Conservation Genetics Resources (2019) 11:253–257
https://doi.org/10.1007/s12686-018-1001-8

TECHNICAL NOTE

# SNP discovery in European lobster (*Homarus gammarus*) using RAD sequencing

**Tom L. Jenkins[1] · Charlie D. Ellis[2] · Jamie R. Stevens[1]**

Received: 17 January 2018 / Accepted: 20 January 2018 / Published online: 31 January 2018


## Abstract

The European lobster (*Homarus gammarus*) is a decapod crustacean with a high market value and therefore their fisheries are of major importance to the economies they support. However, over-exploitation has led to profound stock declines in some regions such as Scandinavia and the Mediterranean. To manage this resource sustainably, knowledge of population structure and connectivity is crucial to inform management about dispersal, recruitment, stock identification and food traceability. We used restriction-site associated DNA sequencing to develop novel SNP markers from 55 individuals encompassing much of the species range; SNPs were quality filtered, ranked using *F*-statistics and the top 96 SNPs adequate for primer design were retained. SNP markers were developed with the aim of maximising the power to detect genetic differentiation between: (i) Atlantic and Mediterranean lobsters and (ii) Atlantic lobsters. This panel of SNPs provides a useful resource for future studies of population genetic structure and assignment in *H. gammarus*.

**Keywords** Conservation genetics · Fisheries management · *Homarus gammarus* · Population assignment · RAD-seq · Single nucleotide polymorphism

The European lobster (*Homarus gammarus*) is a decapod crustacean belonging to the family Nephropidae. They are found on hard substrates hiding in crevices or on compressed muds, typically at depths from the low tide mark to 50 m, but they can occur at depths up to 150 m. *Homarus gammarus* is widely distributed, ranging from Morocco to Arctic Norway, including Skagerrak, and also in the Mediterranean where they are generally found more sparsely. The species' high market value makes it a highly-prized seafood product, so its fisheries are of great importance to the local and regional economies they support. However, current and historical over-exploitation has led to stock declines, some of which have been quite profound in several regions (e.g. Scandinavia, Mediterranean) and from which recovery has been slow or stagnant (Kleiven et al. 2012). This has led to the rearing of *H. gammarus* larvae in lobster hatcheries to produce juveniles which are released into the wild to supplement productive stocks where the risk of over-exploitation is high (Ellis et al. 2015).

Over the last decade, genetic diversity and population structure has been investigated in *H. gammarus* using traditional molecular markers including random amplification of polymorphic DNA (RAPDs) (Ulrich et al. 2001), allozymes (Jorstad et al. 2005), mtDNA restriction fragment length polymorphisms (RFLPs) (Triantafyllidis et al. 2005) and microsatellites (Huserbraten et al. 2013; Watson et al. 2016; Ellis et al. 2017). However, single nucleotide polymorphisms (SNPs) are becoming the marker of choice in molecular ecology studies, particularly for non-model organisms without a well-annotated genome, because they are (i) abundant and generally widespread in the genome, (ii) eligible for high-throughput screening and automation, and (iii) reproducible across labs (Seeb et al. 2011). Moreover, genomics now enables thousands to tens of thousands of SNPs to be discovered in non-model marine organisms, meaning we have greater power over previous genetic markers to resolve spatial patterns of genetic differentiation, which is thought

**Electronic supplementary material** The online version of this article (https://doi.org/10.1007/s12686-018-1001-8) contains supplementary material, which is available to authorized users.

