Papetti *et al. BMC Res Notes (2016) 9:238*
DOI 10.1186/s13104-016-2039-x

BMC Research Notes

SHORT REPORT Open Access

Microsatellite markers for the notothenioid fish *Lepidonotothen nudifrons* and two congeneric species

Chiara Papetti[1*], Lars Harms[1], Jutta Jürgens[1], Tina Sandersfeld[1,2], Nils Koschnick[1], Heidrun Sigrid Windisch[1,3], Rainer Knust[1], Hans-Otto Pörtner[1] and Magnus Lucassen[1]

Abstract

Background: Loss of genetic variability due to environmental changes, limitation of gene flow between pools of individuals or putative selective pressure at specific markers, were previously documented for Antarctic notothenioid fish species. However, so far no studies were performed for the Gaudy notothen *Lepidonotothen nudifrons*. Starting from a species-specific spleen transcriptome library, we aimed at isolating polymorphic microsatellites (Type I; i.e. derived from coding sequences) suitable to quantify the genetic variability in this species, and additionally to assess the population genetic structure and demography in nototheniids.

Results: We selected 43,269 transcripts resulting from a MiSeq sequencer run, out of which we developed 19 primer pairs for sequences containing microsatellite repeats. Sixteen loci were successfully amplified in *L. nudifrons*. Eleven microsatellites were polymorphic and allele numbers per locus ranged from 2 to 17. In addition, we amplified loci identified from *L. nudifrons* in two other congeneric species (*L. squamifrons* and *L. larseni*). Thirteen loci were highly transferable to the two congeneric species. Differences in polymorphism among species were detected.

Conclusions: Starting from a transcriptome of a non-model organism, we were able to identify promising polymorphic nuclear markers that are easily transferable to other closely related species. These markers can be a key instrument to monitor the genetic structure of the three *Lepidonotothen* species if genotyped in larger population samples. When compared with anonymous loci isolated in other notothenioids, i.e. Type II (isolated from genomic libraries), they offer the possibility to test how the effects of occurring environmental change influence the population genetic structure in each species and subsequently the composition of the entire ecosystem.

Keywords: Antarctica, Gaudy notothen, *Lepidonotothen larseni*, *Lepidonotothen squamifrons*, Nototheniidae, Population differentiation, Southern ocean

Findings

Research background

Lepidonotothen nudifrons (Lönnberg 1905), the Gaudy notothen [1] is a benthic and moderately active fish species belonging to the family Nototheniidae (red-blood species, [2]). *L. nudifrons* is especially abundant along the Islands of the Scotia Arc until the Bransfield Strait (Antarctic Peninsula) [1, 3]. This species plays an ecologically important role as prey to piscivorous fish species in this region [3], however, little is known about its genetic variability and the demography. Since the strongest local effects of climate change in polar regions were recorded for the Antarctic Peninsula [4], it is essential to investigate if *L. nudifrons* is experiencing any loss of genetic variability or if gene flow is limited among pools of individuals. Genetic differentiation among populations has already been demonstrated for another ecologically important notothenioid (*Pleuragramma antarctica*) in the same area [5].

These investigations require microsatellites, which have proven useful and reliable genetic markers in similar

*Correspondence: chiara.papetti@unipd.it
[1] Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Am Handelshafen 12, Bremerhaven 27570, Germany
Full list of author information is available at the end of the article

studies [5, 6]. In the past, the most common procedure to isolate microsatellites required the identification of repeat-containing sequences from libraries of genomic DNA enriched for microsatellite motives (Type II loci, [6–9]). In comparison to the traditional but labour-intensive and costly approach, currently these markers are also isolated from transcriptomic libraries obtained by next generation sequencing (NGS) that has proven to be more cost effective [7, 10]. However, markers obtained with this latter approach are gene-associated simple sequence repeats (SSR, Type I, functional loci, either called expressed sequence tags EST-linked markers [6, 9, 11]).

Due to their position inside or their flanking coding gene sequence and due to functional constraints, EST-linked SSRs are expected to have higher probability to be under selective pressure [6]. Gene-associated microsatellites have been isolated across many taxa in recent years [10 and references therein], however, only one study has reported isolation of microsatellites from a transcriptomic database in notothenioids [12].

The primary goal of this study is to expand the panel of available EST-linked loci in notothenioids for future studies on the evolution and demographic history of Antarctic fish. Moreover, we aim to isolate specific markers for *Lepidonotothen* spp. in order to compare their variability with anonymous Type II markers isolated in other notothenioid species.

