletsky, 1998). Forward primers were tagged on the 5' end with the universal M13 sequence (5'-TG-TAAAACGACGGCCAGT-3') to allow fluorescent labelling of PCR products using a 3-primer protocol detailed below. Additionally, a 5' GTTTCTT "pigtail" was added to the reverse primers derived from the RAD-Seq library to improve consistency in amplicon size (Brownstein et al., 1996). PCR conditions were initially optimised on a subset of individuals (n = 8) after which 18 loci (Table 1) were more fully characterised by genotyping the 42 individuals. Each locus was individually amplified in a 10 μl reaction containing ~50 ng template DNA, 5 μl Biomix (Bioline, London, UK), 0.025 UM forward (M13 tagged) primer, 0.25 UM reverse (untagged) primer, and 0.25 UM universal M13 primer labelled with a FAM, VIC, NED, or PET fluorescent dye. The same PCR thermoprofile consisting of an initial denaturation step (95°C for 3 min) followed by 35 cycles of 95°C for 30 s, 50°C for 30 s and 72°C for 30 s, and a final cool down step (4°C for 60 s), was used for all loci. Amplicons were separated on

Table 1. Primer sequences and characteristics of 18 microsatellite loci developed for *Lucanus cervus*, including repeat motif in the sequence/clone used to develop the locus (see GenBank accession). Allele numbers (N_a) and size range, observed (H_O) and expected (H_E) heterozygosities and P values for tests of Hardy-Weinberg equilibrium (P_{HW}) calculated from analysis of 42 individuals.

Locus/GenBank Accession	Primer sequence (5'–3')	Repeat motif in sequenced allele	N_a	Size range (bp)	H_O	H_E	P_{HW}
Lcerv-1 MK050797	F: TGCAGATTATGAACACGTG R: TGGTCCCATTCGAACCAC	$(GT)_{10}$	5	101–111	0.59	0.59	0.41
Lcerv-3 MK050798	F: TGCAGTTTCATTTATAAATGTG R: TGTGAAAGAGCCAAGATACACG	$(CA)_9$	4	88–98	0.33	0.41	0.11
Lcerv-4 MK050799	F: TGCAGCTTTTTATTATTTCTTGC R: AACCACCGTCGTGCAGTTAG	$(CA)_8$	2	102–104	0.38	0.46	0.28
Lcerv-6 MK050800	F: TGCAGTCTAATCTGAATTGAG R: TGTCGCCTGAAAATAACTTGTC	$(GA)_{11}$	3	108–114	0.21	0.20	0.89
Lcerv-7 MK050801	F: TGCAGGGATGCAAAAACG R: CGCTTGTTTATAACTCAATTTCC	$(TA)_{12}$	6	92–104	0.50	0.45	0.85
Lcerv-8 MK050802	F: CGATGGTAAAATGCCGTTTC R: ACATGTGCCCCACTTACTCG	$(GT)_8$	6	95–105	0.67	0.73	0.35
Lcerv-9 MK050803	F: CTCCACCTTCAATAATTCTCC R: GGCGGATCGATGAAGTAAAAC	$(CT)_{12}$	5	110–118	0.33	0.63	<u><0.01</u>
Lcerv-16 MK050804	F: TGCAGCAGCAGCAACATG R: TGAAGTACGAGGTTGAAGGAAG	$(GT)_{16}$	4	106–130	0.50	0.46	0.84
Lcerv-17 MK050805	F: CTCCTAATGGAGCGTGAACC R: TTGACCCGAAGAAGTGAGTG	$(GT)_9$	4	208–214	0.42	0.66	0.01
Lcerv-20 MK050806	F: GACCGAACATGCTCTCGAAC R: GGGTTGTTCTCCACCGTACC	$(GT)_{12}$	5	98–108	0.32	0.61	<u><0.01</u>
Lcerv-21 MK050807	F: CGAGCCATTTAAATTACAAAACG R: TTCAATATTATTTCAGGGATGTCG	$(AG)_{12}$	2	92–96	0.33	0.46	0.08
Lcerv-25 MK050808	F: TGCAGCAATTTCAAGTCC R: ATCCCGAATTCCCGGTCT	$(GT)_{11}$	7	95–113	0.69	0.74	<u><0.01</u>
Lcerv-26 MK050809	F: CTGCAGTTTGAGGAGTGTGT R: CAACATCTGTACAGGTACACTCA	$(GT)_{22}$	6	95–113	0.24	0.73	<u><0.01</u>
Lcerv-28 MK050810	F: TGCAGTGGAAAATTACAGG R: TTGACGTCAGAGAATTTTACTTGC	$(GAA)_8$	4	93–108	0.25	0.28	0.95
Lcerv-29 MK050811	F: TGCAGGTAACCGACGATGA R: TTTGCACAAGTAGCCTTTCG	$(AAT)_{13}$	4	91–109	0.31	0.38	<u>0.01</u>
Lcerv-30 MK050812	F: GCAGTAGGATCGGAATGAATG R: GCCGTATCGAAAGTGCTGAT	$(CAT)_9$	2	92–95	0.06	0.35	<u><0.01</u>
Lcerv-31 MK050813	F: TGCAGTGTATGACGTTTGG R: GCCGATTTAGAAAGCAGCAC	$(TTA)_8$	3	94–100	0.51	0.52	0.48
Lcerv-36 MK050814	F: CTATTGACAAAATTTCATCTAATCTA R: TGAAGATCAGCAGACCGGATA	$(CTAT)_5$	3	164–176	0.62	0.75	<0.01