✉ Tom L. Jenkins
t.l.jenkins@exeter.ac.uk

[1] Department of Biosciences, College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4QD, UK

[2] National Lobster Hatchery, South Quay, Padstow, Cornwall PL28 8BL, UK

**Table 1** Summary information for the 96 SNP markers developed for the European lobster (*Homarus gammarus*)

| Locus ID | Sequence length (bp) | SNP | $H_o$ | $H_e$ | MAF | $F_{IS}$ | $P_{HWE}$ |
|---|---|---|---|---|---|---|---|
| H_gam_03441 | 442 | G/A | 0.407 | 0.460 | 0.355 | 0.045 | 0.640 |
| H_gam_04173 | 496 | C/T | 0.599 | 0.492 | 0.409 | −0.457 | 0.273 |
| H_gam_06157 | 264 | G/C | 0.383 | 0.425 | 0.282 | −0.094 | 0.701 |
| H_gam_07502 | 97 | C/T | 0.568 | 0.499 | 0.445 | −0.184 | 0.574 |
| H_gam_07892 | 97 | A/T | 0.204 | 0.268 | 0.155 | 0.089 | 0.273 |
| H_gam_08953 | 496 | G/T | 0.222 | 0.423 | 0.308 | 0.470 | 0.018 |
| H_gam_09441 | 496 | A/G | 0.414 | 0.378 | 0.264 | −0.080 | 0.691 |
| H_gam_11071 | 400 | G/A | 0.179 | 0.251 | 0.145 | 0.304 | 0.239 |
| H_gam_11183 | 130 | A/G | 0.537 | 0.496 | 0.445 | −0.072 | 0.716 |
| H_gam_11291 | 270 | T/G | 0.167 | 0.213 | 0.120 | −0.274 | 0.306 |
| H_gam_12971 | 496 | A/G | 0.395 | 0.426 | 0.309 | 0.056 | 0.702 |
| H_gam_14047 | 496 | C/T | 0.401 | 0.417 | 0.300 | −0.252 | 0.759 |
| H_gam_14742 | 496 | G/A | 0.216 | 0.331 | 0.222 | 0.217 | 0.097 |
| H_gam_15109 | 496 | T/A | 0.265 | 0.423 | 0.300 | 0.215 | 0.087 |
| H_gam_15128 | 142 | C/T | 0.383 | 0.425 | 0.291 | −0.094 | 0.532 |
| H_gam_15435 | 496 | C/T | 0.173 | 0.190 | 0.109 | 0.029 | 0.611 |
| H_gam_15531 | 122 | G/A | 0.284 | 0.476 | 0.391 | 0.074 | 0.029 |
| H_gam_15581 | 107 | A/G | 0.290 | 0.365 | 0.236 | 0.013 | 0.298 |
| H_gam_18512 | 496 | G/T | 0.290 | 0.337 | 0.218 | −0.179 | 0.426 |
| H_gam_18652 | 201 | A/G | 0.451 | 0.473 | 0.364 | 0.037 | 0.712 |
| H_gam_19266 | 175 | C/T | 0.284 | 0.296 | 0.182 | −0.125 | 0.759 |
| H_gam_19460 | 247 | C/T | 0.432 | 0.477 | 0.382 | 0.016 | 0.646 |