We therefore mined microsatellite sequences from the publicly available spleen transcriptome of *L. nudifrons* [13] and screened several promising candidate loci for amplification and polymorphism. We were able to isolate sixteen species-specific SSR markers from assembled transcript sequences of *L. nudifrons.* Loci identified from *L. nudifrons* were also tested in two congeneric species: *L. squamifrons* and *L. larseni*.

Methods

Gene-associated SSR markers were identified among 112,477 spleen transcripts of *L. nudifrons* [13]. The raw sequences were obtained from an Illumina MiSeq sequencer. Adapter clipping and quality trimming was performed using Trimmomatic v.0.32 [14] with following parameters: seed mismatch of 2, palindrome clip threshold of 30, simple clip threshold of 10, a minimum adapter length of 2, keep both reads parameter set to true, headcrop of 7, leading and trailing quality of 3, sliding window size of 4 with an average quality of 15 and a minimum sequence length of 50 bases. The subsequent *de novo* assembly was performed using the trinity genome-independent transcriptome assembler [release 17 July 2014, 15] with a minimum transcript length of 300 bases (for further details see [13]). Transcripts were screened using SciRoKo v. 3.4 [16]. Di-, tri-, tetra-, penta-, hexa-, hepta-, octa-, nona-, and deca-nucleotide repeat motifs were searched, setting the minimum repeat units to five, for all motif categories. Among transcripts containing microsatellites, we selected non-redundant SSRs with sufficiently large flanking sequences (>50 bp) on each side of the repeated units as "Potentially Amplifiable Loci" i.e. PAL [10, 17]. For subsequent analyses, we randomly selected 19 loci among all PAL. Primer pairs were designed with FastPCR v. 6.0 [18] to avoid primer dimers, self-annealing and hairpin formation when multiplexing loci during PCR. In addition, primers were designed to have very similar melting temperatures to avoid increased complexity of optimization protocol and to facilitate multiplexing.

Primer validation was carried out on genomic DNA extracted from 7 specimens of *L. nudifrons* collected in March/April 2012 near Elephant Island at 70–322 m depth during the RV Polarstern expedition ANT-XXVIII/4. The fish are not considered endangered at the sampling site. Sampling of the animals has been performed in accordance with the Antarctic treaty, and permission was given by the respective national authority (Umweltbundesamt: permit number I 3.5-94003-3/274). All treatments with live animals were in accordance with German law and approved by the competent national authority (Freie Hansestadt Bremen, Germany, permit number AZ: 522-27-11/02-00 (93)).

Total DNA was purified from 10 to 20 mg of muscle tissue following the standard protocol of the DNeasy Blood and Tissue Kit (Qiagen, Germany). Quality and quantity of DNA extractions were assessed using a NanoDrop™ 2000c spectrophotometer (Thermo Fisher Scientific, USA) before samples were stored at −20 °C.

Initially, PCR primer pairs were tested as single-locus PCR in 20 µl volume, containing 1X reaction buffer (5 Prime, Hamburg, Germany), 70 µM dNTPs, 0.25 µM of each primer, 1 unit Taq polymerase (5 units/µl, 5 Prime, Hamburg, Germany) and ~30 ng of genomic DNA. PCR conditions were: initial denaturation at 94 °C for 1 min, followed by 30 cycles of 94 °C for 30 s (denaturation), 54 °C for 40 s (annealing), 72 °C for 40 s (extension) and a final single extension step at 72 °C for 5 min. Electrophoresis was carried out at 100 V on 1.5 % agarose gels containing GelRed Nucleic Acid Gel Stain (Biotium, Hayward, USA) for a preliminary qualitative polymorphism detection. Only loci that provided a clear PCR product were retained (for details, see Results section) and their polymorphism was verified on an Applied Biosystems 3130 XL automated sequencer (Life Technologies, USA, ROX500 as size standard) using a larger sample of 21 *L. nudifrons* individuals (collected in March/April 2012 near Elephant Island, Research Vessel RV Polarstern, expedition ANT-XXVIII/4, DNA

extraction as described above). To this purpose, forward primers were labelled with fluorescent dyes FAM, HEX, and TAMRA (Applied Biosystems, USA) and loci were combined in multiplex PCRs designed with Multiplex Manager v. 1.2 [19]. Amplification reactions were optimised in terms of volumes and concentrations of reagents to reduce the total genotyping cost. Amplifications were carried out in 10 µl reaction volume, using the Multiplex PCR kit (Qiagen, Germany) in accordance with manufacturer's instructions. PCR conditions were: initial denaturation at 95 °C for 15 min, followed by 30 cycles of 94 °C for 30 s (denaturation), 57 °C for 90 s (annealing), 72 °C for 60 s (extension) and a final single extension step at 60 °C for 30 min.

