Variation of body weight and seasonal reproduction in populations from different latitudes: genetic analyses using wild-derived medaka fish

異なる緯度に由来する集団における体重と季節繁殖の変異：野生由来メダカを利用した遺伝学的解析

IMAEDA YASSUMOTO Tamiris

Summary of doctoral dissertation submitted to the Graduate School of Bioagricultural Sciences of Nagoya University

Supervisor: Professor YOSHIMURA Takashi

In this study, I use forward genetic approaches to search for the genetic mechanisms responsible for different body weights and suppression of reproduction under short days in medaka fish from different latitudes. Medaka fish (*Oryzias latipes* species complex) has wild populations distributed from high to low latitudes in Japan, and the latitudinal range becomes even wider if we consider close-related *Oryzias spp*. from outside. Different latitudes result in different selective pressures that may drive evolution, such as differences in temperature, resources availability and predation. For instance, higher latitudes have a shorter growing season, as environmental temperature is theoretically expected to limit growth rate in ectotherms (Angilletta et al., 2007). In medaka, there is a latitudinal cline in juvenile growth rate with populations from higher latitudes growing faster than those of lower latitudes (Yamahira and Takeshi, 2008), and, alternatively, ectotherms that grow slower in cold environments, but delay maturation and attain a bigger body size in adulthood, have also been reported (Angilletta et al., 2007; Trip et al., 2014). A latitudinal cline in body size is an ecological pattern often found in mammals (Blackburn et al., 1999), although its genetic mechanism remains unknown. Quantitative Trait Loci (QTL) analysis is a popular method developed to identify loci that affect quantitative traits such as body weight or gonadal weight, and therefore was one of the strategies chosen to investigate the phenotypes of interest in this study. In chapter I of this thesis, I make a general introduction on important topics for this study, such as QTL analysis, seasonal reproduction and medaka fish.

In chapter II, my objective was to investigate the genetic basis of body weight difference in medaka populations from different latitudes. I compared two wild medaka stocks originating from different latitudes: Maizuru and Ginoza. The Maizuru population from higher latitudes exhibited greater body weight than Ginoza. To better understand the genetic basis of body weight, those populations were crossed and quantitative trait locus (QTL) analysis was performed using 126 F_2 individuals. Simple interval mapping was performed with 371 RAD-markers distributed throughout all chromosomes, and Bayesian credible interval (95%) was computed using R/qtl software. One statistically significant QTL for body weight on medaka chromosome 4 was detected. With a search in literature, twelve candidate genes related to body weight or growth were identified in that QTL region. Nine of these had at least one SNP causing amino acid substitution in protein-coding regions, and the effects of those amino acid substitutions were estimated using PROVEAN (Protein Variation Effect Analyzer). The substitution on one of them was predicted to affect the protein function, and its protein sequence was compared with those of other vertebrate species using the sequence alignment tool ClustalW (version 2.1, DNA Data Bank of Japan), which showed that the region is much conserved. The QTL found in that analysis was estimated to explain 14% of the variance between the two populations. That result may help to uncover what loci are under selective pressures that result in the latitudinal clines of growth and body weight that we find in the wild.

A similar approach was made in the investigation of suppression of reproduction under short days in chapter III. In higher latitudes, seasonal breeders suppress reproduction in autumn and avoid producing offspring in the harsh winter conditions. In lower latitudes, however, the favorable season for reproduction tends to be longer. Therefore, suppression of reproduction under short days (SD) varies with latitude, and may determine whether an animal will be a seasonal or a non-seasonal breeder. In previous research in our research group (Shinomiya and Nakatsukasa, unpublished data), QTL analysis was performed with F_2 generation of the cross between Kiyosu medaka, from higher latitude and responsive to short days, and Miyazaki medaka, from lower latitude and non-responsive to short days. A significant QTL on chromosome 9 with 349 genes was found. The goal was to unveil the genetic mechanism that determines whether a population will be a seasonal or a non-seasonal breeder based on forward genetic approaches. In this current study, I selected 32 candidate genes from the QTL region accordingly to known connections to reproduction or circadian rhythms, and analyzed the

SNPs between the two populations. Previously sequenced genomes of Kiyosu (n = 6) and Miyazaki (n = 6) from an investigation of our research group provided SNPs data. Only SNPs homozygous in all Kiyosu and all Miyazaki were considered in the analysis performed with GENETYX software (version 13, GENETICS Inc., Tokyo, Japan) in search of variants that caused an amino acid substitution inside coding regions. The amino acid substitutions were analyzed by PROVEAN, and substitutions were found in two candidate genes. Therefore, their protein sequences were compared with those of other vertebrate species using the sequence alignment tool ClustalW. Both had the amino acid substitutions happening in conserved regions. Moreover, a substitution in one of genes was inside a sterol-sensing domain, which may increase the chances of affecting phenotypes related to reproduction. In addition to the SNPs in the QTL regions, RNA-Seq was also performed to obtain information on differently expressed genes (DEGs) between the two populations, as this could be another way to generate different phenotypes. Kiyosu and Miyazaki adults were exposed to short days (SD) (10L 14D) or long days (LD) (14L 10D) for 14 days, and sampled at zeitgeber time (ZT) 0 or ZT10, and their brain region containing the hypothalamus and pituitary was sampled for RNA-Seq. The expression data of RNA sequencing was normalized as transcripts per million (TPM) for comparisons between samples, and genes with a fold change higher than 1.5 TPM and FDR p-value lower than 0.05 were considered differentially expressed genes (DEGs) between Kiyosu and Miyazaki. To validate physiological conditions of two populations, I performed a gene ontology (GO) analysis. Top enriched GO terms for DEGs between Kiyosu and Miyazaki under SD condition highlighted the sensory perception of light, steroid metabolic process, and steroid hormone biosynthesis, supporting the differences in the reproductive status of those two populations. Then, I focused on the DEGs inside the QTL region. Among the 349 inside the QTL, 60 were differentially expressed between Kiyosu and Miyazaki at some time point, in LD or SD. Five of the DEGs in short-day condition are in the list of candidate genes in QTL region related to reproduction. One of them is particularly interesting because, as aforementioned, an amino acid substitution was observed between Kiyosu and Miyazaki.

