Mediterranean Marine Science

Vol 21, No 2 (2020)

Vol 21, n2

Development and potential application of new set of Atlantic bluefin tuna EST-SSRs in the survival success during farming cycle

IVANA RADONIĆ, ŽELJKA TRUMBIĆ, TANJA ŠEGVIĆ- BUBIĆ, LEON GRUBIŠIĆ, IVONA MLADINEO

doi: 10.12681/mms.19025

To cite this article:

RADONIĆ, I., TRUMBIĆ, ŽELJKA, ŠEGVIĆ- BUBIĆ, T., GRUBIŠIĆ, L., & MLADINEO, I. (2020). Development and potential application of new set of Atlantic bluefin tuna EST-SSRs in the survival success during farming cycle. *Mediterranean Marine Science*, *21*(2), 298–307. https://doi.org/10.12681/mms.19025

Supplementary Data

Mediterranean Marine Science
Indexed in WoS (Web of Science, ISI Thomson) and SCOPUS
The journal is available on line at http://www.medit-mar-sc.net
DOI: http://dx.doi.org/10.12681/mms.19025

Development and potential application of a new set of Atlantic bluefin tuna EST-SSRs in survival success during the farming cycle

Ivana RADONIĆ, Željka TRUMBIĆ, Tanja ŠEGVIĆ- BUBIĆ, Leon GRUBIŠIĆ and Ivona MLADINEO

Mediterranean Marine Science, 2020, 21 (2)

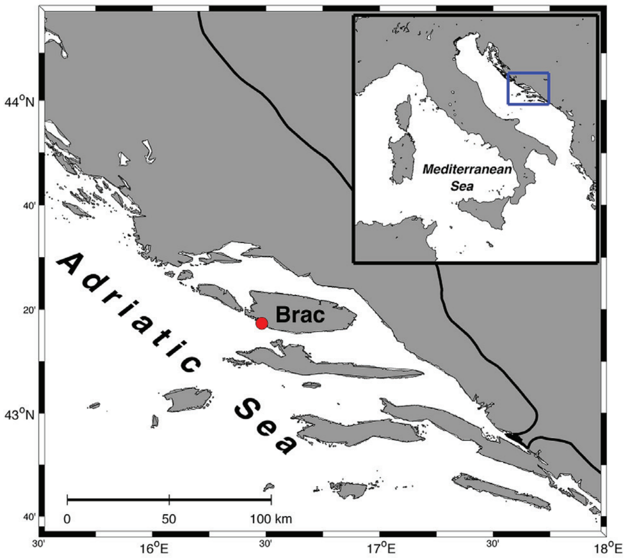

Fig. S1: Geographical location of the tuna facility in the Adriatic Sea (Croatia).

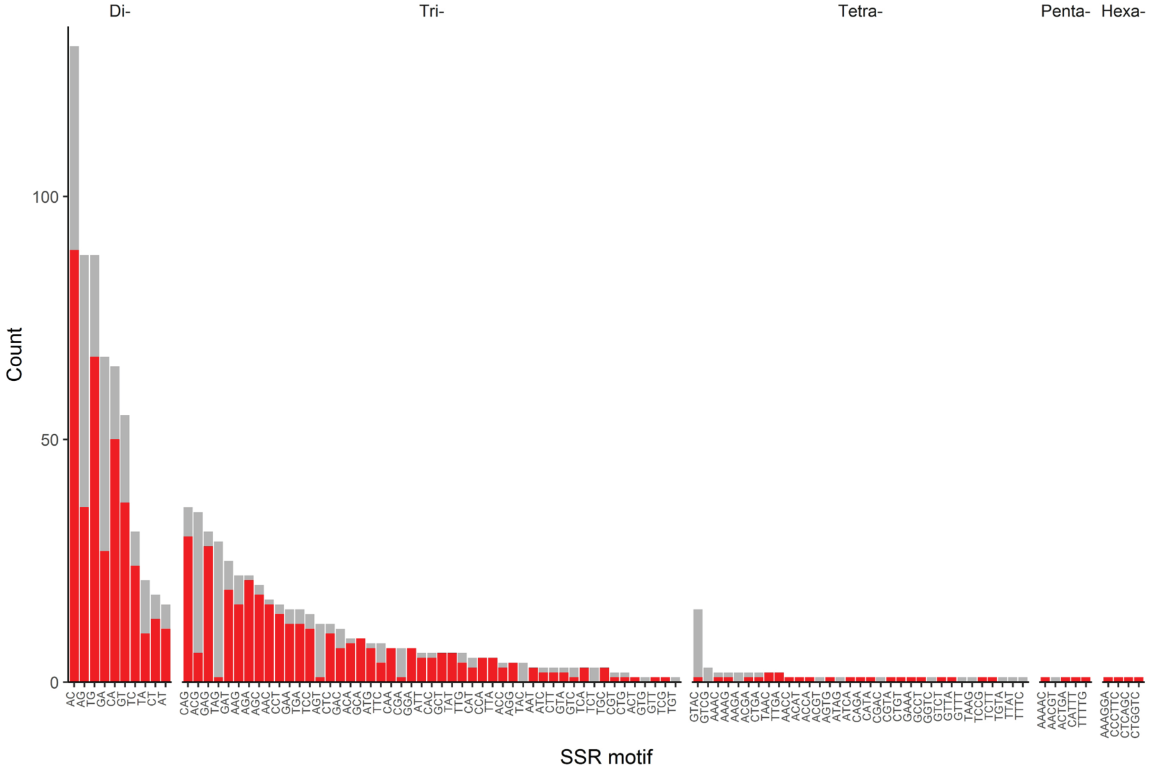

Fig. S2: Most common EST-SSR motifs found for the Atlantic bluefin tuna *T. thynnus* per each repeat category. Red bars represent motif counts in perfect repeats only, while light grey bars represent counts of specific motifs in perfect and compound formation combined.