Specimens of *L. squamifrons* (n = 20) and *L. larseni* (n = 18) were tested to assess loci variability in congeneric species (DNA extraction and amplification conditions as described above). Individuals of these two species were collected around Elephant Island in March/April 2012 (RV Polarstern expedition ANT-XXVIII/4).

Allele sizes were assigned using GeneMarker v. 2.6.3 (Soft-Genetics, Pennsylvania). Binning was automated with Flexibin v. 2 [20]. All input files for statistical analysis were produced with Create v. 1.37 [21]. All standard and basic statistics were produced with DiveRsity v. 1.9 [22]. The estimation of departure from Hardy–Weinberg equilibrium (HWE) was obtained with Genepop (online version, exact test) [23]. The presence of null alleles was assessed using the program ML-NullFreq [24]. Genotypic disequilibrium for pairs of loci (Fishers' exact test) was tested with Genepop (online version) [23]. Correction for multiple testing (HWE and genotypic disequilibrium) was accomplished using the standard Bonferroni technique [25, 26].

To annotate the transcripts corresponding to the final panel of markers, homology searches were performed using Blastx [27] against the UniProtKB/Swiss-Prot database with an e-value cut-off of 10^{-9}.

Results

We identified 43,269 transcripts containing SSRs with 2–10 repeat units out of 128 Mb transcripts of a spleen transcriptome (Table 1). This result indicated an average amount of one SSR per 2.9 Kb. The number of transcripts containing SSR represented approximately 38.46 % of the total sequenced transcriptome. The largest proportion of SSRs (82.50 %) consisted of di-nucleotide repeats. SSRs with nucleotide repeat of higher complexity (e.g. tri- to deca-nucleotide repeat motifs) were present in progressively smaller numbers (Table 1). A list of all the transcripts containing SSR motifs can be provided upon request. 484 transcripts resulted to be suitable for primer design (PAL), and 19 of these sequences (containing a di-nucleotide repeat motif) were randomly chosen for primer design. Of these 19 loci, two did not provide any amplified fragment, even after attempting to optimize annealing temperatures. One locus showed a longer fragment than expected based on the original sequence (length determined by agarose gel-sizing), possibly suggesting the presence of an intron and therefore making allele sizes unpredictable [11]. These three markers were discarded from subsequent analyses. The remaining 16 markers showed a clearly defined amplified band on agarose gel and could be easily genotyped with an automatic sequencer in three multiplex PCRs containing eight, six, and two loci. The final panel of markers (Tables 2 and 3) consists of five monomorphic and eleven polymorphic loci for *L. nudifrons* with no missing genotypes. The allele number per locus ranged from 2 to 17 (Table 3), with an average value of 5.1 (±4.4 SD). Mean observed (H_o) and expected heterozygosities (H_e) were 0.43 (±0.34 SD) and 0.43 (±0.31 SD). Ten loci out of eleven were in HWE after correction for multiple testing (nominal significance level $\alpha = 0.05$, Table 3). Hardy–Weinberg disequilibrium was detected for locus Ln_42016, due to excess of homozygosity. This result could be due to several reasons, such as single locus stochasticity due to small sample size (n = 21) or occurrence of a non-amplified allele. The presence of a null allele was suggested by ML-NullFreq [24] at this locus with a frequency of 4.9 %. Loci with null alleles might have impacts on estimates of population differentiation [28] and are generally not recommended for

Table 1 Summary of Simple Sequence Repeats (SSR) found in the transcriptome of *Lepidonotothen nudifrons* [13]

Type of repeat (length of motif in number of nucleotides)	No. of transcripts
SSR (all motifs)	*43,269*
SSR (2)	35,697
SSR (3)	6905
SSR (4)	549
SSR (5)	32
SSR (6)	23
SSR (7)	5
SSR (8)	15
SSR (9)	26
SSR (10)	17
PAL SSR with ≥ 8 repeats (all motifs)	*484*
PAL SSR (2)	478
PAL SSR (3)	5
PAL SSR (4)	1

