Mapping and Expression of Candidate Genes for Development Rate in Rainbow Trout (Oncorhynchus mykiss)

MATTHEW C. HALE, JOHN A. COLLETTI, SCOTT A. GAHR, JULIE SCARDINA, FRANK P. THROWER, MATTHEW HARMON, MEGAN CARTER, RUTH B. PHILLIPS, GARY H. THORGAARD, CAIRD E. REXROAD III, AND KRISTA M. NICHOLS

From the Department of Biological Sciences, Purdue University, West Lafayette, IN (Hale, Colletti, Scardina, Harmon, Carter, and Nichols); the Biology Department, St. Vincent College, Latrobe, PA (Gahr); Alaska Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Juneau, AK (Thrower); the Department of Biological Sciences, Washington State University, Vancouver, WA (Phillips); the Department of Biological Sciences and Center for Reproductive Biology, Washington State University, Pullman, WA (Thorgaard); the United States Department of Agriculture, Agricultural Research Service, National Center for Cool and Coldwater Aquaculture, Leetown, WV (Rexroad); the Department of Forestry and Natural Resources, Purdue University, West Lafayette, IN (Nichols); and the National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Northwest Fisheries Science Center, Conservation Biology Division, 2725 Montlake Boulevard East, Seattle, WA 98112 (Nichols).

Address correspondence to Krista M. Nichols at the address above, or e-mail: krista.nichols@noaa.gov.

Data deposited at Dryad: http://dx.doi.org/doi:10.5061/dryad.m4cc1

Abstract

Development rate has important implications for individual fitness and physiology. In salmonid fishes, development rate correlates with many traits later in life, including life-history diversity, growth, and age and size at sexual maturation. In rainbow trout (Oncorhynchus mykiss), a quantitative trait locus for embryonic development rate has been detected on chromosome 5 across populations. However, few candidate genes have been identified within this region. In this study, we use gene mapping, gene expression, and quantitative genetic methods to further identify the genetic basis of embryonic developmental rate in O. mykiss. Among the genes located in the region of the major development rate quantitative trait locus (GHR1, Clock1a, Myd118-1, and their paralogs), all were expressed early in embryonic development (fertilization through hatch), but none were differentially expressed between individuals with the fast- or slow-developing alleles for a major embryonic development rate quantitative trait locus. In a follow-up study of migratory and resident rainbow trout from natural populations in Alaska, we found significant additive variation in development rate and, moreover, found associations between development rate and allelic variation in all 3 candidate genes within the quantitative trait locus for embryonic development. The mapping of these genes to this region and associations in multiple populations provide positional candidates for further study of their roles in growth, development, and life-history diversity in this model salmonid.

Subject areas: Genomics and gene mapping; Quantitative genetics and Mendelian inheritance

Key words: qPCR, quantitative genetics, quantitative trait loci, salmonid

A central aim in the field of evolutionary genetics is to identify genes responsible for phenotypic variation. Typically, quantitative trait locus (QTL) approaches have been used to identify regions of the genome segregating for traits of interest, followed by fine mapping and subsequent identification of positional candidate genes from available genome sequences in QTL regions. For example, early QTL studies in threespine stickleback (Gasterosteus aculeatus) identified a genome region associated with plate armor (Peichel et al. 2001); with the later availability of a genome sequence and forward genetic approaches, a single gene (ectodysplasin, Eda) was identified as a major gene controlling morphology in multiple populations (Colosimo et al. 2004, 2005). Modern genomic sequencing technologies are rapidly advancing the
fields of ecological and evolutionary genetics, making this goal more realistic than ever before (Stapley et al. 2010). Frequently, however, the thousands of markers generated by genotype-by-sequencing approaches (such as Restriction site-Associated DNA [RAD] sequencing) are unannotated with respect to their functions (e.g., Amores et al. 2011; Everett et al. 2012; Hecht et al. 2012; Miller et al. 2012). The use of candidate genes with known function for a trait of interest to fine map a QTL, combined with gene expression approaches, is an attractive alternative or parallel approach.

In *Oncorhynchus mykiss* and related salmonid fishes, the rate of embryonic development has important implications in later life and is specifically associated with growth rate, age of first feeding, and age at sexual maturity (Allendorf et al. 1983). The phenotypic association between embryonic development rate and later life-history traits has been further corroborated by finding genetic associations between these traits. In particular, QTL studies in several populations of *O. mykiss* have found that a single chromosomal region in this species is associated with these multiple life-history traits (chromosome Omy5; Robison et al. 2001; Sundin et al. 2005; Nichols et al. 2007, 2008; Miller et al. 2012). In fact, genetic variation within this region of the genome is responsible for a substantial portion (21–28%) of the phenotypic variation in development rate across these populations (Robison et al. 2001; Sundin et al. 2005; Nichols et al. 2007; Miller et al. 2012). This region of the genome or markers associated with it have also been linked to both variation in the propensity to migrate to the ocean (or remain resident; Nichols et al. 2008) as well as the timing of return migration and spawning (O’Malley et al. 2002), not only in *O. mykiss* but also in other salmonids including arctic char (*Salvelinus alpinus*; Moghadam et al. 2007) and coho salmon (*Oncorhynchus kisutch*; McClelland and Naish 2010). The colocalization of trait associations in this genomic region suggests that the underlying genes encoding variation in these traits are either pleiotropic or in linkage disequilibrium in many populations.

Because previous QTL studies in salmonid fishes have used largely unannotated markers, we have a limited understanding of the genes and molecular mechanisms underlying variability in development rate and life-history diversity in *O. mykiss*. This is due in part to the lack of a complete, annotated reference salmonid genome. Although the *O. mykiss* linkage map contains some candidate genes that localize to the major development rate QTL (Danzmann et al. 2005; Leder et al. 2006; Phillips et al. 2006; Xu et al. 2011), the total number of mapped candidate genes is small. A single study using microarrays has examined differential expression of genes in backcross individuals selected for the major embryonic development rate QTL (Xu et al. 2011). Though many genes were differentially expressed between individuals with fast- and slow-development alleles at this QTL, mapping of a subset of those differentially expressed genes found that only 2 of 12 genes, *vtgr* (very-low-density lipoprotein) and *ppia* (peptidyl prolyl isomerase A), mapped to development rate QTL region. The remaining genes mapped to other chromosomes, suggesting a possible dual role of cis- and trans-acting effects on the genetic basis of development rate.

