Rapid, broad-scale gene expression evolution in experimentally harvested fish populations

Running title: Harvest-induced gene expression evolution

Silva Uusi-Heikkilä, Tiina Sävilammi, Erica Leder, Robert Arlinghaus and Craig R. Primmer

a Department of Biology, University of Turku, 20014 University of Turku, Finland

b Natural History Museum, University of Oslo, P.O. Box 1172 Blindern, 0318 Oslo, Norway

c Department of Biology and Ecology of Fishes, Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Müggelseedamm 310, 12587 Berlin, Germany

d Division of Integrative Fisheries Management, Department for Crop and Animal Sciences, Faculty of Life Sciences, Humboldt-Universität zu Berlin, Philippstrasse 13, Haus 7, 10115 Berlin, Germany

e Current address: Department of Biosciences, 00014 University of Helsinki, Finland

* To whom correspondence should be addressed: Silva Uusi-Heikkilä, Division of Genetics and Physiology, Department of Biology, University of Turku, 20014 University of Turku, Finland, tel. +358 44 32 38 174, fax: +358 2 333 5040, email: silva.uusi-heikkila@utu.fi.

Keywords: Gene expression | Fisheries-induced evolution | Transcriptome sequencing | Size selection | Adaptation | Captive breeding
ABSTRACT

Gene expression changes potentially play an important role in adaptive evolution under human-induced selection pressures but this has been challenging to demonstrate in natural populations. Fishing exhibits strong selection pressure against large body size, thus potentially inducing evolutionary changes in life-history and other traits that may be slowly reversible once fishing ceases. However, there is a lack of convincing examples regarding the speed and magnitude of fisheries-induced evolution and thus the relevant underlying molecular-level effects remain elusive. We use wild-origin zebrafish (*Danio rerio*) as a model for harvest-induced evolution. We experimentally demonstrate broad-scale gene expression changes induced by just five generations of size-selective harvesting, and limited genetic convergence following the cessation of harvesting. We also demonstrate significant allele frequency changes in genes that were differentially expressed after five generations of size-selective harvesting. We further show that nine generations of captive breeding induced substantial gene expression changes in control stocks likely due to inadvertent selection in the captive environment. The large extent and rapid pace of the gene expression changes caused by both harvest-induced selection and captive breeding emphasizes the need for evolutionary enlightened management towards sustainable fisheries.
Introduction

Gene expression changes play an important role in adaptive evolution and have been a topic of great interest in recent years (Manceau et al. 2011; Jones et al. 2012; Papakostas et al. 2014). A number of studies have documented that the role of gene expression changes in adaptation of phenotypic traits is influenced strongly by just one or few genes with large effect in a broad range of species (Shapiro et al. 2004; Miller et al. 2007; Manceau et al. 2011). However, the extent to which adaptations in polygenic traits are mediated by gene expression changes has been studied less intensively (Fraser et al. 2010; McKinney et al. 2015). Because gene expression can be regulated at many steps (from DNA to protein), and the regulatory networks governing gene expression are complex and modular, gene expression changes can be more labile than DNA sequence changes (Huminiecki & Wolfe 2004). Therefore, it has been suggested that gene expression changes may be particularly important for fostering (or constraining) genetic adaptation to strong selection pressures over short time scales (Jones et al. 2012; Papakostas et al. 2014).

Human activities, such as habitat fragmentation, harvesting and domestication, are affecting countless animal and plant species around the globe by changing the environment in which they occur and creating new adaptive landscapes (Palumbi 2001; Sullivan et al. 2017). For example, wildlife harvesting not only exceeds natural mortality rates (e.g., Jørgensen et al. 2007) but it is also selective in terms of sex, age, size and a range of other phenotypes (Allendorf & Hard 2009; Arlinghaus et al. 2017; Kuparinen & Festa-Bianchet 2017). A classic example of selective wildlife harvesting is fishing, where the largest individuals, cohort after cohort, are selectively removed from the population, often at very high rates. The resulting selection against large body size is predicted to favor fast life-histories characterized by the
evolution of earlier maturation at smaller size and higher energy investment into reproduction, which reduces post-maturation growth (Stearns 1992; Heino et al. 2015). Despite the accumulating empirical evidence from observational phenotypic data collected in the wild (Jørgensen et al. 2007), there has been a long controversy on the question of whether phenotypic changes in time series are truly evolutionary or merely reflect plastic life-history responses (Browman et al. 2008; Hilborn & Minte-Vera 2008) and what is the rate of fisheries-induced evolution (Andersen & Brand 2009). Furthermore, the impact of fisheries-induced selection on gene expression (how many genes and how fast expression changes) is unknown. This is a major shortcoming for understanding the molecular mechanisms involved in harvest-induced evolution.

Another question which has been rarely studied empirically, is whether or not populations can recover from harvest-induced evolutionary changes (e.g., following a fishing moratorium) and if so, at what pace? Size-selective harvesting can be a more intensive selection force than natural selection (e.g., predators, competitors; Darimont et al. 2009). Therefore, populations might not adapt to the cessation of size-selective harvesting as rapidly as they adapt to the fisheries selection (Conover et al. 2009; Enberg et al. 2009). Also, intensive selection together with high mortality (and genetic drift) might deplete population genetic diversity (Pinsky & Palumbi 2014; Marty et al. 2015). This can reduce the adaptive potential of a population and hamper its recovery (Allendorf et al. 2008). There are empirical examples of wild fish populations not fully recovering demographically after the cessation of fishing (Neubauer et al. 2013) potentially due to harvest-induced decline in genetic variation (Swain et al. 2007), however no studies have examined the recovery pattern at the molecular level. Understanding the extent and mechanisms of evolutionary
responses in exploited populations is crucial for the design of management tools aimed at sustainable exploitation of natural biological resources.

