Conservation Genet Resour (2017) 9:587–589
DOI 10.1007/s12686-017-0728-y

TECHNICAL NOTE

Single nucleotide polymorphism markers for analysis of historical and contemporary samples of Arctic char (*Salvelinus alpinus*)

Magnus W. Jacobsen[1] · Camilla Christensen[1] · Rikke Madsen[1] · Rasmus Nygaard[2] · Bjarni Jónsson[3] · Kim Præbel[4] · Michael M. Hansen[1]

Received: 9 March 2017 / Accepted: 10 March 2017 / Published online: 18 March 2017

Abstract Arctic regions are predicted to experience drastic temperature increases in the future, which may impact species negatively by changing their environments and habitats. One such species is the Arctic Char (*Salvelinus alpinus*), which is an important target for local fisheries and the top predator in many freshwater systems. Here we present 53 SNPs for population genetic inferences in the species. They are tested in contemporary and historical samples and will provide a useful resource for future studies of genetic structure and for genetic monitoring of Arctic char populations.

Keywords SNPs · ddRAD sequencing · Arctic char · *Salvelinus alpinus* · Genetic monitoring · Historical samples

Introduction

Anthropogenic climate change is considered one of the most severe challenges to wild plants and animals (Bellard et al. 2012), and it is predicted that Arctic regions will experience the most drastic temperature increases (Hinzman et al. 2005). Arctic char (*Salvelinus alpinus*) is a salmonid fish that is distributed across cold regions of the Northern Hemisphere. The species is considered significantly vulnerable to the effects of climate change, and it has been estimated that by 2100 it may have lost >70% of its range in Sweden (Hein et al. 2012). In Arctic regions like Greenland and Canada the species is an important target for local fisheries and is often the top predator in lakes and rivers. Due to its ecological and economic importance, monitoring efforts are warranted that can readily detect negative changes in key demographic parameters. Genetic monitoring (Schwartz et al. 2007) fulfills these requirements, and if historical samples are available this furthermore allows for retrospective monitoring (Nielsen and Hansen 2008). Here we present a resource of SNPs (Single nucleotide polymorphisms), identified using ddRAD (double digest Restriction site Associated DNA) sequencing, and adapted for the Fluidigm Biomark platform (Fluidigm Corporation, San Francisco, CA, USA). We validate the markers by analyzing contemporary and historical samples of anadromous and landlocked populations from Greenland, Iceland and Norway.

Materials and methods

ddRAD sequencing

ddRAD sequencing (Peterson et al. 2012) using the restriction enzymes SbfI and MspI was conducted for five populations from the Nuuk Fjord region, Greenland (Fig. 1) and for six family crosses involving parent fish from one of the populations, Kobbefjord. Ilumina HiSeq2000 sequencing

Electronic supplementary material The online version of this article (doi:10.1007/s12686-017-0728-y) contains supplementary material, which is available to authorized users.

✉ Magnus W. Jacobsen
magnus.jacobsen@bios.au.dk

[1] Department of Bioscience, Aarhus University, Ny Munkegade 114, 8000 Aarhus C, Denmark

[2] Greenland Institute of Natural Resources, Kivioq 2, P.O. Box 570, 3900 Nuuk, Greenland

[3] Northwest Iceland Nature Research Centre, Adalgata 2, 550 Saudárkrókur, Iceland