Here, we present a series of studies that aim to further our understanding of the genetic basis of embryonic development rate in rainbow and steelhead trout. To meet these goals, we used several populations of *O. mykiss* that included fast- and slow-developing individuals. In so doing, we evaluated whether 1) candidate genes mapped to a major QTL for development rate in this species, 2) gene expression of some of these candidate genes were differentially expressed over time and between fast- and slow-developing lines, 3) there was a significant heritable genetic component for development rate (time to hatch) in crosses between resident and migrant individuals sampled from a natural population of *O. mykiss* in Alaska, and 4) sequence variation in candidate genes, mapped to the major development rate QTL, was associated with development rate in this same population.

**Materials and Methods**

**Development Rate Candidate Gene Selection and Mapping Panels**

Twenty-one genes were selected for mapping, based on known function in other species or on the chromosome containing the major development rate QTL in this species (see Table 1 for details and references). These genes were selected for their putative roles in growth, development rate, or life-history diversity or were previously known to map to the major region associated with development rate in *O. mykiss*. Sequences from GenBank, or the literature, were used to design primers for sequencing (see Supplementary File 1 and Supplementary Table S1 online) and to identify polymorphisms within a doubled-haploid cross between 2 clonal lines: a hatchery-raised Shasta-type rainbow trout from Oregon State University (OSU) and a hatchery population of migratory steelhead trout from Clearwater, ID (CW) (see Young et al. 1996 for details of these lines). This doubled-haploid cross has been used for many mapping studies in *O. mykiss* (Nichols, Bartholomew, et al. 2003; Phillips et al. 2006, 2012; Nichols et al. 2007, 2008). The cross consists of 2 mapping panels, one with 110 individuals and another with 99 individuals. The first panel has primarily been used for mapping genes and traits associated with migratory phenotypes (Nichols et al. 2008) and the second panel for genes and traits associated with disease resistance and immune function (Nichols, Bartholomew, et al. 2003; Phillips et al. 2006, 2012). The 2 panels were combined in this study to increase the number of potential informative meioses for linkage mapping.

**Sequencing of Candidate Genes**

For single nucleotide polymorphism (SNP) discovery in new genes, primers were designed to span introns or the 3’ untranslated region of genomic DNA, with a target fragment size of 500–100 bp. For SNP discovery in genes that had been previously sequenced and identified in *O. mykiss*, primers from prior studies were used to amplify the gene fragments in the mapping families. Gene fragments were amplified in 10 μL.
for each sample consisting of 100 ng of DNA, 1.0 μm each primer, 200 μm dNTPs, 1× polymerase chain reaction (PCR) buffer, 1.5 mM MgCl2, and 0.5 U of Taq polymerase (Bioline, Taunton, MA). The thermal cycle consisted of 3 min at 94 °C followed by 35 cycles of 30 s at 94 °C, 35 s at 55 °C, and 1.25 min at 72 °C. PCR products were either cleaned using SAP-Exo (Applied Biosystems, Foster City, CA) following manufacturer's instructions or extracted from a 1.2% agarose gel stained with ethidium bromide using QIAquick columns (Qiagen, Valencia, CA) following manufacturer’s instructions. Cycle sequencing was performed with BigDye version 3.1 (Applied Biosystems) and sequence separation was conducted on an ABI 3130xl. Forward and reverse sequences were aligned using SeqMan (DNASTar, Madison, WI). SNPs were chosen if there were fixed differences between the 2 clonal lines (as doubled-haploid lines were used, there were no heterozygous positions). SNPs were genotyped in the doubled-haploid individuals using ABI SNaPshot (Applied Biosystems) with SNP-specific primers. Genotypes were scored using GeneMapper V 3.7 (Applied Biosystems).

Table 1  Details of genotyped SNPs with known function to development rate or smoltification, including marker name, the chromosome and linkage group of where the marker was mapped, the function of the gene, and the original mapping study or origin of sequence data