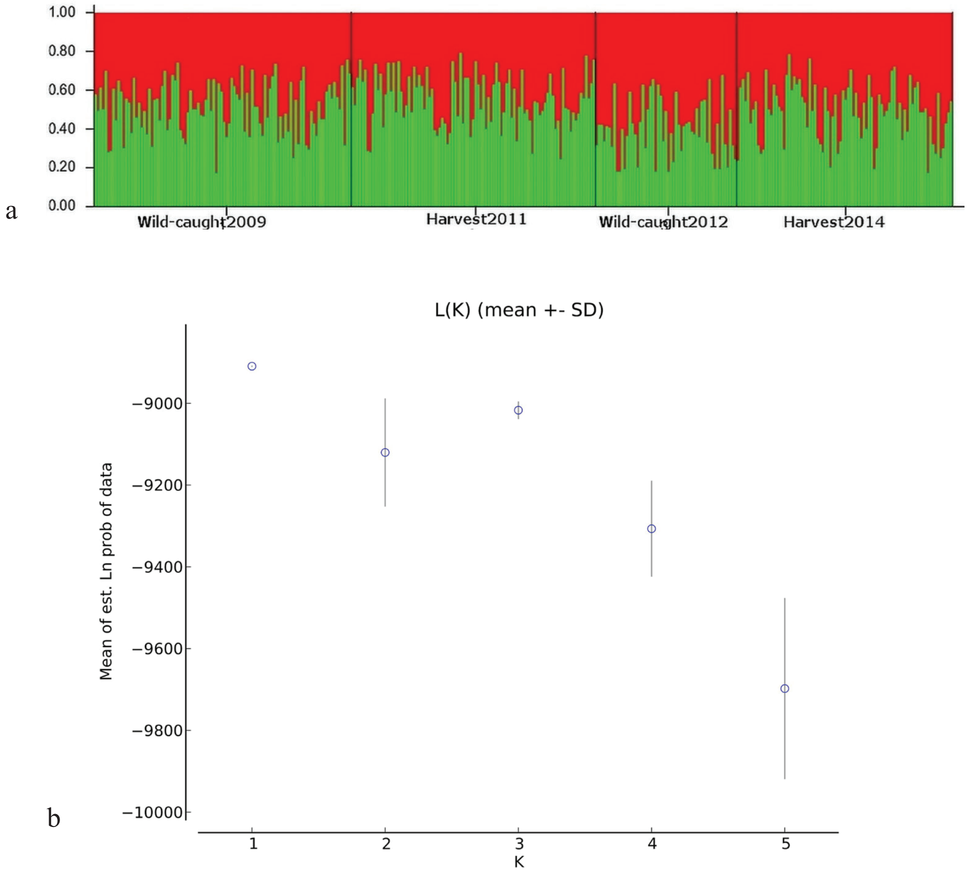

Fig. S3: Estimated population structure for the Atlantic bluefin *T. thynnus* sampled during two consecutive farming cycles at the tuna facility in the Adriatic Sea. a) Bar-plot showing the results of STRUCTURE analysis. Each individual of *T. thynnus* is represented by a vertical line, and the colour indicates the coefficient of the relationship (scale at left) to each cluster (K); b) Plot of the Log posterior probability of data for each value of K.

Table S1. Parameters of genetic diversity of four Atlantic bluefin tuna *T. thynnus* groups sampled during two consecutive farming cycles at the tuna facility in the Adriatic Sea.

a)	Wild-caught2009 (N = 100)							Harvest2011 (N = 95)						
Locus	**N**	**H_o**	**H_e**	**N_{pr}**	**A_r**	**F_{IS}**	**HWEP**	**N**	**H_o**	**H_e**	**N_{pr}**	**A_r**	**F_{IS}**	**HWEP**
ATT1	8	0.440	0.441	1	7.51	0.002	NS	7	0.547	0.569	0	6.38	0.039	NS
ATT2	3	0.280	0.282	0	3.00	0.008	NS	4	0.221	0.213	1	3.50	-0.037	NS
ATT3	5	0.950	0.654	0	4.73	-0.454^x	*	5	0.947	0.641	0	4.88	-0.480^x	*
ATT5	3	0.110	0.104	0	2.47	-0.048	NS	3	0.115	0.111	0	2.87	-0.043	NS
ATT11	4	0.616	0.569	1	3.48	-0.082	NS	4	0.652	0.569	1	3.50	-0.147	NS
ATT12	5	0.710	0.714	0	4.99	0.006	*	5	0.715	0.682	0	4.99	-0.049	*
ATT13	18	0.533	0.871	1	15.56	0.389$^+$	*	18	0.717	0.909	1	16.52	0.212$^+$	*
ATT15	2	0.110	0.104	0	1.99	-0.053	NS	4	0.157	0.176	2	3.47	0.107	NS
ATT18	10	0.744	0.772	0	8.94	0.035	NS	10	0.768	0.741	0	9.01	-0.037	NS
ATT19	3	0.300	0.325	0	2.73	0.078	NS	3	0.336	0.387	0	2.88	0.131	NS
ATT22	6	0.434	0.419	0	5.33	-0.035	NS	7	0.578	0.505	1	5.98	-0.146	NS
ATT23	4	0.535	0.554	0	3.98	0.034	NS	5	0.673	0.541	1	4.23	-0.245	NS
ATT25	6	0.660	0.735	1	5.46	0.104	NS	6	0.831	0.739	1	5.44	-0.125	NS
Mean ± SD	5.92 ± 4.25	0.49 ± 0.25	0.50 ± 0.25	0.30 ± 0.48	5.41 ± 3.69	0.019 ± 0.18		6.23 ± 4.02	0.55 ± 0.26	0.52 ± 0.24	0.61 ± 0.65	5.67 ± 3.68	-0.070 ± 0.17	