SSR were searched among 112,477 transcripts. PAL, Potentially Amplifiable Loci [10, 17]

Table 2 Characteristics of 16 SSR loci in *Lepidonotothen nudifrons* and two additional congeneric species

Locus name	Accession #	Blastx Top Hit	Repeat content	Primers (5′–3′)	Fluorescent label	Multiplex mix	Ln	Ls	Ll
Ln_22268	HACN01006381	–	$(GA)_9$	F: TGGATGGTTTTTCGTTTGACCAA R: TGAAGAGCTTTGGAGCAACAA	TAMRA	1	M	M	M
Ln_22517	HACN01006568	–	$(AT)_9$	F: TGATTCATCCTCTGGGTACCAT R: AGCTCGTAGCTATTCCAGCA	HEX	2	P	M	–
Ln_23194	HACN01007070	–	$(TA)_9$	F: TCCGGAGGCTGCGTTCAGAG R: CTGCTTGAGGACCGGCTCAG	HEX	1	M	P	P
Ln_32298	HACN01014077	Dolichol-phosphate mannosyltransferase subunit 3 (Q3ZC71)	$(AC)_9$	F: AAAGTAGTCCGTGCATCCCTT R: GTCTTACGACCTGTTGGAGC	TAMRA	2	M	M	M
Ln_34878	HACN01015998	–	$(AG)_{10}$	F: GTTGCAGTCACGTTGAAAGC R: ACGAATCCAAACCTTGCAACAT	FAM	3	P	–	–
Ln_35217	HACN01016251	–	$(CT)_9$	F: CCACTACACAGCTGAAATGGT R: GGCATGATTGGAGCACTTCC	HEX	1	P	P	P
Ln_36100	HACN01016925	–	$(AG)_9$	F: TAAAGCTCCAGCCGATACTGG R: TGGGAATTGAACAGATGTCACC	TAMRA	1	P	M	P
Ln_36156	HACN01016972	–	$(GT)_9$	F: GTCTGCTCATGTCCAACACAT R: AGGTGTTTCTGTAAACCCTCG	FAM	2	P	P	P
Ln_40551	HACN01020461	–	$(TA)_9$	F: TAACTGATGCATGCCCAGGAACTT R: GCTTCCTAAAAGCTCCACAGAT	FAM	1	P	P	P
Ln_41246	HACN01021014	–	$(AT)_9$	F: GCTTAAATCCAACCATAGACGCA R: CTCATCAG-GACTTCACGTGC	FAM	2	M	–	–
Ln_41281	HACN01021039	Uncharacterized protein C1orf21 homolog (Q8K207)	$(GT)_9$	F: TGTGTTTTTGGACGTATGGCA R: GCTTCCATTTTGGGAGTACCT	HEX	1	P	P	P

Table 2 continued

Locus name	Accession #	Blastx Top Hit	Repeat content	Primers (5′–3′)	Fluorescent label	Multiplex mix	Ln	Ls	Ll
Ln_42016	HACN01021779	–	$(GT)_9$	F: ACATCACTTG CAAATAGGCCGGTG R: CCGTCATGCACCTGG GAGTT	HEX	3	P	P	P
Ln_42233	HACN01021981	Serine/threonine-protein kinase N2 (A7MBL8)	$(AT)_9$	F: ATTCCTGACTGC CAAAGACG R: TGTTGACTTGCA GACTGAACG	FAM	1	P	M	P
Ln_42701	HACN01022446	Hepatic leukemia factor (Q8BW74)	$(AT)_9$	F: TGGACTTCACAA GCGTGGCT R: ATTGAAAGTGTGCAT AGTCCGT	FAM	1	M	M	M
Ln_45257	HACN01024934	–	$(TG)_9$	F: AAGACCAGGGC GAGTCTGAC R: CTCTAACCAGAACAC CGCCT	HEX	2	P	P	P
Ln_45589	HACN01025236	–	$(AG)_9$	F: CGTGTCTAGGAGCTA CAGCA R: CCAGGTTGGTG GCATCTCTT	TAMRA	2	P	P	P
						Total	11	8	10