<table>
<thead>
<tr>
<th>Marker ID</th>
<th>Chromosome</th>
<th>Linkage group</th>
<th>Name/annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myd118-2</td>
<td>1</td>
<td>6</td>
<td>Myeloid differentiation response protein 2</td>
</tr>
<tr>
<td>ARNT</td>
<td>2</td>
<td>27</td>
<td>Aryl hydrocarbon receptor nuclear translocator</td>
</tr>
<tr>
<td>IGF1Rα</td>
<td>2</td>
<td>27</td>
<td>Insulin-like growth factor 1 receptor alpha</td>
</tr>
<tr>
<td>POMCβ</td>
<td>4</td>
<td>24</td>
<td>Proopiomelanocortin B</td>
</tr>
<tr>
<td>TNN</td>
<td>4</td>
<td>24</td>
<td>Troponin I, fast skeletal muscle (isoform 3)</td>
</tr>
<tr>
<td>TSHRβ</td>
<td>4</td>
<td>24</td>
<td>Thyroid stimulating hormone receptor beta</td>
</tr>
<tr>
<td>UroI</td>
<td>4</td>
<td>24</td>
<td>Urotensin 1</td>
</tr>
<tr>
<td>Clock1α</td>
<td>5</td>
<td>8</td>
<td>Circadian rhythm gene</td>
</tr>
<tr>
<td>GHR1</td>
<td>5</td>
<td>8</td>
<td>Growth hormone receptor 1</td>
</tr>
<tr>
<td>Myd118-1</td>
<td>5</td>
<td>8</td>
<td>Myeloid differentiation response protein 1</td>
</tr>
<tr>
<td>Vgr</td>
<td>5</td>
<td>8</td>
<td>Very-low-density lipoprotein receptor precursor</td>
</tr>
<tr>
<td>Ppia</td>
<td>5</td>
<td>8</td>
<td>Peptidyl-prolyl cis–trans isomerase A</td>
</tr>
<tr>
<td>ATP51b</td>
<td>7</td>
<td>12</td>
<td>Sodium potassium ATPase beta subunit 1b</td>
</tr>
<tr>
<td>MyoG</td>
<td>7</td>
<td>12</td>
<td>Myogenin</td>
</tr>
<tr>
<td>DIA</td>
<td>7</td>
<td>12</td>
<td>Diaphorase (NADH)</td>
</tr>
<tr>
<td>TSHRα</td>
<td>7</td>
<td>12</td>
<td>Thyroid stimulating hormone beta</td>
</tr>
<tr>
<td>RXα</td>
<td>8</td>
<td>23</td>
<td>Retinoic X receptor alpha</td>
</tr>
<tr>
<td>TSHRα</td>
<td>8</td>
<td>23</td>
<td>Thyroid stimulating hormone receptor alpha</td>
</tr>
<tr>
<td>FTFZ1</td>
<td>10</td>
<td>20</td>
<td>Fushi tarazu factor 1</td>
</tr>
<tr>
<td>GCR</td>
<td>10</td>
<td>20</td>
<td>Glucocorticoid receptor</td>
</tr>
<tr>
<td>p53</td>
<td>10</td>
<td>20</td>
<td>Tumor suppressor protein p53</td>
</tr>
<tr>
<td>Rpl36a</td>
<td>10</td>
<td>20</td>
<td>Ribosomal protein 36a</td>
</tr>
<tr>
<td>mrlb</td>
<td>10</td>
<td>20</td>
<td>Myosin light chain 2</td>
</tr>
<tr>
<td>UCP2α</td>
<td>10</td>
<td>20</td>
<td>Uncoupling protein 2a</td>
</tr>
<tr>
<td>OXCT1</td>
<td>11</td>
<td>19</td>
<td>3-Oxoadic CoA transferase 1</td>
</tr>
<tr>
<td>Colla</td>
<td>12</td>
<td>9</td>
<td>Alpha 1 type I collagen</td>
</tr>
<tr>
<td>Epd2</td>
<td>12</td>
<td>9</td>
<td>Ependymin 2</td>
</tr>
<tr>
<td>GSDF</td>
<td>12</td>
<td>9</td>
<td>Gonadal soma-derived growth factor</td>
</tr>
<tr>
<td>HSP70</td>
<td>12</td>
<td>9</td>
<td>Heat shock protein 70</td>
</tr>
<tr>
<td>NDKPB</td>
<td>12</td>
<td>9</td>
<td>Nucleoside diphosphate kinase B</td>
</tr>
<tr>
<td>Hnrdpl</td>
<td>12</td>
<td>9</td>
<td>Heterogeneous nuclear ribonucleoprotein D</td>
</tr>
<tr>
<td>IgL</td>
<td>12</td>
<td>9</td>
<td>Immunoglobulin light chain</td>
</tr>
<tr>
<td>PRLR</td>
<td>12</td>
<td>9</td>
<td>Prolactin receptor</td>
</tr>
<tr>
<td>IGfTb</td>
<td>13</td>
<td>2</td>
<td>Immunoglobulin heavy chain Tb</td>
</tr>
<tr>
<td>mlc-2</td>
<td>13</td>
<td>2</td>
<td>Myosin regulatory light chain 2</td>
</tr>
<tr>
<td>ATOX1</td>
<td>14</td>
<td>3</td>
<td>Copper transport protein ATOX1</td>
</tr>
<tr>
<td>THRβ</td>
<td>14</td>
<td>3</td>
<td>Thyroid hormone receptor beta</td>
</tr>
<tr>
<td>THRα</td>
<td>16</td>
<td>22</td>
<td>Thyroid hormone receptor alpha</td>
</tr>
<tr>
<td>TRH</td>
<td>16</td>
<td>22</td>
<td>Thyrotopin-releasing hormone precursor</td>
</tr>
<tr>
<td>YB1</td>
<td>16</td>
<td>22</td>
<td>Y-box-binding protein</td>
</tr>
<tr>
<td>Dio</td>
<td>19</td>
<td>14</td>
<td>Type II iodothyronine deiodinase</td>
</tr>
<tr>
<td>POMCA2</td>
<td>19</td>
<td>14</td>
<td>Proopiomelanocortin A2</td>
</tr>
<tr>
<td>DECR1</td>
<td>27</td>
<td>11</td>
<td>Dienoyl-CoA deductase</td>
</tr>
<tr>
<td>TF</td>
<td>28</td>
<td>13</td>
<td>Transferrin</td>
</tr>
<tr>
<td>POMCA1</td>
<td>29</td>
<td>25</td>
<td>Proopiomelanocortin A1</td>
</tr>
<tr>
<td>TGFP1</td>
<td>Sex</td>
<td>1</td>
<td>Transforming growth factor beta 1</td>
</tr>
<tr>
<td>GHR2</td>
<td>Sex</td>
<td>1</td>
<td>Growth hormone receptor 2</td>
</tr>
</tbody>
</table>

Leder et al. (2006)
Linkage Map Construction

Linkage maps were constructed by combining both OSU × CW mapping panels in JoinMap 4.0 (Van Ooijen 2011). Markers showing segregation distortion or whose genotypes deviated significantly from an expected 1:1 ratio were removed. First, 2 separate maps were constructed using the Kosambi map function with a minimum log of the odds (LOD) score of 3.0. Marker order was checked using maximum likelihood, and the most likely marker order for each linkage group was determined. Maps were then merged using the “combine maps” function in JoinMap (Van Ooijen 2011), for a combined map based on 209 individuals. JoinMap uses the loci in common to construct a framework map and then adds the loci typed in 1 panel in the most likely position based on the recombination fraction. This was done for each of the 29 O. mykiss linkage groups. Nomenclature of the 29 linkage groups followed Voorrips et al. (2002) and Phillips et al. (2006). Maps were drawn in MapChart 2.1 (Voorrips 2002).