b)	Wild-caught2012 (N = 55)							Harvest2014 (N = 84)						
Locus	**N**	**H_o**	**H_e**	**N_{pr}**	**A_r**	**F_{IS}**	**HWEP**	**N**	**H_o**	**H_e**	**N_{pr}**	**A_r**	**F_{IS}**	**HWEP**
ATT1	7	0.236	0.293	0	6.60	0.196	NS	7	0.488	0.530	0	6.48	0.080	NS
ATT2	3	0.111	0.107	0	2.98	-0.036	NS	3	0.261	0.236	0	3.00	-0.106	NS
ATT3	6	0.945	0.628	1	5.87	-0.512^x	*	5	0.976	0.709	0	5.00	-0.378^x	*
ATT5	1	/	/	/	/	/	/	3	0.083	0.080	/	2.56	-0.032	NS
ATT11	3	0.818	0.576	0	3.00	-0.424^x	*	4	0.785	0.622	1	3.81	-0.265^x	NS
ATT12	5	0.781	0.655	1	4.87	-0.194	*	6	0.630	0.603	2	5.14	-0.046	NS
ATT13	13	0.875	0.840	1	13.0	-0.042	NS	15	0.786	0.860	3	14.53	0.086	*
ATT15	2	0.072	0.070	0	2.00	-0.029	NS	3	0.059	0.125	1	2.98	0.526$^+$	*
ATT18	9	0.796	0.782	0	8.76	-0.018	*	9	0.761	0.738	0	8.07	-0.032	NS
ATT19	2	0.327	0.276	0	2.00	-0.187	NS	3	0.369	0.331	1	2.92	-0.114	NS
ATT22	4	0.647	0.492	0	4.00	-0.319	NS	6	0.535	0.470	2	5.29	-0.140	NS
ATT23	4	0.784	0.541	0	3.99	-0.456^x	*	5	0.674	0.563	1	4.54	-0.199	NS
ATT25	4	0.854	0.672	0	4.00	-0.274	*	5	0.833	0.739	1	5.00	-0.128	NS
Mean ± SD	4.84 ± 3.28	0.60 ± 0.32	0.49 ± 0.25	0.25 ± 0.45	4.77 ± 3.23	-0.224 ± 0.21		5.69 ± 3.32	0.55 ± 0.29	0.50 ± 0.24	1.00 ± 0.95	5.33 ± 2.95	-0.097 ± 0.12	

Note: ATT - Adriatic *Thunnus thynnus*; N - number of alleles per locus; H_o - observed heterozygosity; H_e - expected heterozygosity; N_{pr} - number of private alleles; A_r - allelic richness; F_{IS} - fixation index; HWEP - result for Hardy-Weinberg Equilibrium test; NS = non-significant; * p < 0.001 after sequential Bonferroni correction; $^{+/x}$, p-value for F_{IS} within samples based on, 1040 randomisations. Indicative adjusted nominal level (5%) for one table = 0.00096. $^+$ - Proportion of randomisations that gave a larger F_{IS} than the observed, ^{X} - Proportion of randomisations that gave a smaller F_{IS} than the observed.

Table S2. Pairwise F_{ST} values among two consecutive farming cycles of the Atlantic bluefin tuna *T. thynnus* sampled at the tuna facility in the Adriatic Sea.

Population	Wild-caught 2009	Harvest 2011	Wild-caught 2012
Harvest 2011	0.00097		
Wild-caught 2012	0.00344	<u>0.00945</u>	
Harvest 2014	0.00238	0.00115	<u>0.00603</u>

Note: significant F_{ST} values are underlined (following Bonferroni correction).

Table S3. Results of BayeScan analysis performed to identify the newly developed EST-SSR loci under natural selection.

Loci	prob	Log10(PO)	qval	alpha	fst
ATT1	0.061	-1.187	0.814	-0.004	0.009
ATT2	0.084	-1.035	0.781	-0.029	0.009
ATT3	0.090	-1.004	0.725	-0.055	0.009
ATT5	0.089	-1.007	0.748	-0.026	0.009
ATT11	0.086	-1.025	0.766	-0.000	0.010
ATT12	0.081	-1.052	0.804	-0.025	0.009
ATT13	0.948	1.264	0.051	-1.552	0.002
ATT15	0.118	-0.873	0.489	0.082	0.011
ATT18	0.466	-0.058	0.292	-0.664	0.006
ATT19	0.095	-0.976	0.694	-0.030	0.009
ATT22	0.101	-0.948	0.652	-0.078	0.009
ATT23	0.084	-1.036	0.793	-0.039	0.009
ATT25	0.105	-0.930	0.590	-0.071	0.009

Note: prob - the posterior probability; log10(PO) - the logarithm of Posterior Odds; qval - FDR analogue of the p-value; alpha - locus specific component (negative alpha suggests balancing selection, while positive alpha suggests directional selection); Fst - Fixation Indices subpopulation to total population.