The table reports: the name of each locus (Ln = *Lepidonotothen nudifrons* and the transcript number), Accession number of the *de novo* assembled transcripts of *L. nudifrons* (European Nucleotide Archive -ENA- project number PRJEB8919) [13], the putatively annotated loci (Blastx Top Hit, e-value < 10^{-9}), the repeat content, the forward (F) and reverse (R) primer sequences, the fluorescent label, the multiplex pool (mix). Variability, expressed in terms (P) polymorphic, (M) monomorphic or (-) unsuccessful locus, was assessed on 21 *L. nudifrons* (Ln), 20 *L. squamifrons* (Ls) and 18 *L. larseni* (Ll) individuals. Total number of polymorphic loci is reported in the last row

Table 3 Summary statistics of 16 SSR loci genotyped in *Lepidonotothen nudifrons*, *L. squamifrons* and *L. larseni*

Locus name	*Lepidonotothen nudifrons* (21)					*Lepidonotothen squamifrons* (20)					*Lepidonotothen larseni* (18)				
	Size range (bp)	Na	H_o	H_e	pHWE	Size range (bp)	Na	H_o	H_e	pHWE	Size range (bp)	Na	H_o	H_e	pHWE
Ln_22268	108	1	0	0	–	108	1	0	0	–	108	1	0	0	–
Ln_22517	251–263	4	0.29	0.26	1	245	1	0	0	–	–	–	–	–	–
Ln_23194	92	1	0	0	–	81–92	2	0.05	0.05	1	91–94	2	0.22	0.28	0.3899
Ln_32298	275	1	0	0	–	275	1	0	0	–	275	1	0	0	–
Ln_34878	126–138	7	0.86	0.8	0.0052	–	–	–	–	–		–	–	–	–
Ln_35217	242–246	3	0.33	0.42	0.1606	237–256	6	0.7	0.63	1	240–252	6	0.83	0.68	0.8929
Ln_36100	258–264	2	0.1	0.09	1	263	1	0	0	–	240–263	3	0.22	0.42	0.0326
Ln_36156	186–244	17	1	0.92	0.9627	191–241	14	0.8	0.86	0.1197	184–248	16	0.89	0.92	0.6330
Ln_40551	291–329	6	0.52	0.6	0.4682	272–294	2	0.5	0.38	0.2770	287–314	10	0.17	0.87	*P < 0.0001*
Ln_41246	89	1	0	0	–	–	–	–	–	–	–	–	–	–	–
Ln_41281	314–320	2	0.1	0.09	1	314–320	2	0.35	0.44	0.3492	314–320	3	0.22	0.2	1
Ln_42016	127–139	7	0.62	0.73	*P < 0.0001*	135–139	2	0	0.45	*P < 0.0001*	135–139	3	0.11	0.32	*0.0029*
Ln_42233	212–242	2	0.1	0.09	1	214	1	0	0	–	212–242	3	0.22	0.29	0.0611
Ln_42701	107	1	0	0	–	107	1	0	0	–	107	1	0	0	–
Ln_45257	98–100	2	0.05	0.13	0.0731	78–92	3	0.05	0.43	*P < 0.0001*	78–101	3	0.17	0.25	0.0392
Ln_45589	132–146	4	0.76	0.64	0.1007	129–171	8	0.7	0.63	1	131–137	4	0.61	0.55	0.9367

For each species (sample size in brackets), the table shows the size range of amplified fragments in bp, the number of alleles (Na) detected, the observed (H_o) and expected (H_e) heterozygosities and the Hardy–Weinberg equilibrium probability (pHWE). Significant p-values, after correction for multiple testing using the standard Bonferroni technique [25, 26] are in italics (nominal significance level $\alpha = 0.05$)

use in population genetic inference. However, it has been shown that the influence of null alleles in studies of population genetics might be marginal compared to other factors such as the number of loci and strength of population differentiation [29]. Departure from HWE due to excess of homozygosity could also be caused by other factors either to scoring errors or amplification artefacts. Departure from HWE could be generated by the pressure of actually occurring evolutionary forces (e.g. selection, local adaptation) or admixture of genetically distinct populations (i.e. Wahlund effect [30]). This can affect loci at different magnitude with some triggers being locus-specific and others being sample-specific [31]. In some circumstances, the usual approach of consistently removing loci in Hardy–Weinberg disequilibrium may be too conservative leading to the exclusion of ecologically informative markers [31, 32].