an AB3500 genetic analyser (Applied Biosystems, Warrington, UK) with alleles genotyped using GENEMAPPER (Applied Biosystems) software. The software GENALEX 6.5 (Peakall & Smouse, 2006) was used to calculate standard diversity indices and perform tests of linkage disequilibrium and Hardy Weinberg Equilibrium (HWE) using default parameters.

RESULTS AND DISCUSSION

The 18 microsatellite loci were all polymorphic with allele sizes differing in expected multiples of their repeated motifs. Of the 18 loci only 2 (Lcerv-17 and Lcerv-36) were obtained using the traditional ELC method. Overall the ELC method (i) reported a low proportion of microsatellite containing sequences (6 out of 100 sequences) and (ii) yielded candidate loci for which PCR produced multiple banding patterns that could not be genotyped. In contrast, the RAD-Seq method identified a large number of microsatellite containing sequences with all candidate loci that were tested producing clear PCR amplicons. Similar difficulties in the isolation of microsatellite loci using traditional (i.e. non-NGS) enrichment and cloning protocols have been reported for some other coleopteran species (Salle et al., 2007; but see Drag et al., 2013) as well as Lepidoptera (Zhang, 2004). The data therefore suggest that NGS based methods may be more efficient and ultimately more cost effective means to develop informative suits of microsatellites for coleopteran species.

Diversity indices for each locus and their observed allele size ranges are presented in Table 1. The mean number of alleles per locus was 4 and both observed and expected heterozygosities were low. An important consideration here is that the genotyped samples were from a single area in Britain and other British areas may harbour more genetically variable individuals/populations. Furthermore, British populations may have experienced historical founder effects and loss of genetic variability associated with postglacial colonisation processes and recurrent isolation from European mainland populations as reported for other species (Rossiter et al., 2000). Therefore, levels of variability at these loci may be higher among mainland European samples. More extensive spatial/temporal genetic surveys will be needed to assess historical and recurrent influences on genetic diversity among British and mainland European populations. While there is the possibility of ascertainment bias here, a number of other saproxylic species such as the bark beetle (*Ips typographus*) (Salle et al., 2007), the Asian longhorned beetle (*Anoplohora glabripennis*) (Carter et al., 2009) and the Rosalia Longicorn (*Rosalia alpine*) (Drag et al., 2013, 2015) have revealed similarly low levels of intrasample microsatellite variability over wider geographical ranges suggesting a general pattern of low genetic variability among saproxylics.

No significant linkage disequilibrium was detected between any locus pair indicating that they are segregating independently and can each be considered independent markers. While most loci conformed to HWE, a number exhibited significant deviations due to heterozygote deficits (Table 1). Heterozygote deficits have been reported for other invertebrates (Shaw et al., 2010) and can result from technical artefacts such as null alleles and/or biological processes such as selection, inbreeding, and Wahlund effects (i.e. sampling of individuals from genetically distinct groups). Data from paternity studies of British samples report no evidence of null alleles or other technical artefacts at these loci (N.J. McKeown, unpubl. data). Furthermore, while selection effects have been detected at microsatellite loci (McKeown et al., 2017), it must be considered unlikely that selection could be driving the heterozygote deficits at multiple loci observed here. Therefore inbreeding and/or Wahlund effects seem the most likely explanation which ergo points to underlying spatial/temporal heterogeneity.