[4] UiT The Arctic University of Norway, 9037 Tromsø, Norway

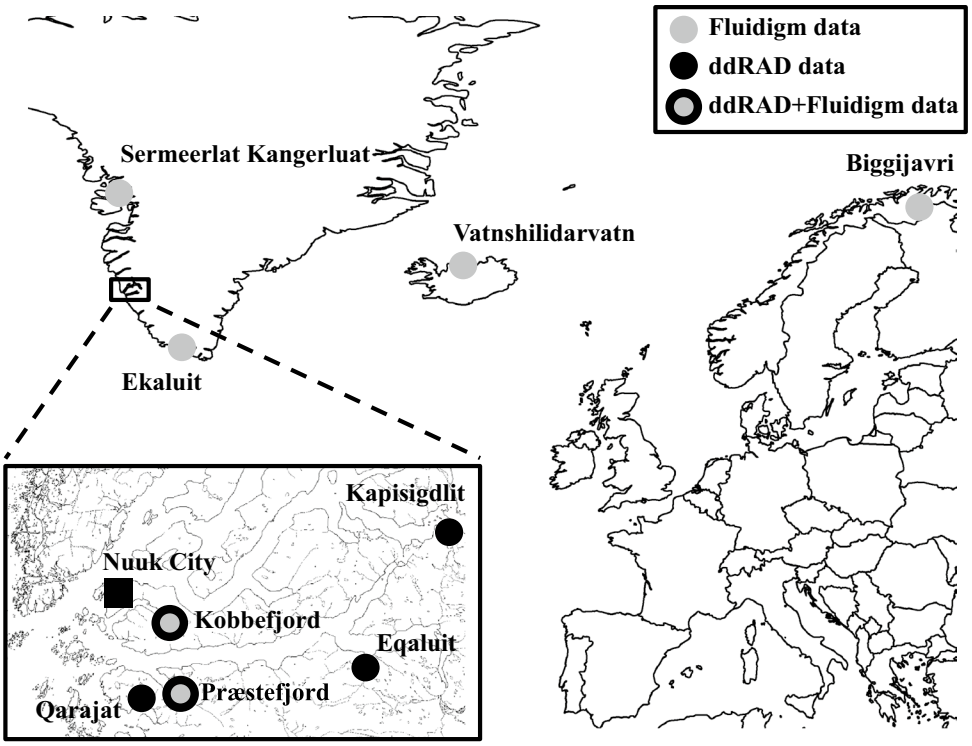

Fig. 1 Overview over the populations used for identifying SNPs using ddRAD sequencing and for subsequent validation of SNPs using the Fluidigm Biomark platform

(100 bp paired-end) of pooled multiplexed libraries (≈100 uniquely barcoded individuals per library) was outsourced to Beijing Genomic Institute (HongKong, China). Details on samples, DNA extraction and the protocol for ddRAD sequencing are provided in Notes 1 and 2 in Supporting Information.

SNP identification

Sequences and associated SNPs for Fluidigm genotyping were identified in the ddRAD dataset. In short, a "synthetic genome" was constructed based on RAD loci in parents of the family crosses by making use of overlap between paired reads. Subsequently, all reads from populations and family crosses were aligned to the "synthetic genome". SNPs were subsequently called and filtered in order to remove sequencing errors and paralogs, the latter being of special importance in salmonids due to ancient tetraploidy (Allendorf et al. 2015). Finally, SNPs with Minor Allele Frequency (MAF) <0.05 were removed and only SNPs with minimum 50 bp flanking sequence on each side were retained. Mendelian inheritance of SNPs was assessed based on the family crosses, but 31 additional SNPs not represented in the family dataset were also incorporated. Two additional SNPs were included that represent a non-synonymous substitution and variation in an untranslated region, respectively, in the immune-related Cathelicidin 2 gene (Kapralova et al. 2013) (GenBank accession numbers KC590659 and KC596075). Finally, primers were designed using the software D3 (https://d3.fluidigm.com). More details about the pipeline and sequences are provided in Supporting Information Note 3.

FLUIDIGM genotyping, scoring and testing of SNP panel

The SNP panel was tested in one historical (DNA extracted from otoliths collected in 1952) and three contemporary population samples from Greenland, and two landlocked populations from Iceland and Norway, respectively (see Fig. 1; Table 1). SNPs were genotyped on a 96.96 Dynamic Array (Fluidigm Corporation, San Francisco, CA, USA) using the Fluidigm EP1 instrumentation according to the manufacturer's recommendations. Observed and expected heterozygosity and Hardy–Weinberg equilibrium were analyzed using Genepop 4.6 (Rousset 2008).