| H_gam_20354 | 142 | C/T | 0.469 | 0.430 | 0.309 | −0.194 | 0.626 |
| H_gam_21197 | 163 | C/T | 0.525 | 0.498 | 0.436 | −0.197 | 0.759 |
| H_gam_21880 | 496 | A/C | 0.481 | 0.503 | 0.463 | 0.054 | 0.706 |
| H_gam_22323 | 439 | G/A | 0.586 | 0.491 | 0.418 | −0.358 | 0.465 |
| H_gam_22365 | 176 | A/T | 0.340 | 0.449 | 0.318 | −0.038 | 0.291 |
| H_gam_22740 | 138 | T/C | 0.370 | 0.386 | 0.245 | −0.328 | 0.689 |
| H_gam_23146 | 174 | T/C | 0.315 | 0.442 | 0.300 | 0.156 | 0.206 |
| H_gam_23447 | 114 | T/C | 0.358 | 0.472 | 0.382 | 0.229 | 0.275 |
| H_gam_23481 | 137 | T/A | 0.296 | 0.369 | 0.236 | 0.159 | 0.307 |
| H_gam_23677 | 228 | A/G | 0.370 | 0.488 | 0.418 | 0.202 | 0.229 |
| H_gam_23787 | 496 | T/G | 0.185 | 0.246 | 0.155 | 0.078 | 0.267 |
| H_gam_24020 | 496 | C/G | 0.216 | 0.223 | 0.127 | −0.076 | 0.759 |
| H_gam_25229 | 230 | C/G | 0.259 | 0.227 | 0.118 | −0.395 | 0.759 |
| H_gam_25580 | 101 | C/T | 0.630 | 0.497 | 0.436 | −0.302 | 0.276 |
| H_gam_25608 | 97 | C/T | 0.185 | 0.264 | 0.164 | 0.268 | 0.161 |
| H_gam_27329 | 97 | T/C | 0.407 | 0.504 | 0.464 | 0.122 | 0.462 |
| H_gam_28357 | 496 | G/A | 0.444 | 0.497 | 0.436 | 0.175 | 0.587 |
| H_gam_29410 | 97 | T/C | 0.420 | 0.476 | 0.391 | 0.001 | 0.553 |
| H_gam_29801 | 496 | A/G | 0.179 | 0.267 | 0.164 | 0.100 | 0.113 |
| H_gam_29889 | 496 | A/G | 0.383 | 0.483 | 0.400 | 0.172 | 0.451 |
| H_gam_30339 | 496 | G/A | 0.228 | 0.231 | 0.136 | −0.285 | 0.759 |
| H_gam_31462 | 140 | C/A | 0.327 | 0.455 | 0.345 | 0.211 | 0.198 |
| H_gam_31618 | 496 | A/G | 0.333 | 0.369 | 0.236 | 0.059 | 0.595 |
| H_gam_31967 | 195 | A/C | 0.302 | 0.429 | 0.318 | 0.203 | 0.180 |
| H_gam_31979 | 182 | G/T | 0.259 | 0.380 | 0.245 | −0.080 | 0.223 |
| H_gam_32358 | 496 | G/A | 0.074 | 0.198 | 0.109 | −0.169 | 0.036 |
| H_gam_32362 | 213 | C/T | 0.630 | 0.497 | 0.436 | −0.302 | 0.276 |
| H_gam_32435 | 496 | T/C | 0.210 | 0.246 | 0.145 | 0.070 | 0.489 |