QTL Mapping of Development Rate Traits

We reanalyzed the QTL for 3 development rate traits previously published in the OSU × CW cross (Nichols et al. 2007, 2008). Time to hatch (tth), an often-used proxy for embryonic development rate (Ferguson et al. 1985), was calculated for each fish as accelerated temperature units (ATUs), where ATU = (incubation temperature − natural log of the ng input of the RNA) × (calendar days from fertilization to hatch) (Nichols et al. 2007). Instantaneous growth rate in standard length during the spring of the second year (sprgrsl; Nichols et al. 2008) and instantaneous growth rate in weight during the spring of the second year (sprgrwt; Nichols et al. 2008) were calculated between February and June 2003. Instantaneous growth rates were calculated using:

\[
growth rate = \frac{\ln(L_2) - \ln(L_1)}{t_2 - t_1} \times 100,
\]

where \(L_1\) and \(L_2\) are standard lengths (sprgrsl) or weights (sprgrwt) at times \(t_1\) and \(t_2\) (Nichols et al. 2008). A subset \((n = 110)\) of fish used for prior development rate QTL mapping was used to localize the development rate QTL with the addition of the markers described herein. QTLs were identified using both interval mapping (IM) (Zeng 1994) and multiple trait QTL mapping (MQM) (Kao et al. 1999). Permutation tests (Churchill and Doerge 1994) were run to determine genome- and chromosome-wide LOD significance thresholds at the 95% level \((n = 1000\) permutations). On those linkage groups showing significant QTL with IM, the markers with the largest significant LOD score were selected as cofactors in the MQM models. QTL identified by MQM were drawn on the map using MapChart 2.1 (Voorrips 2002). Two LOD support intervals were estimated for putative QTL as a measure of confidence of the localization of QTL (estimated at around 95% confidence; Visscher et al. 1994). All QTL analyses were performed in MapQTL 5 (Van Ooijen 2004).

Expression of Candidate Genes

Quantitative PCR (qPCR) was used to evaluate expression of candidate genes during early development with 2 objectives: 1) to evaluate fine-scale temporal expression of genes from fertilization through hatch and 2) to evaluate differences in expression of genes between a fast- and a slow-developing line of rainbow trout at fewer time points. The genes used for expression analysis were those that mapped to the major QTL for embryonic development rate. Because salmonids have a recently duplicated genome (Allendorf and Thorgaard 1984), the duplicated copies of some genes were also examined. Myd118-1 (Phillips et al. 2006) and Clock1a (Leder et al. 2006; Paibomesai et al. 2010) were previously shown to map to Omy5, and GHR1 was mapped to Omy5 in this study (see below). GHR2 and Myd118-2 were identified as duplicates of GHR1 and Myd118-1, respectively. No duplicate was investigated for Clock1a.

For a fine-scaled examination of the time course of expression of candidate genes, rainbow trout embryos were collected at 20 time points \((0–24\) days postfertilization \(\text{dpf})\). Daily samples of 5 embryos were collected from fertilization \((0\) dpf) through 14 dpf and then samples were collected every other day through 24 dpf (hatch). Collections of embryos were made from 2 lots of embryos previously described by Coulibaly et al. (2006): each lot consisted of eggs pooled from 2 to 3 females and fertilized with the sperm of a single male, such that 2 unique males were used. RNA was extracted from embryos and cDNA synthesized as described by Coulibaly et al. (2006). Samples were run in triplicate on the same plate with a negative control that lacked cDNA and with a reference gene (TATA box–binding protein a; \(TBP\a\); Filby and Tyler 2007) that was set up in triplicate for each sample. qPCR reactions consisted of \(1\times\) SYBR green, \(0.36\) µM of each primer, \(1\) µL of template cDNA, and water in a total volume of 10 µL. The thermal profile for all genes was \(95\) °C for 10 min, followed by 40 cycles of \(95\) °C for 15 s and \(58 \) °C for 1 min. Finally, a melting curve analysis was conducted from \(50\) to \(90\) °C with \(0.5\) °C increases per cycle for a total of 80 cycles to ensure that there was no misannealing or contaminated genomic DNA in the sample. In the developmental series, gene expression was quantified using the relative standard curve method (Rutledge and Cote 2003). We used a serial dilution (8 standards) of a mixed tissue cDNA to construct a standard curve for each assay plate. The standard curve was constructed by plotting the threshold cycle \((C_t)\) versus the natural log of the ng input of RNA into the reaction. The \(C_t\) values were determined using automated settings on the ABI 7900. The abundance of each transcript was calculated using this curve for each sample. This value was normalized to the reference gene \((TBP\a)\) expression to control for differences in the cDNA concentrations in the reaction. Each sample was run in triplicate on a single plate, and each plate was run in duplicate. Mean relative expression and standard deviation across the 2 pools of embryos were calculated for each time point.

Expression in a fast- and slow-developing backcrossed introgressed line of rainbow trout was evaluated at 3 times during embryonic development. A fourth generation...
backcross family was made at Washington State University with the aid of marker-assisted selection to select third-generation backcross individuals possessing the allele on chromosome Omy5 for fast development from the CW clonal line of _O. mykiss_. A single male with the fast-developing allele was used for crossing with a female from the same slow-developing line used for previous generation backcrosses from the OSU rainbow trout clonal line. The microsatellite markers OMM1009 and OmyFGT12TFU, linked to the major development rate QTL on chromosome 5 (Sundin et al. 2005; Nichols et al. 2007), were used to identify third-generation males possessing the CW (fast-developing) allele, as described by Xu et al. (2011). The male chosen was heterozygous for alleles from CW and OSU at both loci and was crossed with a single OSU female, producing progeny that segregated roughly 50% homozygous for the OSU alleles and 50% heterozygous for the OSU/CW alleles. Molecular markers underlying the major embryonic development rate QTL indicated that individual embryos were OSU/OSU or OSU/CW, as described by Xu et al. (2011). Eight embryo samples taken from 3 temporal collections (15, 19, and 28 dpf) were used for gene expression analysis. qPCR was set up as described above, except that β-actin was used as a reference gene (Hale et al. 2011; Xu et al. 2011). The qRT-PCR data were analyzed using Pfaffl’s ΔΔCT method with expression patterns compared with those of OSU/OSU samples at 15 dpf (Pfaffl 2001). General linear models were used to determine if expression differences between OSU and CW were significant. Gene expression, y was modeled as a function of sex (S), time (T), and genotype (G) as follows:

\[ y_{ijl} = S_i + T_j + G_k + (S_i T_j) + (S_i G_k) + (T_j G_k) + e_{ijl}. \]

Sex was determined by genotyping individuals using OmyY1 (Brunelli et al. 2008), T is developmental time (15, 19, and 28 dpf), and G is the QTL genotype of the individual (OSU/OSU or OSU/CW). Models were evaluated in SAS 9.2 (SAS Statistical Institute, Cary, NC).