Although our loci are located in transcribed sequences, only three loci could be putatively annotated through similarity search (Blastx, Table 2) [27]. A fourth marker resulted in a similarity to an "uncharacterized protein C1orf21 homolog" (Table 2). These transcripts are likely portions of protein-coding genes and future studies should aim at combining information about allelic frequencies, gene expression and function. All polymorphic loci were in linkage equilibrium after correction for multiple testing (nominal significance level $\alpha = 0.01$, data not shown).

In the congeneric species, *L. squamifrons* and *L. larseni*, two loci failed to amplify consistently despite attempts to optimize the PCR conditions. Locus Ln_22517 was successfully amplified in *L. squamifrons* (monomorphic), but it did not provide a consistent amplification and genotyping result for *L. larseni*. The remaining 13 markers were successfully genotyped (Table 3). Eight loci were polymorphic in *L. squamifrons*, while in *L. larseni* we scored multiple alleles in ten loci (Tables 2 and 3). In particular, Ln_23194, monomorphic in *L. nudifrons*, turned out to be polymorphic in *L. squamifrons* and *L. larseni* (Tables 2 and 3). We obtained a panel of loci that worked for all three species at the same PCR conditions, which facilitates the rapid implementation and application at large sample size scale.

Our results confirm that polymorphic SSR markers can be effectively isolated from transcriptomes of non-model organisms [12, 33]. For *Lepidonotothen* spp. these markers are also easily transferable among different, but phylogenetically closely related species.

Additional tests should be implemented to verify whether these loci could be considered candidates for being influenced by selection. An effective method to verify this, is to compare F_{ST} values to search for loci showing a significantly high level of genetic differentiation between pairs of species as applied for the *Chionodraco* genus [6]. Agostini et al. [6] indicated that out of 21 microsatellites two Type I markers and one Type II locus were putatively under selection.

Moreover, we aim to use our microsatellites in conjunction with other markers such as mitochondrial and nuclear sequences and on a larger sample size. This would effectively help monitoring the genetic structure of the three *Lepidonotothen* species, profiling demographic events occurred in the past, and identifying signatures of local adaptation in genetically different populations. To understand the significance of these events and the climate challenges that marine organisms in polar regions are facing, is essential to model and project future population viability for species management and conservation [34].

Abbreviations

ENA: European nucleotide archive; EST: expressed sequence tags; H_e: expected heterozygosity; H_o: observed heterozygosity; HWE: Hardy–Weinberg equilibrium; NGS: next generation sequencing; PAL: potentially amplifiable loci; RV: research vessel; SSR: simple sequence repeats.

Authors' contributions

CP, ML and HOP conceived and designed the study, CP, TS, NK, RK and ML performed fish sampling, JJ and NK performed microsatellite lab work, CP, LH and HW performed SSR identification, primer design, genotyping and allele scoring and basic statistics description. HW and CP advised on the laboratory protocols. CP, LH, HW and ML wrote the paper. All authors critically read and approved the final manuscript.

Author details

[1] Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Am Handelshafen 12, Bremerhaven 27570, Germany. [2] University of Bremen, BreMarE-Bremen Marine Ecology, Leobener Str. NW2, 28359 Bremen, Germany. [3] Institute for Cell Biology and Zoology, Heinrich-Heine-University, Universitätsstrasse 1, Düsseldorf 40225, Germany.

Acknowledgements

Authors thank the captain and crew of the RV Polarstern for their precious help during samples collection (expedition ANT-XXVIII/4, Antarctic Peninsula). Authors are grateful to Imke Fries, Nancy Lange, and Lars Henning (AWI) for their support in the Marie Curie project development and management. Authors thank M. La Mesa and S. Frickenhaus for critically revising an early version of the manuscript. We also thank Anna Dowden-Williams for copy-editing the English.

Availability of data and materials

The *de novo* assembled transcripts of *L. nudifrons* spleen are deposited in the European Nucleotide Archive' (ENA) at the 'European Molecular Biological Laboratory—European Bioinformatics Institute' (EMBL-EBI) under accession number PRJEB8919. Individual transcripts accession numbers range from HACN01000001 to HACN01112477.

Competing interests

The authors declare that they have no competing interests.

Consent to publish

Not applicable.

Ethics

All treatments with live animals were in accordance with German law and approved by the competent national authority (Freie Hansestadt Bremen, Germany, permit number AZ: 522-27-11/02-00 (93)).