Table S4. Gene ontology and the description of gene function of each newly developed microsatellite marker derived from EST libraries of the Atlantic bluefin tuna *T. thynnus,* sampled during two consecutive farming cycles at the tuna facility in the Adriatic Sea.

Locus name	Description	Gene symbol	Position	Gene ontology	The function of a gene	Reference
ATT1	Solute carrier family 6 (neurotransmitter transporter, creatine), member 8	*slc6a8*	Coding	Biological Process: GO:0006810:transport; GO:0006836:neurotransmitter transport Cellular Component: GO:0005887:integral component of plasma membrane; GO:0016020:membrane; GO:0016021:integral component of membrane Molecular Function: GO:0005309:creatine:sodium symporter activity; GO:0005328:neurotransmitter: sodium symporter activity; GO:0015293:symporter activity	Encodes for sodium- and chloride- dependent creatine transporter 1 which transports creatine into and out of the target cells. Creatine acts as an energy buffer by its involvement in maintenance of adequate ATP levels, hence can be found in tissues with high energy demand and flux, such as skeletal muscles or brain.	(Borchel *et al.*, 2014; Brosnan & Brosnan, 2016)
ATT2	Nicotinamide phosphoribosyltransferase-like	*Nampta*	Coding	Biological Process: GO:0009435:NAD biosynthetic process Molecular Function: GO:0004514:nicotinate-nucleotide diphosphorylase (carboxylating) activity	Encodes a protein that participates in the biosynthesis of nicotinamide adenine dinucleotide (NAD$^+$) from nicotinamide. It belongs to the nicotinamide phosphoribosyltransferase (NAMPT) family, also known as visfatin/pre-B-cell colony-enhancing factor. It has ubiquitous expression and a broad range of functions associated with NAD$^+$ levels, from cellular viability to stress and inflammatory response.	(Rongvaux *et al.*, 2008; Fang *et al.*, 2015)"container-title":"Journal of Immunology (Baltimore, Md.: 1950
ATT3	Sodium channel, voltage-gated, type IV, beta a	*scn4ba*	5'	Biological Process: GO:0006810:transport; GO:0006811:ion transport Cellular Component: GO:0016020:membrane; GO:0016021:integral component of membrane	Encodes a protein that is one of several sodium channel beta subunits which participate in ion transport of sodium ions through the plasma membrane, mediating regenerative cell membrane depolarization and conduction of electrical signaling in nerves and muscles.	(Diss *et al.*, 2004)

continued

Table S4 continued

Locus name	Description	Gene symbol	Position	Gene ontology	The function of a gene	Reference
ATT5	Polymerase (RNA) II (DNA directed) polypeptide D	*polr2d*	Coding	Biological Process: GO:0000288:nuclear-transcribed mRNA catabolic process, deadenylation-dependent decay, GO:0006351:transcription, DNA-templated; GO:0006367:transcription initiation from RNA polymerase II promoter GO:0031990:mRNA export from nucleus in response to heat stress GO:0034402:recruitment of 3'-end processing factors to RNA polymerase II holoenzyme complex; GO:0044237:cellular metabolic process; GO:0045948:positive regulation of translational initiation Cellular Component: GO:0000932:P-body; GO:0005622:intracellular; GO:0005665:DNA-directed RNA polymerase II, core complex Molecular Function: GO:0000166:nucleotide binding; GO:0003676:nucleic acid binding; GO:0003697:single-stranded DNA binding; GO:0003727:single-stranded RNA binding; GO:0003824:catalytic activity; GO:0003899:DNA-directed 5'-3' RNA polymerase activity; GO:0031369:translation initiation factor binding	Encodes the fourth largest subunit of RNA polymerase II, the polymerase responsible for synthesizing messenger RNA in eukaryotes.	(Cramer, 2001)
ATT11	Complement factor I	*cfi*	3'	Biological Process: BP:GO:0006508:proteolysis BP:GO:0061304:retinal blood vessel morphogenesis Cellular Component: GO:0016020:membrane Molecular Function: GO:0004252:serine-type endopeptidase activity GO:0005044:scavenger receptor activity; GO:0008233:peptidase activity GO:0008236:serine-type peptidase activity; GO:0016787:hydrolase activity	The orthologue of this gene encodes a serine proteinase that is essential for regulating the complement cascade. This heterodimeric glycoprotein can cleave and inactivate the complement components C4b and C3b, thus inhibiting the amplification loop of the alternative pathway of complement activation. Complement system is an important part of innate immune system, also acting as an effector mechanism for the adaptive immune response.	(Holland & Lambris, 2002; Fremeaux-Bacchi *et al.*, 2004; Fagerness *et al.*, 2009)