**Heritability and Candidate Gene Association Analysis of Development Rate in a Natural Population of _O. mykiss_**

In 2008, 12 experimental crosses were made from migratory anadromous and resident _O. mykiss_ from Sashin Creek and Sashin Lake, Alaska. Fish from Sashin Lake, which flows into Sashin Creek, share a common ancestry with Sashin Creek, which flows into Sashin Lake, Alaska. Fish from Sashin Lake, which flows into Sashin Creek and Sashin Lake, Alaska. Fish from Sashin Creek and Sashin Lake, Alaska. Fish from Sashin Creek and Sashin Lake, Alaska. Fish from Sashin Creek and Sashin Lake, Alaska. Fish from Sashin Creek and Sashin Lake, Alaska. Fish from Sashin Creek and Sashin Lake, Alaska. Fish from Sashin Creek and Sashin Lake, Alaska. Fish from Sashin Creek and Sashin Lake, Alaska. Fish from Sashin Creek and Sashin Lake, Alaska.

Crosses were made on 3 dates (27 May, 3 June, and 10 June 2008) when mature adults were available. We sampled about 400 eggs from each of 192 families across the 12 cross types. Of these, about 100 fertilized eggs from 171 families (a subset showing good survival to the eyed stage) were placed into hatch grids with each embryo placed in its own cell to minimize the effects of hatching enzymes between embryos. Each family was monitored for time to hatch at 4-h intervals, as described by Robison et al. (2001). At hatching, individuals were removed from the grid, and time to hatch was calculated as the number of ATUs from fertilization to hatch.

To evaluate whether development rate differed between cross types, a mixed model was used:

\[ \text{devrate}_{ij} = \text{spawn} + \text{offspring line}_{ij} + \text{male}_{i} + \text{female}_{j} + e_{ij}, \]

where \( i \) indicates spawning date (most cross types were produced on multiple spawning dates), \( j \) indicates cross type, \( k \) indicates the identity of the male used in the cross, \( l \) is the identity of the female used, and \( m \) indexes the individual phenotype for time to hatch. Spawning date and cross type were treated as fixed effects, whereas sex was included as a random effect in the model. Least-squares means estimated for offspring line and spawning were calculated, and Tukey’s honest significant difference (HSD) was used to test the hypothesis that pairs of means were significantly different. A line-cross analysis was used to evaluate whether the means across each of the cross types were distributed in a way consistent with the additive action of genes (see Supplementary File 2 online; Lynch and Walsh 1998). Statistical tests were performed in SAS (SAS Statistical Institute). Type I error rate was set at 0.05.

**Heritability**

Heritability and maternal effects for development rate in the Alaska crosses were calculated using the animal model (Kruuk 2004):

\[ \text{residual devrate}_{ij} = \text{ID}_{i} + \text{dam}_{j} + e_{ij}, \]

where \( i \) indicates individual and \( j \) is the female used in the cross.

The dependent variable, residual devrate, was calculated as the residual development rate (number of ATUs) accounting for spawning date. Residual ATUs within each spawning time were used, rather than using spawning time as a fixed effect because there were differences in the timing and temperature of the 3
spawning and because individual females were confounded with spawning time. This latter point leads to singularity in the design matrix such that maternal effects cannot be reliably tested in a model in which spawning time also occurs as a fixed effect. We used a general linear model to calculate residual ATUs, with development rate as the dependent variable and spawning date as the only independent predictor (SAS PROC GLM). In the animal model, ID (identity of the embryo) and dam were modelled as random effects, with the variance in ID taking into account the additive genetic relationship matrix among individuals. ASReml 2.0 (VSN International, Hemel Hempstead, UK) was used for the analysis. Heritability ($h^2$) was calculated as the ratio of additive genetic variance ($V_A$) to total phenotypic variance ($V_P$), and the proportion of phenotypic variation due to maternal effects ($m^2$) was calculated as the ratio of variance among dams to the total phenotypic variance. To test whether variation due to maternal effects ($m^2$) or additive genetic effects ($h^2$) was nonzero, we used a likelihood ratio (LR) test to test the difference with and without those effects in the model; the LR of the full and reduced models was compared with a chi-square distribution with 1 degrees of freedom to test for significance with a type I error rate of 0.05.

**Candidate Gene Association Analysis From the Same Families Used for Heritability Analyses**

Twelve individuals from 8 families from the heritability analysis described above were used to examine whether allelic variation in *Clock1a*, GHR1, and *Myd118-1* was associated with development rate. These 8 families consisted of 4 families with some of the largest recorded ATUs and 4 with some of the smallest recorded ATUs; full-sib families that did not share parents were chosen (see Supplementary File 2 online). DNA was extracted using the Wizard SV Genomic DNA Kit following manufacturer’s recommendations (Promega, Madison, WI). Gene fragments were amplified by PCR (see Supplementary File 1 and Supplementary Table S1 online for primer sequences) and were subsequently cleaned and sequenced as described above (see Sequencing of Candidate Genes). One fish from each family was initially sequenced to identify polymorphism within the amplified product. Forward and reverse sequences were assembled using SeqMan (DNAStar), and polymorphisms within the genes were identified. SNP and insertion/deletion (indel) polymorphisms were genotyped using ABI SNaPshot with SNP-specific primers and following manufacturer’s instructions. Genotypes were scored using GeneMapper 3.7 (Applied Biosystems). Associations between genotype (independent variable) and development rate (dependent variable) were tested with a mixed model ANOVA (PROC MIXED) in SAS 9.2 (SAS Statistical Institute) with a type I error rate of 0.05. ATUs were modelled with the fixed effects of spawning date and genotype and the random effect of family; pairwise comparisons of least-squares means for each genotype were made with a Tukey’s HSD test.

**Data Archiving**

In fulfillment of data archiving guidelines (Baker 2013), we have deposited the primary data underlying these analyses with Dryad.