Funding
This work was supported by the European Marie Curie project "Polarexpress" Grant No. 622320 to CP. CP acknowledges also financial support of the Hanse-Wissenschaftskolleg, Institute for Advanced Study in Delmenhorst, Germany. This paper is a contribution to the research programme PACES II, Topic 1, Workpackage 6 of the Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research (AWI). Funding bodies had no decisional role in design, collection, analysis, and interpretation of data, in the writing of the manuscript, and in the decision to submit the manuscript for publication.

Received: 14 December 2015 Accepted: 14 April 2016
Published online: 26 April 2016

References

1. Gon O, Heemstra PC. Fishes of the Southern Ocean. Grahamstown, South Africa: J.L.B. Smith Institute of Ichthyology; 1990. pp. 299–300.
2. Eakin RR, Eastman, JT. Notothenioid species list 2015. http://www.oucom.ohiou.edu/dbms-eastman/Accessed Accessed 15 March 2015.
3. Llompart F, Delpiani M, Lattuca E, Delpiani G, Cruz-Jiménez A, Orlando P, et al. Spatial patterns of summer demersal fish assemblages around the Antarctic Peninsula and South Shetland Islands. Antarct Sci. 2015;27(02):109–17. doi:10.1017/S0954102014000352.
4. IPCC. Climate Change 2014: Impacts, adaptation, and vulnerability. Part A: Global and sectoral aspects. Contribution of Working group II to the fifth assessment report of the intergovernmental panel on climate change. Cambridge, United Kingdom and New York, NY, USA; 2014.
5. Agostini C, Patarnello T, Ashford JR, Torres JJ, Zane L, Papetti C. Genetic differentiation in the ice-dependent fish *Pleuragramma antarctica* along the Antarctic Peninsula. J Biogeogr. 2015;42(6):1103–13. doi:10.1111/jbi.12497.
6. Agostini C, Papetti C, Patarnello T, Mark F, Zane L, Marino IM. Putative selected markers in the *Chionodraco* genus detected by interspecific outlier tests. Polar Biol. 2013;36(10):1509–18. doi:10.1007/s00300-013-1370-0.
7. Fernandez-Silva I, Whitney J, Wainwright B, Andrews KR, Ylitalo-Ward H, Bowen BW, et al. Microsatellites for next-generation ecologists: a post-sequencing bioinformatics pipeline. PLoS One. 2013;8(2):e55990. doi:10.1371/journal.pone.0055990.
8. Zane L, Bargelloni L, Patarnello T. Strategies for microsatellite isolation: a review. Mol Ecol. 2002;11(1):1–16. doi:10.1046/j.0962-1083.2001.01418.x.
9. O'Brien SJ, Womack JE, Lyons LA, Moore KJ, Jenkins NA, Copeland NG. Anchored reference loci for comparative genome mapping in mammals. Nat Genet. 1993;3(2):103–12. doi:10.1038/ng0293-103.
10. Rico C, Normandeau E, Dion-Côté A-M, Rico MI, Côté G, Bernatchez L. Combining next-generation sequencing and online databases for microsatellite development in non-model organisms. Sci Rep. 2013;3:3376. doi:10.1038/srep03376.
11. Liu Z. Microsatellite markers and assessment of marker utility. In: aquaculture genome technologies. Blackwell Publishing Ltd; 2007: p. 43–58; doi: 10.1002/9780470277560.ch5.
12. MolecularEcologyResourcesPrimerDevelopmentConsortium, Agostini C, Agudelo PA, BÂ K, Barber PA, Bisol PM, et al. Permanent Genetic Resources added to Molecular Ecology Resources Database 1 October 2010–30 November 2010. Mol Ecol Resour. 2011;2:418–21. doi:10.1111/j.1755-0998.2010.02970.x.
13. Papetti C, Harms L, Windisch HS, Frickenhaus S, Sandersfeld T, Jürgens J, et al. A first insight into the spleen transcriptome of the notothenioid fish *Lepidonotothen nudifrons*: resource description and functional overview. Mar Genomics. 2015;24(3):237–9. doi:10.1016/j.margen.2015.06.006.
14. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20. doi:10.1093/bioinformatics/btu170.
15. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 2011;29(7):644–52. doi:10.1038/nbt.1883.