continued

Table S4 continued

Locus name	Description	Gene symbol	Position	Gene ontology	The function of a gene	Reference
ATT12	Mitogen-activated protein kinase 6	*mapk6*	5’	Biological Process: GO:0006468:protein phosphorylation; GO:0010468:regulation of gene expression; GO:0016310:phosphorylation; GO:0071310:cellular response to organic substance Cellular Component: GO:0005634:nucleus; GO:0005737:cytoplasm Molecular Function: GO:0004672:protein kinase activity; GO:0004707:MAP kinase activity; GO:0005524:ATP binding; GO:0016301:kinase activity; GO:0016740:transferase activity	Encodes the protein that is a member of the Ser/Thr protein kinase family, member of MAPK sequential phosphorylation signaling cascades. They are involved in diverse cellular processes such as growth, proliferation, differentiation, survival, regulation of inflammation and development.	(Krens *et al.*, 2006; Arthur & Ley, 2013)
ATT13	Cathepsin F	*ctsf*	3’	Biological Process: GO:0006508:proteolysis; GO:0051603:proteolysis involved in cellular protein catabolic process Cellular Component: GO:0005615:extracellular space; GO:0005764:lysosome Molecular Function: GO:0004197:cysteine-type endopeptidase activity; GO:0004869:cysteine-type endopeptidase inhibitor activity; GO:0008234:cysteine-type peptidase activity	A member of papain family cysteine proteinases that represent a major component of the lysosomal proteolytic system. In fish, cathepsin F is also known to confer important role in oocyte maturation, growth and development as well as immune system (antigen-presenting).	(Kim *et al.*, 2010; Lee *et al.*, 2013)
ATT15	NFKB activating protein-like	*nkap*	Coding	Biological Process: GO:0007219:Notch signaling pathway GO:0045892:negative regulation of transcription Molecular Function: GO:0003682:chromatin binding DNA-templated	This gene encodes a protein that acts as activator of the ubiquitous transcription factor NF-kappaB, regulates the Notch signaling pathway and is required for T-cell development. Additionally, a role of murine NKAP is predicted in splicing and RNA metabolism.	(Burgute *et al.*, 2014)

continued

Table S4 continued

Locus name	Description	Gene symbol	Position	Gene ontology	The function of a gene	Reference
ATT18	Growth factor receptor-bound protein 2b	*grb2b*	3’	Biological Process: GO:0007169:transmembrane receptor protein tyrosine kinase signaling pathway; GO:0016477:cell migration; GO:0030154:cell differentiation; GO:0038083:peptidyl-tyrosine autophosphorylation; GO:0042127:regulation of cell proliferation; GO:0045087:innate immune response Cellular Component: GO:0031234:extrinsic component of cytoplasmic side of plasma membrane Molecular Function: GO:0004715:non-membrane spanning protein tyrosine kinase activity; GO:0005102:receptor binding	An adaptor protein involved in signal transduction and cell communication. It is ubiquitously expressed and activates Ras and MAP kinases in growth factor receptor signaling with many important physiological roles, including lymphocyte differentiation, as well as control of secondary humoral immune responses.	(Ackermann *et al.*, 2011)
ATT19	TRAF2 and NCK interacting kinase a	*tnika*	3’	Biological Process: GO:0006468:protein phosphorylation; GO:0007346:regulation of mitotic cell cycle; GO:0007369:gastrulation; GO:0008360:regulation of cell shape GO:0016310:phosphorylation; BP:GO:0031098:stress-activated protein kinase signaling cascade BP:GO:0042981:regulation of apoptotic process BP:GO:0048812:neuron projection morphogenesis BP:GO:0060027:convergent extension involved in gastrulation Cellular Component: GO:0005737:cytoplasm Molecular Function: GO:0004672:protein kinase activity; GO:0005524:ATP binding; GO:0008349:MAP kinase kinase kinase kinase activity; GO:0016301:kinase activity; GO:0016740:transferase activity	Belongs to the Germinal Center Kinases (GCKs) that can activate c-Jun N-terminal kinase pathway or associate with Rap2 to regulate cytoskeleton organization and cell spreading. Also implicated in neuronal dendrite extension. New evidence indicate its essential role in colon cancer growth through aberrant WNT signaling pathway.	(Yu *et al.*, 2014)

continued

Table S4 continued

Locus name	Description	Gene symbol	Position	Gene ontology	The function of a gene	Reference
ATT22	Peroxisome proliferator-activated receptor delta a	*pparda*	Coding	Biological Process: GO:0006351:transcription, DNA-templated; GO:0006355:regulation of transcription, DNA-templated; GO:0014032:neural crest cell development GO:0014823:response to activity; GO:0043401:steroid hormone mediated signaling pathway Cellular Component: GO:0000785:chromatin; GO:0005634:nucleus Molecular Function: GO:0003677:DNA binding; GO:0003700:DNA binding transcription factor activity; GO:0003707:steroid hormone receptor activity; GO:0004879:nuclear receptor activity; GO:0008270:zinc ion binding GO:0043565:sequence-specific DNA binding; GO:0046872:metal ion binding	Encodes a member of the peroxisome proliferator-activated receptor (PPAR) family, nuclear hormone receptors, PPARs are important factors in maintenance of energy homeostasis and regulators of lipid metabolism. Pparda is a zebrafish orthologue with possible role in early adipocyte development.	(Den Broeder *et al.*, 2015)
ATT23	SHC-transforming protein 1-like	*sch1*	3’	Biological Process: GO:0001525:angiogenesis; GO:0007165:signal transduction GO:0007169:transmembrane receptor protein tyrosine kinase signaling pathway; GO:0035556:intracellular signal transduction Cellular Component: GO:0005886:plasma membrane Molecular Function: GO:0030971:receptor tyrosine kinase binding	Encodes SHC protein that appears in three isoforms in humans. They are adapter proteins in signal transduction pathways and participate in the activation of Ras and MAPKs downstream of many receptors for various growth factors, cytokines and hormones. Shc mediates early angiogenesis by integrating signals of integrins and VEGF.	(Sweet *et al.*, 2012)
ATT25	Low density lipoprotein receptor a	*ldlra*	Coding	Biological Process: GO:0090118:receptor-mediated endocytosis involved in cholesterol transport Cellular Component: GO:0016020:membrane GO:0016021:integral component of membrane Molecular Function: GO:0005041:low-density lipoprotein receptor activity GO:0005509:calcium ion binding	LDLR gene family consists of cell surface proteins involved in receptor-mediated endocytosis and lysosomal delivery of specific ligands (lipoproteins, proteases and protease inhibitors) that bind to their extracellular domain.	(Gotthardt *et al.*, 2000)