**Results**

**Linkage and QTL Mapping**

A total of 559 markers were analyzed, of which 455 had significant linkage (LOD score > 3.0) with at least one other marker in the combined map and were placed in 29 linkage groups. Of these, 434 have been previously published in other *O. mykiss* linkage maps (Nichols, Bartholomew, et al. 2003; Nichols, Young, et al. 2003; Nichols et al. 2004, 2007, 2008; Phillips et al. 2006; Xu et al. 2011). The 2 mapping panels shared 153 markers that served as anchor points for the markers unique to each panel. The joint linkage map consisted of 167 markers from the Nichols et al. (2008) map and 135 markers from the Nichols et al. (2004) maps. The total combined sex averaged map distance was 1320 cM, with an average intermarker distance of 2.9 cM. This is somewhat shorter than reported in previous published maps of *O. mykiss* (c.g., 2927.1 cM, Rexroad et al. 2008; 2627.5 cM, Young et al. 1998; 2750 cM, Guyomard et al. 2006) and undoubtedly is due to using doubled-haploid individuals originating from males, which have a reduced recombination rate when compared with standard mixed-sex or female mapping panels. Linkage groups varied in size from 5.8 cM (Omy24) to 96.0 cM (Omy4). The map consisted of 227 amplified fragment length polymorphisms (AFLPs) (see Nichols et al. 2004, 2008), 139 microsatellites (Guyomard et al. 2006; Rexroad et al. 2008), and 89 SNPs. Map locations of SNPs not previously mapped in this cross are reported in Figure 1. Though the map location of GHR2 was previously reported in Phillips et al. (2009), we also include the location here, because it was a paralog of GHR1 mapped in this study.

Three candidate genes were mapped to Omy5, *Clock1a*, GHR1, and *Myd118-1*. Of these, *Clock1a* previously appeared on a rainbow trout linkage map (Leder et al. 2006; Paibomesai et al. 2010), and *Myd118-1* and GHR1 were previously placed on the physical map (Phillips et al. 2006), but details on linkage map placement and association with development rate QTL have not been previously reported (Figure 1). Two other candidate genes, GHR2, which mapped to the sex chromosome (also reported in Phillips et al. 2009), and *Myd118-2*, which mapped to Omy1, were in trans (i.e., away from the QTL for development rate) with the major QTL for development rate. A reanalysis of QTL for 3 development rate traits in the OSU × CW cross identified a major QTL located on Omy5 for all traits and accounted for 35.8% (tth), 17.2% (sprgrs1), and 15.6% (sprgrwt) of the total variance in each trait, respectively (Figure 1 and Table 2). The LOD support intervals for the QTL for these 3 traits overlapped and overlapped with the 3 mapped candidate genes on this linkage group.

**Candidate Gene Expression**

Patterns of expression of GHR1 and GHR2 were similar during embryonic development, with a general trend for increased expression starting at 4–5 dpf (Figure 2). The expression profile for *Clock1a* showed a general increase in expression during embryonic development, but also great
variability within sample times, perhaps owing to the cyclic expression of circadian genes (Figure 2). Expression of *Myd118* showed strikingly different profiles, particularly during the earliest development times. *Myd118-2* showed a steady increase in expression from about 5 dpf to hatching. On the other hand, *Myd118-1* expression rapidly increased from 2 to 4 dpf before a drastic decline in expression between 7 and 10 dpf through hatching (Figure 2).
Figure 1. Continued
Table 2  QTL results of growth and development phenotypes in rainbow trout: time to hatch (tth), spring growth in standard length (sprgrsl), and spring growth in weight (sprgrwt)

<table>
<thead>
<tr>
<th>QTL</th>
<th>Chromosome</th>
<th>Position (cM)</th>
<th>Nearest marker</th>
<th>LOD</th>
<th>% VE</th>
<th>GW sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTH-OMY5</td>
<td>5</td>
<td>15.9</td>
<td>AGCCAG10</td>
<td>11.03</td>
<td>35.8</td>
<td>Y</td>
</tr>
<tr>
<td>TTH-OMY29</td>
<td>29</td>
<td>12.3</td>
<td>AGCGAT5</td>
<td>2.12</td>
<td>4.7</td>
<td>N</td>
</tr>
<tr>
<td>SPRGRSL-OMY5</td>
<td>5</td>
<td>11.4</td>
<td>OMM1423</td>
<td>4.01</td>
<td>17.2</td>
<td>Y</td>
</tr>
<tr>
<td>SPRGRSL-OMY10</td>
<td>10</td>
<td>26.3</td>
<td>AGCGAT19</td>
<td>2.75</td>
<td>11.3</td>
<td>N</td>
</tr>
<tr>
<td>SPRGRWT-OMY5</td>
<td>5</td>
<td>3.9</td>
<td>TC145239</td>
<td>3.14</td>
<td>15.6</td>
<td>Y</td>
</tr>
<tr>
<td>SPRGRWT-OMY10</td>
<td>10</td>
<td>29.9</td>
<td>OMM1438</td>
<td>3.56</td>
<td>15.4</td>
<td>Y</td>
</tr>
</tbody>
</table>

Permutation tests (Churchill and Doerge 1994) were run to determine genome-wide and chromosome-wide LOD significance thresholds at the 95% level (n = 1000 permutations). The peak position of the QTL is recorded, as is the nearest marker to the QTL, the LOD score of the QTL, the percentage of variance explained (% VE) by the QTL, and whether the QTL is significant at the genome-wide (GW sig) level.

Expression Profiles Between OSU and CW Lines

Expression differences of the 5 candidate genes having paralogs on Omy5 (Clock1a, Myd118-1, Myd118-2, GHR1, and GHR2) were tested between a fast-developing (CW, anadromous steelhead) and a slow-developing (OSU, resident rainbow trout) QTL genotype (Supplementary File 1 and Supplementary Table S2 online). The expression of only a few genes differed significantly between individuals with the fast- and slow-developing QTL genotypes (Supplementary File 1, Supplementary Figure S1, and Supplementary Table S2 online). The interaction between dpf and sex was significant for Myd118-1 (P = 0.04), but genotype was not statistically significant either as a main effect or as part of an interaction. Sex by genotype was nearly significant for Clock1a (P = 0.07) but otherwise, none of the other effects or interaction terms was significant. Nevertheless, some notable trends appeared. For example, the expression of Clock1a was similar between the 2 QTL genotypes, except at 28 dpf, when OSU/CW individuals showed nearly a 2-fold increase in expression compared with OSU/OSU individuals. In Myd118-1, OSU/OSU individuals showed upregulated expression compared with OSU/CW, whereas Myd118-2 showed upregulation in the OSU/CW line relative to OSU/OSU.