Due to its role as a flagship and focal species, the monitoring of *L. cervus* deserves high priority for biodiversity conservation in Europe, not least because its conservation may positively impact other taxa within the same habitats (Chiari et al., 2014). However, as the species is holometabolous and adults exhibit differentiation in a number of traits between males and females that may bias inferences derived from capture methods, the species poses a number of challenges to the accurate estimation of population parameters (Harvey et al., 2011b; Chiari et al., 2014). Microsatellite analysis can be used to estimate effective and census population sizes, as well as identifying populations with genetic characteristics and/or reduced probability of persistence that require prioritisation (Allendorf, 2017). Spatial patterns of genetic structure can also be used to identify patterns and processes (e.g. behaviour) of population connectivity/isolation and thus define management units while also providing a neutral genetic background against which phenotypic patterns can be interpreted and inform on roles of environmental plasticity and/or local adaptation, key factors underpinning species/population resilience to environmental change. The loci developed here thus represent tools that can be directly applied to numerous aspects of conservation research. In this study complete genotypes were obtained for all samples collected as prey leftover, or road kill. This highlights the utility of the loci to study non-preserved and potentially degraded samples and suggests they may also be applicable to museum specimens. As body fragments are often collected as part of walk transects (Campanaro et al., 2010; Campanaro & Bardiani, 2012) and road surveys (Harvey et al., 2011b) that are commonly employed monitoring methods, genetic analysis of such samples offers a means by which ecological and genetic data collection could be combined operationally for non-invasive population monitoring. Finally, if the loci produce homologous products in closely related taxa/sister species they may help clarify higher taxonomic relationships by providing a nuclear phylogenetic perspective (Cox et al., 2013).

REFERENCES

ALEXANDER K.N.A. 2008: Tree biology and saproxylic Coleoptera: issues of definitions and conservation language. — *Rev. Ecol. Terre Vie (Suppl.)* **63**: 9–13.

ALLENDORF F.W. 2017: Genetics and the conservation of natural populations: allozymes to genomes. — *Mol. Ecol.* **26**: 420–430.

BAIRD N.A., ETTER P.D., ATWOOD T.S., CURREY M.C., SHIVER A.L., LEWIS Z.A., SELKER E.U., CRESKO W.A. & JOHNSON E.A. 2008: Rapid SNP discovery and genetic mapping using sequenced RAD markers. — *PLoS ONE* **3**(10): e3376, 7 pp.

BROWNSTEIN M.J., CARPTEN J.D. & SMITH J.R. 1996: Modulation of non-templated nucleotide addition by taq DNA polymerase: Primer modifications that facilitate genotyping. — *Biotechniques* **20**: 1004–1006.

CAMPANARO A. & BARDIANI M. 2012: Walk transects for monitoring of *Lucanus cervus* in an Italian lowland forest. Saproxylic beetles in Europe: monitoring, biology and conservation. — *Stud. Forest. Sloven.* **137**: 17–22.

CAMPANARO A., TONI I., HARDERSEN S. & GRASSO D.A. 2011: Monitoring of *Lucanus cervus* by means of remains of predation (Coleoptera: Lucanidae). — *Entomol. Gener.* **33**: 79–89.

CARTER M., SMITH M. & HARRISON R. 2009: Genetic analyses of the Asian longhorned beetle (Coleoptera, Cerambycidae, *Anoplophora glabripennis*), in North America, Europe and Asia. — *Biol. Invas.* **12**: 1165–1182.

CHIARI S., ZAULI A., MAZZIOTTA A., LUISELLI L., AUDISIO P. & CARPANETO G.M. 2013: Surveying an endangered saproxylic beetle, *Osmoderma eremita*, in Mediterranean woodlands: a comparison between different capture methods. — *J. Insect Conserv.* **17**: 171–181.

CHIARI S., ZAULI A., AUDISIO P., CAMPANARO A., DONZELLI P.F., ROMITI F., SVENSSON G.P., TINI M. & CARPANETO G.M. 2014: Monitoring presence, abundance and survival probability of the stag beetle, *Lucanus cervus*, using visual and odour-based capture methods: implications for conservation. — *J. Insect Conserv.* **18**: 99–109.