continued

Table S4 continued

Locus name	Description	Gene symbol	Position	Gene ontology	The function of a gene	Reference
Locus name	Description	Gene symbol	Position	Gene ontology	The function of a gene	Reference
ATT1	Solute carrier family 6 (neurotransmitter transporter, creatine), member 8	*slc6a8*	Coding	Biological Process: GO:0006810:transport; GO:0006836:neurotransmitter transport Cellular Component: GO:0005887:integral component of plasma membrane; GO:0016020:membrane; GO:0016021:integral component of membrane Molecular Function: GO:0005309:creatine:sodium symporter activity; GO:0005328:neurotransmitter: sodium symporter activity; GO:0015293:symporter activity	Encodes for sodium- and chloride- dependent creatine transporter 1 which transports creatine into and out of the target cells. Creatine acts as an energy buffer by its involvement in maintenance of adequate ATP levels, hence can be found in tissues with high energy demand and flux, such as skeletal muscles or brain.	(Borchel *et al.*, 2014; Brosnan & Brosnan, 2016)
ATT2	Nicotinamide phosphoribosyltransferase-like	*Nampt a*	Coding	Biological Process: GO:0009435:NAD biosynthetic process Molecular Function: GO:0004514:nicotinate-nucleotide diphosphorylase (carboxylating) activity	Encodes a protein that participates in the biosynthesis of nicotinamide adenine dinucleotide (NAD^+) from nicotinamide. It belongs to the nicotinamide phosphoribosyltransferase (NAMPT) family, also known as visfatin/pre-B-cell colony-enhancing factor. It has ubiquitous expression and a broad range of functions associated with NAD^+ levels, from cellular viability to stress and inflammatory response.	(Rongvaux *et al.*, 2008; Fang *et al.*, 2015)"container-title":"Journal of Immunology (Baltimore, Md.: 1950
ATT3	Sodium channel, voltage-gated, type IV, beta a	*scn4ba*	5'	Biological Process: GO:0006810:transport; GO:0006811:ion transport Cellular Component: GO:0016020:membrane; GO:0016021:integral component of membrane	Encodes a protein that is one of several sodium channel beta subunits which participate in ion transport of sodium ions through the plasma membrane, mediating regenerative cell membrane depolarization and conduction of electrical signaling in nerves and muscles.	(Diss *et al.*, 2004)

continued

Table S4 continued

Locus name	Description	Gene symbol	Position	Gene ontology	The function of a gene	Reference
ATT5	Polymerase (RNA) II (DNA directed) polypeptide D	*polr2d*	Coding	Biological Process: GO:0000288:nuclear-transcribed mRNA catabolic process, deadenylation-dependent decay, GO:0006351:transcription, DNA-templated; GO:0006367:transcription initiation from RNA polymerase II promoter GO:0031990:mRNA export from nucleus in response to heat stress GO:0034402:recruitment of 3'-end processing factors to RNA polymerase II holoenzyme complex; GO:0044237:cellular metabolic process; GO:0045948:positive regulation of translational initiation Cellular Component: GO:0000932:P-body; GO:0005622:intracellular; GO:0005665:DNA-directed RNA polymerase II, core complex Molecular Function: GO:0000166:nucleotide binding; GO:0003676:nucleic acid binding; GO:0003697:single-stranded DNA binding; GO:0003727:single-stranded RNA binding; GO:0003824:catalytic activity; GO:0003899:DNA-directed 5'-3' RNA polymerase activity; GO:0031369:translation initiation factor binding	Encodes the fourth largest subunit of RNA polymerase II, the polymerase responsible for synthesizing messenger RNA in eukaryotes.	(Cramer, 2001)
ATT11	Complement factor I	*cfi*	3'	Biological Process: BP:GO:0006508:proteolysis BP:GO:0061304:retinal blood vessel morphogenesis Cellular Component: GO:0016020:membrane Molecular Function: GO:0004252:serine-type endopeptidase activity GO:0005044:scavenger receptor activity; GO:0008233:peptidase activity GO:0008236:serine-type peptidase activity; GO:0016787:hydrolase activity	The orthologue of this gene encodes a serine proteinase that is essential for regulating the complement cascade. This heterodimeric glycoprotein can cleave and inactivate the complement components C4b and C3b, thus inhibiting the amplification loop of the alternative pathway of complement activation. Complement system is an important part of innate immune system, also acting as an effector mechanism for the adaptive immune response.	(Holland & Lambris, 2002; Fremeaux-Bacchi *et al.*, 2004; Fagerness *et al.*, 2009)