**Table 1** (continued)

| Locus ID | Sequence length (bp) | SNP | $H_o$ | $H_e$ | MAF | $F_{IS}$ | $P_{HWE}$ |
|---|---|---|---|---|---|---|---|
| H_gam_33066 | 218 | C/A | 0.370 | 0.386 | 0.245 | −0.328 | 0.685 |
| H_gam_33784 | 136 | A/G | 0.463 | 0.504 | 0.491 | 0.002 | 0.715 |
| H_gam_34443 | 302 | G/A | 0.346 | 0.453 | 0.327 | 0.186 | 0.215 |
| H_gam_34818 | 192 | A/C | 0.259 | 0.281 | 0.173 | 0.066 | 0.671 |
| H_gam_35584 | 97 | A/T | 0.346 | 0.445 | 0.336 | 0.149 | 0.306 |
| H_gam_36910 | 97 | A/G | 0.395 | 0.482 | 0.400 | 0.096 | 0.458 |
| H_gam_39107 | 127 | C/T | 0.216 | 0.223 | 0.127 | 0.016 | 0.759 |
| H_gam_39876 | 134 | C/T | 0.296 | 0.312 | 0.200 | −0.155 | 0.574 |
| H_gam_41521 | 97 | A/T | 0.438 | 0.451 | 0.355 | 0.051 | 0.759 |
| H_gam_42395 | 496 | T/C | 0.314 | 0.472 | 0.380 | 0.107 | 0.119 |
| H_gam_42529 | 496 | A/C | 0.364 | 0.365 | 0.227 | −0.166 | 0.759 |
| H_gam_42821 | 190 | G/A | 0.167 | 0.185 | 0.100 | 0.006 | 0.581 |
| H_gam_44670 | 251 | T/C | 0.204 | 0.398 | 0.255 | 0.402 | 0.000 |
| H_gam_45154 | 496 | G/A | 0.377 | 0.470 | 0.373 | 0.207 | 0.472 |
| H_gam_45217 | 496 | G/A | 0.265 | 0.259 | 0.145 | −0.136 | 0.759 |
| H_gam_51159 | 97 | T/G | 0.432 | 0.398 | 0.273 | −0.283 | 0.692 |
| H_gam_51507 | 97 | G/A | 0.308 | 0.357 | 0.224 | −0.250 | 0.443 |
| H_gam_53052 | 496 | A/T | 0.407 | 0.368 | 0.227 | −0.288 | 0.684 |
| H_gam_53263 | 496 | T/A | 0.383 | 0.392 | 0.255 | −0.304 | 0.691 |
| H_gam_53314 | 496 | T/C | 0.327 | 0.345 | 0.218 | −0.114 | 0.691 |
| H_gam_53720 | 96 | C/T | 0.568 | 0.495 | 0.435 | −0.335 | 0.483 |
| H_gam_53889 | 496 | G/C | 0.191 | 0.194 | 0.118 | −0.016 | 0.631 |
| H_gam_53935 | 468 | C/T | 0.284 | 0.476 | 0.391 | 0.074 | 0.018 |
| H_gam_54240 | 97 | A/C | 0.444 | 0.420 | 0.287 | −0.182 | 0.759 |
| H_gam_54762 | 496 | C/T | 0.531 | 0.491 | 0.436 | −0.345 | 0.651 |
| H_gam_55111 | 146 | C/T | 0.488 | 0.503 | 0.500 | −0.222 | 0.759 |
| H_gam_55142 | 178 | T/G | 0.327 | 0.490 | 0.426 | 0.270 | 0.164 |
| H_gam_55564 | 496 | G/A | 0.370 | 0.503 | 0.482 | 0.128 | 0.264 |
| H_gam_56423 | 182 | C/T | 0.420 | 0.427 | 0.291 | 0.127 | 0.705 |
| H_gam_56785 | 99 | T/C | 0.444 | 0.497 | 0.436 | 0.175 | 0.575 |
| H_gam_57131 | 97 | T/G | 0.377 | 0.407 | 0.282 | −0.090 | 0.698 |
| H_gam_57989 | 408 | A/T | 0.451 | 0.450 | 0.336 | −0.027 | 0.759 |
| H_gam_58053 | 97 | A/G | 0.049 | 0.179 | 0.100 | 0.046 | 0.000 |
| H_gam_59503 | 97 | T/A | 0.593 | 0.492 | 0.427 | −0.274 | 0.335 |
| H_gam_59586 | 201 | G/T | 0.296 | 0.487 | 0.394 | 0.261 | 0.062 |
| H_gam_59967 | 178 | C/T | 0.358 | 0.382 | 0.255 | 0.090 | 0.693 |
| H_gam_60546 | 167 | C/A | 0.333 | 0.494 | 0.427 | 0.252 | 0.166 |
| H_gam_63140 | 496 | C/T | 0.321 | 0.341 | 0.209 | −0.070 | 0.683 |
| H_gam_63267 | 97 | G/C | 0.395 | 0.381 | 0.255 | −0.085 | 0.759 |
| H_gam_63581 | 139 | T/C | 0.426 | 0.437 | 0.318 | −0.101 | 0.705 |
| H_gam_63605 | 132 | T/C | 0.451 | 0.487 | 0.409 | −0.147 | 0.716 |
| H_gam_63771 | 97 | A/G | 0.346 | 0.454 | 0.343 | 0.227 | 0.287 |
| H_gam_63798 | 188 | G/A | 0.568 | 0.486 | 0.418 | −0.237 | 0.443 |
| H_gam_65064 | 496 | C/A | 0.370 | 0.386 | 0.245 | −0.328 | 0.685 |
| H_gam_65376 | 496 | C/A | 0.364 | 0.429 | 0.309 | 0.134 | 0.511 |
| H_gam_65576 | 173 | A/C | 0.352 | 0.376 | 0.236 | −0.457 | 0.592 |