16. Kofler R, Schlötterer C, Lelley T. SciRoKo: a new tool for whole genome microsatellite search and investigation. Bioinformatics. 2007;23(13):1683–5. doi:10.1093/bioinformatics/btm157.
17. Castoe TA, Poole AW, de Koning APJ, Jones KL, Tomback DF, Oyler-McCance SJ, et al. Rapid Microsatellite Identification from Illumina paired-end genomic sequencing in two birds and a snake. PLoS One. 2012;7(2):e30953. doi:10.1371/journal.pone.0030953.
18. Kalendar R, Lee D, Schulman AH. Java web tools for PCR, in silico PCR, and oligonucleotide assembly and analysis. Genomics. 2011;98(2):137–44. doi:10.1016/j.ygeno.2011.04.009.
19. Holleley CE, Geerts PG. Multiplex Manager 1.0: a cross-platform computer program that plans and optimizes multiplex PCR. Biotechniques. 2009;46(7):511–7. doi:10.2144/000113156.
20. Amos W, Hoffman JI, Frodsham A, Zhang L, Best S, Hill AVS. Automated binning of microsatellite alleles: problems and solutions. Mol Ecol Notes. 2007;7(1):10–4. doi:10.1111/j.1471-8286.2006.01560.x.
21. Coombs JA, Letcher BH, Nislow KH. Create: a software to create input files from diploid genotypic data for 52 genetic software programs. Mol Ecol Resour. 2008;8(3):578–80. doi:10.1111/j.1471-8286.2007.02036.x.
22. Keenan K, McGinnity P, Cross TF, Crozier WW, Prodöhl PA. DiveRsity: an R package for the estimation and exploration of population genetics parameters and their associated errors. Methods Ecol Evol. 2013;4(8):782–8. doi:10.1111/2041-210X.12067.
23. Raymond M, Rousset F. Genepop (Version 1.2): population genetics software for exact tests and ecumenicism. J Hered. 1995;86(3):248–9.
24. Kalinowski ST, Taper ML. Maximum likelihood estimation of the frequency of null alleles at microsatellite loci. Conserv Genet. 2006;7(6):991–5. doi:10.1007/s10592-006-9134-9.
25. Miller R. Simultaneous Statistical Inference, 2nd edn. Springer, New York. pp. 299. 1981, ISBN-10: 146138124X, ISBN-13: 978-1461381242.
26. Bonferroni C. Teoria statistica delle classi e calcolo delle probabilità. Pubblicazioni del R. Istituto Superiore di Scienze Economiche e Commerciali di Firenze, 1936,8:3–62
27. Gish W, States DJ. Identification of protein coding regions by database similarity search. Nat Genet. 1993;3(3):266–72. doi:10.1038/ng0393-266.
28. Chapuis M-P, Estoup A. Microsatellite null alleles and estimation of population differentiation. Mol Biol Evol. 2007;24(3):621–31. doi:10.1093/molbev/msl191.
29. Carlsson J. Effects of microsatellite null alleles on assignment testing. J Hered. 2008;99(6):616–23. doi:10.1093/jhered/esn048.
30. Wahlund S. Zusammensetzung von Populationen und Korrelationserscheinungen vom Standpunkt der Vererbungslehre aus betrachtet. Hereditas. 1928;11(1):65–106. doi:10.1111/j.1601-5223.1928.tb02483.x.
31. Waples RS. Testing for Hardy–Weinberg proportions: have we lost the plot? J Hered. 2015;106(1):1–19. doi:10.1093/jhered/esu062.
32. Dharmarajan G, Beatty WS, Rhodes OE. Heterozygote deficiencies caused by a Wahlund effect: dispelling unfounded expectations. J Wildl Manage. 2013;77(2):226–34. doi:10.1002/jwmg.458.
33. Postolache D, Leonarduzzi C, Piotti A, Spanu I, Roig A, Fady B, Roschanski A, Liepelt S, Vendramin GG. Transcriptome versus genomic microsatellite markers: highly informative multiplexes for genotyping *Abies alba* Mill. and congeneric species. Plant Mol Biol Rep. 2013;32(3):750–60. doi:10.1007/s11105-013-0688-7.
34. Papetti C, Windisch HS, La Mesa M, Lucassen M, Marshall C, Lamare MD. Non-Antarctic notothenioids: past phylogenetic history and contemporary phylogeographic implications in the face of environmental changes. Mar Genomics. 2016;25:1–9. doi:10.1016/j.margen.2015.11.007.