Development Rate Heritability in a Natural Population

Spawning and offspring line were both significantly associated with development rate (ATUs; P < 0.01 and P < 0.001, respectively). The anadromous line (A × A), which had a larger mean ATU (slower development rate), differed significantly from all other lines, except from one of the F2 crosses (ARAR × ARAR; Figure 3). The resident line (R × R) had the lowest mean ATU (fastest development rate), though this mean did not differ significantly from all of the hybrid crosses, likely owing the greater amount of variation and small sample size (n = 193). None of the F2 and F3 crosses differed significantly from one another, and these and the 2 F1 crosses (A × R and R × A) were intermediate to the parental lines (Figure 3). The line-cross analyses supported the findings that the distributions of the means were consistent with additive gene action (see Supplementary File 1 online for complete description).

With greater than 16 380 progeny produced in 172 families in the Sashin, Alaska crosses, heritability analyses demonstrated that significant additive genetic variation and maternal effects contributed to the observed phenotypic variation in development rate. About 60% of observed variation in development rate among the crosses was attributed to additive genetic effects (h2 = Vg/Vp = 0.603 ± 0.138, P < 0.001). This value was within the range ofheritabilities quantified within each spawning (0.58–0.82; see Supplementary File 1 and Supplementary Table S3 online). Maternal effects also were significant in the full data set (m2 = Vm/Vp = 0.169 ± 0.073, P = 0.035).

Candidate Gene Association Analysis of Development Rate in a Natural Population

Sequence polymorphisms (SNPs or indels) were genotyped in 96 individuals from 8 families. These 8 families included representatives from the A × A, R × R, A × R, AR × AR, and AR × RA crosses (see Supplementary File 2 online for complete details and genotype data). A 5-bp indel within the Clock1a sequence produced a significant association (P < 0.0001) with development rate, with embryos with the 270-bp allele developing significantly faster than those with the 275-bp allele (Figure 4). For Myd118-1, embryos with the A/A genotype developed significantly faster than the heterozygote (A/T) class (P = 0.0018). Though there was a greater amount of variation in the mean development rates for the 3 GHR1 genotypic classes, the model was significant (P = 0.0043; Figure 4).

Discussion

In this study, we have examined mapping, expression, and association of candidate genes with the aim of identifying genes that play a key role in divergence of development rate among populations or strains of rainbow and steelhead trout. Through next-generation sequencing technologies, the field of evolutionary genetics and the identification of genes underlying adaptive traits are moving at a rapid pace. However, a candidate gene approach is also valid, and here, we show that a few positional candidate genes map to a major locus controlling development rate in O. mykiss. These genes were associated with development rate in several crosses and populations, and duplicates of these candidate genes showed, in some cases, different patterns of expression over the course of embryonic development.
Figure 2. Relative expression profiles of (A) GHR1 and GHR2, (B) Clock1a, and (C) Myd118-1 and Myd118-2 in a time course of samples collected from fertilization to the approximate time of hatching. Mean and the standard errors of the means for each of 2 pools at each time point (dpf) are presented (embryos were maintained at 12.9–13.4 °C).
The importance of rainbow trout in the fields of aquaculture, ecology, and evolution has led to the construction of several genetic linkage maps and QTL studies of traits of economic and ecological interest (Sakamoto et al. 2000; Nichols et al. 2004, 2007, 2008; Guyomard et al. 2006; Rexroad et al. 2008; Hecht et al. 2012; Miller et al. 2012). However, most markers in these maps are AFLPs (Nichols et al. 2004, 2007, 2008), microsatellites (Sakamoto et al. 2000; Guyomard et al. 2006; Rexroad et al. 2008), or RAD tags (Hecht et al. 2012; Miller et al. 2012) that are unannotated with respect to linkage to or polymorphisms in genes or regulatory regions that are functionally important for variation in phenotypic traits. To date, this and prior studies have identified 3 genes (Clock1a, Myd118-1, and GHR1) within the QTL for embryonic development rate on Omy5 (Sundin et al. 2005; Nichols et al. 2007; Easton et al. 2011; Miller et al. 2012). These 3 genes map in very close linkage beneath the QTL, but determining their individual contribution to development rate by breaking up the close linkage between the genes is difficult. This is primarily due to both the reduced recombination observed in male-derived doubled-haploid individuals used for mapping in this study and the greatly reduced recombination in this particular region of the O. mykiss genome in general (Danzmann et al. 2005; Rexroad et al. 2008).

The central region on Omy5 in the rainbow trout genome appears to be particularly important in influencing life-history variation. The reduced recombination rate on Omy5 may be directly related to the conservation of advantageous alleles that encode for faster development rate and coupled life-history traits in geographically isolated populations (Nichols et al. 2007; Easton et al. 2011; Miller et al. 2012). QTL for age at sexual maturity (O’Malley et al. 2002; Sundin et al. 2005; Haidle et al. 2008; Easton et al. 2011), growth (Nichols et al. 2008; Wringe et al. 2010), and smoltification-related traits (Nichols et al. 2008; Hecht et al. 2012) also map to the same region on Omy5 in this species. A completed genome sequence for chromosome Omy5 would allow researchers to identify additional positional candidate genes that could be further examined for associations with development rate and other life-history traits in this species. In future studies, however, the primary challenge will not just be in the identification of candidate genes in this region but breaking up the close associations between causal and noncausal variants to identify the truly functional variation promoting developmental life-history diversity in this species.