Sequences and additional SNP information can be found in S4 Supplementary Material

*SNP* single nucleotide polymorphism, $H_o$ observed heterozygosity, $H_e$ expected heterozygosity, *MAF* minor allele frequency, $F_{IS}$ inbreeding coefficient, $P_{HWE}$ *P*-values for Hardy–Weinberg equilibrium corrected for multiple comparisons using the false discovery rate

to be particularly beneficial when studying highly dispersive marine species that exhibit typically weak genetic differentiation (e.g. American lobster, Benestan et al. 2015). These advances have also led to the development of small panels of informative SNPs (e.g. Nielsen et al. 2012; Villacorta-Rath et al. 2016) that are likely to be useful for assessments of genetic structure, population assignment and connectivity.

In this study, we used restriction-site associated DNA (RAD) sequencing to isolate and characterise 96 novel SNP markers in *H. gammarus*. Genomic DNA was extracted from v-notch or pleopod tissue using a modified salting-out protocol (Li et al. 2011) (S1 Supplementary Material). The RAD library was prepared in-house using Illumina Nextera XT barcodes and comprised 55 individuals from 27 geographically separate sampling locations, ranging from the Mediterranean to the British Isles and Skagerrak (S2 Supplementary Material). The library was sequenced on an Illumina HiSeq 100 bp paired-end rapid run platform. Raw reads (available from Dryad, https://doi.org/10.5061/dryad.2pc6v) were cleaned and truncated to 97 bp using the process_radtags program in Stacks v1.45 (Catchen et al. 2013) and RAD loci were built using the denovo_map.pl wrapper script in Stacks using optimised parameters of $m=3$, $M=3$ and $n=3$ following the methods of Paris et al. (2017). The populations program was run using all 55 individuals and initial results indicated genetic differentiation between Mediterranean, Skagerrak and the remaining Atlantic samples (S3 Supplementary Material). Therefore, the program was also re-run using only samples from the Atlantic (excluding Mediterranean and Skagerrak samples). This approach maximised the potential to find SNPs that are most informative for detecting hierarchical genetic differentiation between Atlantic lobsters. Full details of the bioinformatics and parameters used are available in S3 Supplementary Information.

In total, 276 million reads were generated and a mean average of 97.9% across all samples were retained after quality control. After initial filtering in Stacks, 7022 biallelic SNPs were identified using all samples and 4377 biallelic SNPs were identified using only Atlantic samples. These SNPs were then ranked by highest $G''_{ST}$ (Meirmans and Hedrick 2011), sorted by the number of SNPs per RAD locus, and filtered for primer design adequacy and suitability for high-throughput genotyping on a Fluidigm EP1 system. The SNP panel was composed of the highest-ranked remaining SNPs; 21 SNPs were chosen from the dataset composed of all samples (aiming to capture differentiation between Atlantic and Mediterranean lobsters) and 78 SNPs were chosen from the dataset composed of only Atlantic samples (aiming to capture any potential hierarchical differentiation in the Atlantic).

Using these 96 SNP markers and all of our samples, we calculated several population genetic statistics for each locus (Table 1). The observed and expected heterozygosity ranged from 0.049 to 0.630 and 0.179 to 0.504, respectively. The minor allele frequency and the inbreeding coefficient ranged from 0.100 to 0.504 and −0.457 to 0.470, respectively. After false discovery rate correction, six SNPs deviated significantly from Hardy–Weinberg equilibrium ($P<0.05$). To our knowledge, this is the first development of SNP markers in *H. gammarus*, and therefore these novel markers offer a valuable tool for future studies of spatial genetic structure and population assignment in this species.