continued

Table S4 continued

Locus name	Description	Gene symbol	Position	Gene ontology	The function of a gene	Reference
ATT12	Mitogen-activated protein kinase 6	*mapk6*	5’	Biological Process: GO:0006468:protein phosphorylation; GO:0010468:regulation of gene expression; GO:0016310:phosphorylation; GO:0071310:cellular response to organic substance Cellular Component: GO:0005634:nucleus; GO:0005737:cytoplasm Molecular Function: GO:0004672:protein kinase activity; GO:0004707:MAP kinase activity; GO:0005524:ATP binding; GO:0016301:kinase activity; GO:0016740:transferase activity	Encodes the protein that is a member of the Ser/Thr protein kinase family, member of MAPK sequential phosphorylation signaling cascades. They are involved in diverse cellular processes such as growth, proliferation, differentiation, survival, regulation of inflammation and development.	(Krens *et al.*, 2006; Arthur & Ley, 2013)
ATT13	Cathepsin F	*ctsf*	3’	Biological Process: GO:0006508:proteolysis; GO:0051603:proteolysis involved in cellular protein catabolic process Cellular Component: GO:0005615:extracellular space; GO:0005764:lysosome Molecular Function: GO:0004197:cysteine-type endopeptidase activity; GO:0004869:cysteine-type endopeptidase inhibitor activity; GO:0008234:cysteine-type peptidase activity	A member of papain family cysteine proteinases that represent a major component of the lysosomal proteolytic system. In fish, cathepsin F is also known to confer important role in oocyte maturation, growth and development as well as immune system (antigen-presenting).	(Kim *et al.*, 2010; Lee *et al.*, 2013)
ATT15	NFKB activating protein-like	*nkap*	Coding	Biological Process: GO:0007219:Notch signaling pathway GO:0045892:negative regulation of transcription Molecular Function: GO:0003682:chromatin binding DNA-templated	This gene encodes a protein that acts as activator of the ubiquitous transcription factor NF-kappaB, regulates the Notch signaling pathway and is required for T-cell development. Additionally, a role of murine NKAP is predicted in splicing and RNA metabolism.	(Burgute *et al.*, 2014)

continued

Table S4 continued

Locus name	Description	Gene symbol	Position	Gene ontology	The function of a gene	Reference
ATT18	Growth factor receptor-bound protein 2b	*grb2b*	3’	Biological Process: GO:0007169:transmembrane receptor protein tyrosine kinase signaling pathway; GO:0016477:cell migration; GO:0030154:cell differentiation; GO:0038083:peptidyl-tyrosine autophosphorylation; GO:0042127:regulation of cell proliferation; GO:0045087:innate immune response Cellular Component: GO:0031234:extrinsic component of cytoplasmic side of plasma membrane Molecular Function: GO:0004715:non-membrane spanning protein tyrosine kinase activity; GO:0005102:receptor binding	An adaptor protein involved in signal transduction and cell communication. It is ubiquitously expressed and activates Ras and MAP kinases in growth factor receptor signaling with many important physiological roles, including lymphocyte differentiation, as well as control of secondary humoral immune responses.	(Ackermann *et al.*, 2011)
ATT19	TRAF2 and NCK interacting kinase a	*tnika*	3’	Biological Process: GO:0006468:protein phosphorylation; GO:0007346:regulation of mitotic cell cycle; GO:0007369:gastrulation; GO:0008360:regulation of cell shape GO:0016310:phosphorylation; BP:GO:0031098:stress-activated protein kinase signaling cascade BP:GO:0042981:regulation of apoptotic process BP:GO:0048812:neuron projection morphogenesis BP:GO:0060027:convergent extension involved in gastrulation Cellular Component: GO:0005737:cytoplasm Molecular Function: GO:0004672:protein kinase activity; GO:0005524:ATP binding; GO:0008349:MAP kinase kinase kinase kinase activity; GO:0016301:kinase activity; GO:0016740:transferase activity	Belongs to the Germinal Center Kinases (GCKs) that can activate c-Jun N-terminal kinase pathway or associate with Rap2 to regulate cytoskeleton organization and cell spreading. Also implicated in neuronal dendrite extension. New evidence indicate its essential role in colon cancer growth through aberrant WNT signaling pathway.	(Yu *et al.*, 2014)

continued

Table S4 continued

Locus name	Description	Gene symbol	Position	Gene ontology	The function of a gene	Reference
ATT22	Peroxisome proliferator-activated receptor delta a	*pparda*	Coding	Biological Process: GO:0006351:transcription, DNA-templated; GO:0006355:regulation of transcription, DNA-templated; GO:0014032:neural crest cell development GO:0014823:response to activity; GO:0043401:steroid hormone mediated signaling pathway Cellular Component: GO:0000785:chromatin; GO:0005634:nucleus Molecular Function: GO:0003677:DNA binding; GO:0003700:DNA binding transcription factor activity; GO:0003707:steroid hormone receptor activity; GO:0004879:nuclear receptor activity; GO:0008270:zinc ion binding GO:0043565:sequence-specific DNA binding; GO:0046872:metal ion binding	Encodes a member of the peroxisome proliferator-activated receptor (PPAR) family, nuclear hormone receptors, PPARs are important factors in maintenance of energy homeostasis and regulators of lipid metabolism. Pparda is a zebrafish orthologue with possible role in early adipocyte development.	(Den Broeder *et al.*, 2015)
ATT23	SHC-transforming protein 1-like	*sch1*	3’	Biological Process: GO:0001525:angiogenesis; GO:0007165:signal transduction GO:0007169:transmembrane receptor protein tyrosine kinase signaling pathway; GO:0035556:intracellular signal transduction Cellular Component: GO:0005886:plasma membrane Molecular Function: GO:0030971:receptor tyrosine kinase binding	Encodes SHC protein that appears in three isoforms in humans. They are adapter proteins in signal transduction pathways and participate in the activation of Ras and MAPKs downstream of many receptors for various growth factors, cytokines and hormones. Shc mediates early angiogenesis by integrating signals of integrins and VEGF.	(Sweet *et al.*, 2012)
ATT25	Low density lipoprotein receptor a	*ldlra*	Coding	Biological Process: GO:0090118:receptor-mediated endocytosis involved in cholesterol transport Cellular Component: GO:0016020:membrane GO:0016021:integral component of membrane Molecular Function: GO:0005041:low-density lipoprotein receptor activity GO:0005509:calcium ion binding	LDLR gene family consists of cell surface proteins involved in receptor-mediated endocytosis and lysosomal delivery of specific ligands (lipoproteins, proteases and protease inhibitors) that bind to their extracellular domain.	(Gotthardt *et al.*, 2000)

References

Ackermann, J. A., Radtke, D., Maurberger, A., Winkler, T. H., Nitschke, L.,2011. Grb2 regulates B-cell maturation, B-cell-memory responses and inhibits B-cell Ca $^{2+}$ signalling: Grb2-deficient mice. *The EMBO Journal*, 30, 1621-1633.