A link between development rate and migratory phenotype has been suggested in salmonids (O’Malley et al. 2002; Nichols et al. 2008; Miller et al. 2012). Genome-wide QTL and association studies in the Sashin O. mykiss have demonstrated that loci on Omy5 contribute to variation in migration-related traits (Hecht et al. 2012; Hale et al. 2013). Both life-history type and development rate have significant quantitative genetic variation in the Sashin, Alaska populations used herein. Approximately 76% of the variation in migratory behavior is attributed to additive genetic variance (Thrower et al. 2004), and approximately 60% of the variation in development rate is attributed to additive genetic variation in these populations (this study). A formal test for genetic correlation between development rate and migratory phenotype would be difficult to conduct. Nevertheless, in this study, progeny of the pure-resident crosses hatched about 12 ATUs earlier than did progeny of pure-migrant crosses, indicating a faster embryonic development rate in residents than in migrants in a common environment. Similarly, though later in development, another study of Sashin Creek O. mykiss found that residents had higher juvenile growth rates than migrants, even though they were smaller on average than smolts (migrants) (Thrower et al. 2004). In this study, allelic variation in all 3 candidate genes mapping to Omy5 (Clock1a, Myd118-1, and GHR1) were associated with development rate in outcrossed Sashin families; however, caution should be used as markers outside this genomic region were not typed to use as controls.
for population structure in the association analyses carried out here. Though associations between migratory strategy and genotype exist in the Sashin population on this chromosome in the region of the major QTL for development rate, more work is needed to evaluate the functional significance of these and other genes in this region on multiple traits.

A clear pattern of association between the rate of development and the decision to migrate does not emerge when...
evaluated across populations. That is, more rapidly developing embryos are not always associated with the migratory life-history type, and vice versa. For example, for the clonal lines used to generate the backcrosses in this study, the CW clonal line originated from a migratory steelhead trout and had a more rapid development rate, in a common environment, than the OSU clonal resident rainbow trout (Nichols et al. 2007; Miller et al. 2012). This contrasts with our results from the Sashin Creek and Sashin Lake populations, whereby the more rapidly developing fish were the progeny of the resident rainbow trout from Sashin Lake. Though the trends in association between migration and development rate are not clear, it is important to examine development rate as a potential adaptation to water temperature. The resident samples in this study were derived from Sashin Lake, which is colder during embryo development than the shallow stream environment where the migratory steelhead trout spawn (Thrower F., personal observation); the faster development rate in progeny from the resident crosses in the common garden experiment is likely a result of adaptation to colder water. Similarly, the OSU clonal line of rainbow trout, derived from a progenitor from Mt. Shasta, CA, originated from a warmer climate than the CW clonal line of rainbow trout from the Clearwater River, Idaho. Though Narum et al. (2008) found that fish in colder waters produced more residents than migrants, the association between life history and development rate is not consistent in these limited studies. The genetic correlation and possible genetic constraints to covariance in development rate and life-history diversity warrants further study.

It has long been debated that functionally important genetic variation for phenotypes can come in the form of either coding sequence variation or in variation that regulates gene expression (Hoekstra and Coyne 2007; Wray 2007). Though allelic variation in the candidate genes mapping to Omy5 is associated with variation in development rate, we also sought to examine whether those genes were differentially expressed between individuals with fast- versus slow-development rate. Fine-scale temporal expression analysis of candidate genes underlying the development rate QTL, and in some cases their duplicated genes, revealed that the 5 candidate genes were expressed early in embryogenesis, although the pattern of expression and upregulation varies among genes. However, differential expression of these genes underlying the major development rate QTL was not observed in backcross progeny with fast- versus slow-developing alleles. One set of paralogous genes, Myd118-1 (mapped to Omy5) and Myd118-2 (mapped to Omy10), showed very different patterns of expression from fertilization through hatching. The paralog mapping to the major development rate QTL on Omy5, Myd118-1, showed marked upregulation between 5 and 10 dpf compared with Myd118-2, which showed lower expression until hatching (about 23 dpf). Expression can change rapidly during embryonic development. These “spiky” patterns of expression together with a more coarse sampling during embryonic development may have obscured differences in expression in the QTL-selected backcross. A similar difficulty arises with the examination of expression in circadian genes such as Clock, which have a diurnal pattern of expression.

In summary, we have presented the mapping, expression, and association (or QTL) analysis of candidate genes in O. mykiss, demonstrating that allelic variation in candidate genes on Omy5 is associated with embryonic development rate in several populations. The candidate genes (and their paralogs examined) were expressed in O. mykiss embryos prior to hatching, but in this study, we were unable to find differential regulation of these genes in advanced backcrosses selected for this major genomic region of interest. During embryonic development, changes in gene expression occur rapidly, and identifying the genes and timing of expression responsible for divergence between fast- and slow-developing individuals will be difficult. As a completed, annotated genome becomes available, the identification of the total set of genes and regulatory regions in this region will provide an important resource for further candidate gene approaches. In future studies, the challenge remains not only to identify positional candidate genes but also to break down the close associations between genes in this region to identify the causal variant(s) promoting diversification in development rate, growth, and age at sexual maturity.

Supplementary Material
Supplementary material can be found at http://www.jhered.oxfordjournals.org/.

Funding
National Research Council, US Department of Agriculture - National Research Initiative grant (2007-35205-17881), National Science Foundation Career Award (NSF-DEB-0845265) (to K.M.N.); Howard Hughes Medical Institute undergraduate fellowships (to J.A.C. and M.H.).

Acknowledgments
We thank M. Ford and L. Park, Northwest Fisheries Science Center (NWFSC), National Marine Fisheries Service for in-kind support of the earliest work. J. Hard, also at the NWFSC, was very helpful with the line-cross analysis. We are grateful for the resources and personnel from the Little Port Walter Marine Station, National Marine Fisheries Service, Alaska Fisheries Science Center for work conducted in Sashin Creek and Sashin Lake, Alaska, and the staff there that facilitated this work. E. Leder provided primers for POMC gene fragments. Lastly, we are very grateful to 2 anonymous reviewers and S. Grant, whose constructive comments helped improve the manuscript.

References


Guyomard R, Mauger S, Tabet-Canale K, Martineau S, Genet C, Krieg F, Quillet E. 2006. A type I and type II microsatellite linkage map of rainbow trout (Oncorhynchus mykiss) with presumptive coverage of all chromosome arms. BMC Genom. 7:302.


Received July 9, 2013; First decision August 30, 2013; Accepted February 13, 2014

Corresponding Editor: Stewart Grant