**Acknowledgements** We thank the many people who provided lobster tissue samples for this research. We also thank Karen Moore and staff at the Exeter Sequencing Service (Exeter, UK) for constructing the RAD libraries. This research was funded by a Natural Environment Research Council UK GW4 + DTP studentship (Grant No. NE/L002434/1), Natural England (Ref. PO 904130) and the University of Exeter, and forms part of the PhD of TLJ.



## References

Benestan L, Gosselin T, Perrier C et al (2015) RAD-genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species; the American lobster (*Homarus americanus*). Mol Ecol 24:3299–3315. https://doi.org/10.1111/mec.13245

Catchen J, Hohenlohe PA, Bassham S et al (2013) Stacks: an analysis tool set for population genomics. Mol Ecol 22:3124–3140. https://doi.org/10.1111/mec.12354

Ellis CD, Hodgson DJ, Daniels CL et al (2015) European lobster stocking requires comprehensive impact assessment to determine fishery benefits. ICES J Mar Sci 72:i35–i48. https://doi.org/10.1093/icesjms/fsu196

Ellis CD, Hodgson DJ, Daniels CL et al (2017) Population genetic structure in European lobsters: implications for connectivity, diversity and hatchery stocking. Mar Ecol Prog Ser 563:123–137. https://doi.org/10.3354/meps11957

Huserbraten MBO, Moland E, Knutsen H et al (2013) Conservation, spillover and gene flow within a network of northern European Marine Protected Areas. PLoS ONE 8:e73388. https://doi.org/10.1371/journal.pone.0073388

Jorstad KE, Faresteit E, Kelly E, Triantaphyllids C (2005) Allozyme variation in European lobster (*Homarus gammarus*) throughout its distribution range. N Z J Mar Freshw Res 39:515–526. https://doi.org/10.1080/00288330.2005.9517330

Kleiven AR, Olsen EM, Vølstad JH (2012) Total catch of a red-listed marine species is an order of magnitude higher than official data. PLoS ONE 7:1–7. https://doi.org/10.1371/journal.pone.0031216

Li Y, Wang W, Liu X et al (2011) DNA extraction from crayfish exoskeleton. Indian J Exp Biol 49:953–957

Meirmans PG, Hedrick PW (2011) Assessing population structure: $F_{ST}$ and related measures. Mol Ecol Resour 11:5–18. https://doi.org/10.1111/j.1755-0998.2010.02927.x

Nielsen EE, Cariani A, Aoidh EM et al (2012) Gene-associated markers provide tools for tackling illegal fishing and false eco-certification. Nat Commun 3:851. https://doi.org/10.1038/ncomms1845

Paris JR, Stevens JR, Catchen JM (2017) Lost in parameter space: a road map for stacks. Methods Ecol Evol 8:1360–1373. https://doi.org/10.1111/2041-210X.12775

Seeb JE, Carvalho G, Hauser L et al (2011) Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. Mol Ecol Resour 11:1–8. https://doi.org/10.1111/j.1755-0998.2010.02979.x

Triantafyllidis A, Apostolidis AP, Katsares V et al (2005) Mitochondrial DNA variation in the European lobster (*Homarus gammarus*) throughout the range. Mar Biol 146:223–235. https://doi.org/10.1007/s00227-004-1435-2

Ulrich I, Muller J, Schutt C, Buchholz F (2001) A study of population genetics in the European lobster, *Homarus gammarus* (Decapoda, Nephropidae). Crustaceana 74:825–837. https://doi.org/10.1163/15685400152682593

Villacorta-Rath C, Ilyushkina I, Strugnell JM et al (2016) Outlier SNPs enable food traceability of the southern rock lobster, *Jasus edwardsii*. Mar Biol 163:163:223. https://doi.org/10.1007/s00227-016-3000-1

Watson HV, McKeown NJ, Coscia I et al (2016) Population genetic structure of the European lobster (*Homarus gammarus*) in the Irish Sea and implications for the effectiveness of the first British marine protected area. Fish Res 183:287–293. https://doi.org/10.1016/j.fishres.2016.06.015