Arthur, J.S. C., Ley, S. C., 2013. Mitogen-activated protein kinases in innate immunity. *Nature Reviews Immunology*, 13, 679-692.

Borchel, A., Verleih, M., Rebl, A., Kühn, C., Goldammer, T. 2014. Creatine metabolism differs between mammals and rainbow trout (*Oncorhynchus mykiss*). *Springer Plus*, 3, 1-9.

Brosnan, M. E., Brosnan, J. T. 2016. The role of dietary creatine. Amino Acids, 48, 1785-1791.

Burgute, B. D., Peche, V. S., Steckelberg, A. L., Glockner, G., Gassen, B., Gehring, N. H., Noegel, A. A., 2014. NKAP is a novel RS-related protein that interacts with RNA and RNA binding proteins. *Nucleic Acids Research*, 42, 3177-3193.

Cramer, P., 2001. Structural Basis of Transcription: RNA Polymerase II at 2.8 Angstrom Resolution. Science, 292, 1863-1876.

Den Broeder, M. J., Kopylova, V. A., Kamminga, L. M., Legler, J., 2015. Zebrafish as a Model to Study the Role of Peroxisome Proliferating-Activated Receptors in Adipogenesis and Obesity. *PPAR Research*, 2015, 1-11.

Diss, J. K. J., Fraser, S. P., Djamgoz, M. B. A., 2004. Voltage-gated Na+ channels: multiplicity of expression, plasticity, functional implications and pathophysiological aspects. European Biophysics Journal, 33: 180-193.

Fagerness, J. A., Maller, J. B., Neale, B. M., Reynolds, R. C., Daly, M. J., Seddon, J. M. 2009. Variation near complement factor I is associated with risk of advanced AMD. *European Journal of Human Genetics*, 17, 100-104.

Fang, C., Guan, L., Zhong, Z., Gan, X., He, S., 2015. Analysis of the nicotinamide phosphoribosyltransferase family provides insight into vertebrate adaptation to different oxygen levels during the water-to-land transition. *FEBS Journal*, 282, 2858-2878.

Fremeaux-Bacchi, V., Dragon-Durey, M.-A., Blouin, J., Vigneau, C., Kuypers, D., Boudailliez, B., Loirat, C. *et al.* 2004.

Complement factor I: a susceptibility gene for atypical haemolytic uraemic syndrome. *Journal of Medical Genetics*, 41, e84.

Gotthardt, M., Trommsdorff, M., Nevitt, M. F., Shelton, J., Richardson, J. A., Stockinger, W., Nimpf, J., *et al.* 2000. Interactions of the Low Density Lipoprotein Receptor Gene Family with Cytosolic Adaptor and Scaffold Proteins Suggest Diverse Biological Functions in Cellular Communication and Signal Transduction. *Journal of Biological Chemistry*, 275, 25616-25624.

Holland, M. C. H., Lambris, J. D., 2002. The complement system in teleosts. *Fish & Shellfish Immunology*, 12, 399,420.

Kim, Y.O., Park, E.M., Nam, B.H., Kong, H. J., Kim, W.J., Lee, S.J., Kim, K.K., 2010. Molecular cloning and characterization of cathepsin F gene from olive flounder *Paralichthys Olivaceus*. *Genes & Genomics*, 32, 137-142.

Krens, S. F. G., Spaink, H. P., Snaar-Jagalska, B.E., 2006. Functions of the MAPK family in vertebrate-development. *FEBS Letters*, 580, 4984-4990.

Lee, J.W., Lee, Y. M., Yang, H., Noh, J. K., Kim, H. C., Park, C.J., Park, J.W. *et al.* 2013. Expression Analysis of Cathepsin F during Embryogenesis and Early Developmental Stage in Olive Flounder (*Paralichthys olivaceus*). *Development & Reproduciton*, 17, 221-229.

Rongvaux, A., Galli, M., Denanglaire, S., Van Gool, F., Drèze, P. L., Szpirer, C., Bureau, F., *et al.* 2008. Nicotinamide phosphoribosyl transferase/pre-B cell colony-enhancing factor/visfatin is required for lymphocyte development and cellular resistance to genotoxic stress. *Journal of Immunology (Baltimore, Md.: 1950)*, 181, 4685-4695.

Sweet, D. T., Chen, Z., Wiley, D. M., Bautch, V.L., Tzima, E., 2012. The adaptor protein Shc integrates growth factor and ECM signaling during postnatal angiogenesis. *Blood*, 119, 1946-1955.

Yu, D.H., Zhang, X., Wang, H., Zhang, L., Chen, H., Hu, M., Dong, Z. *et al.* 2014. The essential role of TNIK gene amplification in gastric cancer growth. *Oncogenesis*, 3, 1-8.