COX K., THOMAES A., ANTONINI G., ZILIOLI M., DE GELAS K., HARVEY D., SOLANO E., AUDISIO P., MCKEOWN N., SHAW P. ET AL. 2013: Testing the performance of a fragment of the COI gene to identify western Palaearctic stag beetle species (Coleoptera, Lucanidae). — *Zookeys* **365**: 105–126.

DRAG L., ZIMA J. & CIZEK L. 2013: Characterization of nine polymorphic microsatellite loci for a threatened saproxylic beetle *Rosalia alpina* (Coleoptera: Cerambycidae). — *Conserv. Genet. Resour.* **5**: 903–905.

DRAG L., HAUCK D., BERCES S., MICHALCEWICZ J., JELASKA L.S., AURENHAMMER S. & CIZEK L. 2015: Genetic differentiation of populations of the threatened saproxylic beetle Rosalia longicorn, *Rosalia alpina* (Coleptera: Cerambycidae) in central and South-east Europe. — *Biol. J. Linn. Soc.* **116**: 911–925.

FAIRCLOTH B.C. 2008: MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. — *Mol. Ecol. Resour.* **8**: 92–94.

HARVEY D.J., GANGE A.C., HAWES C.J., RINK M., ABDEHALDEN M., AL FULAIJ N., ASP T., BALLERIO A., BARTOLOZZI L., BRUSTEL H. ET AL. 2011a: Bionomics and distribution of the stag beetle, *Lucanus cervus* (L.) across Europe. — *Insect Conserv. Diver.* **4**: 23–38.

HARVEY D.J., HAWES C.J., GANGE A.C., FINCH P., CHESMORE D. & FARR I. 2011b: Development of non-invasive monitoring methods for larvae and adults of the stag beetle, *Lucanus cervus*. — *Insect Conserv. Diver.* **4**: 4–14.

MCKEOWN N.J. & SHAW P.W. 2008: Polymorphic nuclear microsatellite loci for studies of brown crab, *Cancer pagurus* L. — *Mol. Ecol. Resour.* **8**: 653–655.

MCKEOWN N.J., HAUSER L. & SHAW P.W. 2017: Microsatellite genotyping of brown crab *Cancer pagurus* reveals fine scale selection and 'non-chaotic' genetic patchiness within a high gene flow system. — *Mar. Ecol. Prog. Ser.* **566**: 91–103.

NIETO A. & ALEXANDER K.N.A. 2010: *European Red List of Saproxylic Beetles.* Publications Office of the European Union, Luxembourg, 54 pp.

PEAKALL R. & SMOUSE P.E. 2006: GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. — *Mol. Ecol. Resour.* **6**: 288–295.

ROSSITER S.J., JONES G., RANSOME R.D. & BARRATTT E.M. 2000: Genetic variation and population structure in the endangered greater horseshoe bat *Rhinolophus ferrumequinum*. — *Mol. Ecol.* **9**: 1131–1135.

ROZEN S. & SKALETSKY H. 2000: Primer3 on the WWW for general users and for biologist programmers. — *Methods Mol. Biol.* **132**: 365–386.

SALLE A., ARTHOFER W., LIEUTIER F., STAUFFER C. & KERDELHUE C. 2007: Phylogeography of a host-specific insect: genetic structure of *Ips typographus* in Europe does not reflect past fragmentation of its host. — *Biol. J. Linn. Soc.* **90**: 239–246.

SHAW P.W., HENDRICKSON L., MCKEOWN N.J., STONIER T., NAUD M.J. & SAUER W.H.H. 2010: Discrete spawning aggregations of loliginid squid do not represent genetically distinct populations. — *Mar. Ecol. Prog. Ser.* **408**: 117–127.

SOLANO E., THOMAES A., COX K., CARPANETO G.M., CORTELLESSA S., BAVIERA C., BARTOLOZZI L., ZILIOLI M., CASIRAGHI M., AUDISIO P. & ANTONINI G. 2016: When morphological identification meets genetic data: the case of *Lucanus cervus* and *L. tetradon* (Coleoptera, Lucanidae). — *J. Zool. Syst. Evol. Res.* **54**: 197–205.

THOMAES A., KERVYN T. & MAES D. 2008: Applying species distribution modelling for the conservation of the threatened saproxylic stag beetle (*Lucanus cervus*). — *Biol. Conserv.* **141**: 1400–1410.

ZHANG D.X. 2004: Lepidoptern microsatellite DNA; redundant but promising. — *Trends Ecol. Evol.* **19**: 507–509.