Another gene that was considered a strong candidate was “gene X”. That gene is located close to the QTL peak and has an amino acid substitution between the two populations. It also has connections both to circadian rhythms and to reproduction in mammals. Therefore, it was chosen to be further investigated in chapter IV. The strategy

I chose was an investigation of loss of function using the genome editing tool CRISPR/Cas9 to create gene X-KO medaka. Additionally, as gene X was hypothesized to have a role in stress response due to responsiveness to glucocorticoids, I performed a behavioral experiment to investigate the difference between gene X-KO and wild-type in anxiety-like behavior. For the generation of gene X-KO medaka, I performed microinjection (Sawamura et al., 2017) in medaka embryos at the one-cell stage. Cas9 protein, tracrRNA and three crRNAs were simultaneously injected into the embryos according to the manufacturer's protocol for zebrafish embryos (Essner, 2016). Effectiveness of the designed gRNA on generating mosaic mutants, and also the germline transmission, were both confirmed by amplifying a gene X segment that includes the target sequences, and then genotyping with microchip electrophoresis. Direct sequencing was later performed by Eurofins (Tokyo, Japan) on DNA fragments amplified by PCR. Five different mutations were detected in G_1 from G_0 mosaic mutants. One of them was a 17-base deletion predicted to cause a frameshift mutation at the beginning of exon 1 of gene X, before the DNA-binding domain. Therefore, G_1 siblings from family #13 and G_1 siblings from family #14, which had the exact same 17-base deletion, were chosen. Those siblings generated G_2 by mass-mating of G_1 in each family. 119 G_2 grew up under the LD (14L 10D) condition, and were genotyped when they reached adult size. 18% were gene X-KO (-/-), 52% were heterozygous (+/-), and 30% were wild-type (WT) (+/+). Gene X-KO medaka were able to mate and produce healthy offspring, which means gene X is not essential for reproduction in both males and females. To test whether gene X is essential for the response to short days, a SD-response experiment was performed with 6-months old medaka. After two weeks in SD condition, eight gene X-KO, eight wild-type and eight heterozygous females had their gonadosomatic index (GSI) measured. If gene X had an essential role on gonadal regression induced by SD, KO fish would not show clear regression of ovaries, and a significantly higher GSI in gene X-KO females would be expected in comparison with wild-type females. However, no significant difference was found between KO females and WT females. Since the number of fish (eight of each genotype) examined was not sufficient, those results require to be confirmed. Genetic compensation cannot be excluded as well, as it has been reported to influence results in some loss-of-function studies in fish (Rossi et al., 2015). Those possibilities still need to be further investigated to completely exclude gene X as a candidate gene. For the investigation on gene X role in stress response, I used the "novel tank test". In this test, an individual is placed in a novel tank. The more time they spend at the bottom, the more

anxiety-like behavior they are displaying in response to the perceived uncertainty caused by a new environment (Cachat et al., 2010). This test was performed with 4-months old G_3 medaka: 32 gene X-KO (females = 14; males = 18) and 32 WT (females = 16; males = 16). The medaka were individually placed in the novel tank and recorded with video cameras. Then, the time each medaka spent in the top and bottom areas was quantified by Smart tracking software (Panlab Harvard Apparatus). As there are many reports of anxiety and other mood disorders affecting males and females differently (Palanza, 2001), Student's t-test was used to compare the effect of genotype in males and females separately. Gene X-KO females spent significantly less time at the top zone than wild-type females ($p < 0.05$). Therefore, gene X-KO females exhibited higher anxiety-like behavior in a new environment than wild-type females. While gene X-KO males also spent less time at top zone than wild-type males, the difference was not significant. Although results so far suggest gene X is not responsible for the response to short days, the difference in anxiety-like behavior that I found opens an investigation on the role of gene X in modulation of stress-response in medaka.

Responding or not to short days may determine seasonality in reproduction. The genetic mechanism behind that difference is still unknown in animals. In this study, I found interesting candidate genes inside the QTL for response to SD, and also learnt about their differences in expression levels. Those analyses provided a list of genes that, so far, I conclude are the most promising candidates for the response to short days. Although results for gene X were negative, further investigations on the other candidate genes may lead us to unveil the genetic mechanism that determines whether a population will be a seasonal or a non-seasonal breeder.