Results and discussion

Of the 96 candidate SNPs, 53 proved useful for analysis on the Fluidigm platform (sequences provided in Supporting Information, Note 4). Among the discarded 43 SNPs, 14 showed no variation, 9 provided too little separation of genotypes and 20 showed highly significant heterozygote excess, with either all individuals being homozygotes (11) or with many heterozygotes and only one of the two homozygote classes present (9). In total 12 of these loci

Table 1 Information of average observed (H_o) and expected (H_e) heterozygosity within each of the six population analyzed for the 53 SNPs

Population	Country	Ecotype	No. of individuals	No. of variable loci	Average H_o	Average H_e
Biggijavri	Norway	Landlocked	16	12	0.0602	0.0786
Vatnshilidarvatn	Iceland	Landlocked	20	13	0.0778	0.0772
Sermeerlat Kangerluat	Greenland	Anadromous	20	43	0.2669	0.2638
Kobbefjord	Greenland	Anadromous	20	51	0.2746	0.2904
Præstefjord	Greenland	Anadromous	20	49	0.2796	0.2972
Ekaluit[a]	Greenland	Anadromous	22	46	0.1895	0.1993

[a]Historical sample from 1952, consisting of DNA extracted from otoliths

showed Mendelian inheritance in crosses while eight represented randomly chosen SNPs (not significantly different from random expectations: $\chi^2 = 1.357$; $0.10 < P < 0.25$; $df = 1$). We interpret this to represent paralogs, as the short primers for SNP analysis may in some cases target two or more paralogous loci, resulting in multi-site variants (MSVs).

Four significant deviations from Hardy–Weinberg equilibrium were observed across the 53 SNPs and and six populations (Supporting Information, Table S1). Three involved heterozygote excess in Cathelicidin 2 and might reflect selection. Genotypes at the two Cathelicidin SNPs were nearly always in concordance, indicating strong linkage. Hence, it is preferable only to include one of them in population analyses. Heterozygosity was quite similar across anadromous populations, but considerably lower and with a much lower proportion of polymorphic loci in the two landlocked populations (Table 1). This reflects well-established patterns of lower variation in freshwater as compared to anadromous fish populations (DeWoody and Avise 2000).

The SNPs worked well for analyzing historical samples (Table 1, Supporting Information S1). This opens the door for not only using the markers in studies of contemporary genetic structure, but also to estimate changes over time of demographic parameters such as effective population size (Schwartz et al. 2007; Nielsen and Hansen 2008), thereby providing important tools for monitoring Arctic char populations and the extent to which they are negatively affected by climate change.

Acknowledgements We thank the Danish Council for Independent Research, Natural Sciences for funding (Grant No. 1323-00158A to MMH).

References

Allendorf FW, Bassham S, Cresko WA, Limborg MT, Seeb LW, Seeb JE (2015) Effects of crossovers between homeologs on inheritance and population genomics in polyploid-derived salmonid fishes. J Hered 106:217–227

Bellard C, Bertelsmeier C, Leadley P, Thuiller W, Courchamp F (2012) Impacts of climate change on the future of biodiversity. Ecol Lett 15:365–377

DeWoody JA, Avise JC (2000) Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. J Fish Biol 56:461–473

Hein CL, Ohlund G, Englund G (2012) Future distribution of Arctic Char *Salvelinus alpinus* in Sweden under climate change: effects of temperature, lake size and species interactions. Ambio 41:303–312

Hinzman LD, Bettez ND, Bolton WR, Chapin FS, Dyurgerov MB, Fastie CL, Griffith B, Hollister RD, Hope A, Huntington HP, Jensen AM, Jia GJ, Jorgenson T, Kane DL, Klein DR, Kofinas G, Lynch AH, Lloyd AH, McGuire AD, Nelson FE, Oechel WC, Osterkamp TE, Racine CH, Romanovsky VE, Stone RS, Stow DA, Sturm M, Tweedie CE, Vourlitis GL, Walker MD, Walker DA, Webber PJ, Welker JM, Winker K, Yoshikawa K (2005) Evidence and implications of recent climate change in northern Alaska and other arctic regions. Clim Change 72:251–298

Kapralova KH, Gudbrandsson J, Reynisdottir S, Santos CB, Baltanas VC, Maier VH, Snorrason SS, Palsson A (2013) Differentiation at the MHCII alpha and Cath2 loci in sympatric *Salvelinus alpinus* resource morphs in Lake Thingvallavatn. PLoS ONE, 8, e69402

Nielsen EE, Hansen MM (2008) Waking the dead: the value of population genetic analyses of historical samples. Fish Fish 9:450–461

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for De Novo SNP discovery and genotyping in model and non-model species. PLoS ONE, 7, e37135

Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Mol Ecol Resour, 8, 103–106.

Schwartz MK, Luikart G, Waples RS (2007) Genetic monitoring as a promising tool for conservation and management. Trends Ecol Evol 22:25–33