Many fisheries are experiencing over-exploitation and consequently temporary moratoria or no-take reserves are being established to help fish stocks to replenish. Since a moratorium can be considered as a rather extreme management practice to facilitate the recovery of an over-exploited fish stock, supplementary stocking (i.e., the release of hatchery-reared fish to supplement the natural population) is more commonly applied to restore and conserve wild populations, particularly in freshwater environments (Lorenzen et al. 2012). However, in many cases, whether or not hatchery-reared fish actually help population recovery has been questioned (Ryman & Laikre 1991; Araki et al. 2007). This is mainly because unintentional selection pressures brought about by artificial breeding and captive rearing can induce genetic and phenotypic changes that are maladaptive in the natural environment leading to high post-release mortality and low reproductive success (Araki et al. 2007; Lorenzen et al. 2012; Christie et al. 2014). While the fitness consequences of captive breeding on wild populations have been demonstrated, much less is known about the molecular basis of inadvertent selection in captivity (but see Christie et al. 2016). A better understanding of the broad-scale effects of captive breeding at the gene expression level, and identifying the genes upon which domestication selection acts could enable more appropriate recovery programs and management practices for exploited and endangered fish populations.

We studied changes in gene expression at the genome-wide level in experimental fish populations subjected to five generations of either size-selective (large-harvested fish) or random (random-harvested fish) harvesting followed by a six generation no-harvest period simulating a fishing moratorium. We sought to address three questions (Fig. 1): i) how
rapidly and to what extent do experimental fish populations respond to size-selective exploitation at the gene expression level? ii) how do the experimentally exploited fish populations respond to the cessation of harvesting? and iii) to what extent does unintentional selection induced by captive breeding result in a response at the gene expression level?

**Material and methods**

**Experimental design**

We used zebrafish (*Danio rerio*) caught from the wild (in West Bengal, India in 2006) as a model species to study transcriptome-wide gene expression changes in response to five generations of size-selective harvesting (Uusi-Hékkilä *et al.* 2015), followed by six generations of no harvesting, and in response to up to 11 generations of captive breeding. The founder fish were bred randomly for one generation to enable acclimation to captive conditions and reduce possible parental effects before the selection experiment started. Briefly, we subjected replicated (two replicates per harvest treatment, each consisted of 450 individuals) wild-origin zebrafish populations to selection for small body size (hereafter referred to as “large-harvested”, which is the treatment most similar to fisheries), while we harvested the control lines randomly with respect to body size (“random-harvested”). For the large-harvested fish we applied a 75% per-generation harvest rate (based on standard length) and removed the largest 75% of individuals (i.e., 338 individuals) from the population. This is consistent with intensive lethal capture fisheries (Lewin *et al.* 2006), thus the large-harvested fish represented a harvest scheme common in highly exploitive capture
fisheries. In the random-harvested line, the same proportion (75%) of individuals was harvested in each generation, but randomly with respect to body size. In both harvest treatments, the remaining 25% of females and males (i.e., 112 individuals) were used for reproduction (Uusi-Heikkilä et al. 2015). The harvesting regime was conducted for five generations (F1 – F5) and then halted for an additional six generations (F6 – F11), hereafter referred to as the “no-harvest period”. During the no-harvest period, the experimental populations consisted of 110-120 individuals. As fish were not harvested during this period, all fish were allowed to reproduce. Large- and random-harvested fish were reared in a common-garden environment, including similar feeding regimes and rearing densities across generations (Uusi-Heikkilä et al. 2015), thus the potential for different environmental conditions to induce differences in gene expression was minimized. See SI Materials and Methods: Zebrafish rearing conditions and harvesting for more details about fish rearing, feeding, harvesting and breeding.

RNA sampling

The liver transcriptomes were characterized for a total of 48 individuals which included four fish from each treatment replicate sampled in each of three generations: F2- (large- or random-harvested for two generations), F5- (large- or random-harvested for five generations) and F11-generations (six generations of no harvesting following five generations of large or random harvesting). Before the first liver sampling, fish had been reared and bred in the laboratory for three generations (and harvested for two generations). F2-generation samples were the earliest samples available from the selection experiment that were suitable for RNA analyses. Liver was chosen as a target organ for analyses because of its important role in various major physiological processes in fish, such as skeletal and soft
tissue growth, lipid, protein and carbohydrate metabolism, energy storage, maturation, reproduction and development (Devlin et al. 2006). Additionally, many hormones and hormone receptors involved in the regulation of growth and ovarian function (oocyte growth and ovulation), energy homeostasis and metabolism, immunological functions and various behaviors, such as aggression, feeding and foraging behavior are synthesized or expressed in the liver (Bjornsson et al. 2002). Differences in maturation status (i.e., whether fish are immature, maturing or mature) likely lead to differences in expression level of maturation-associated genes given the metabolic and hormonal changes in the liver during maturation (Ng et al. 1986; Soengas et al. 1995). The unusually highly expressed maturation-related genes could lead to significantly lower coverage of other genes and bias the expression estimates of these genes, thus substantially complicating the detection of differentially expressed genes of other biological processes. Therefore, the liver sampling schedule did not follow the age (days post fertilization; DPF) of the experimental fish but rather aimed for sampling the fish at similar developmental stage. The fish matured, and consequently were sampled, at different ages in each generation (F$_2$ at age 116 DPF; F$_5$ at age 69 DPF, and F$_{11}$ at age 97 DPF). However, there were no significant age differences between the random- and large-harvested fish within the F$_2$-, F$_5$- or F$_{11}$-generations (Table S1). A detailed description of the RNA sequencing can be found in SI Materials and Methods: Sampling, sequencing and processing the sequence data; RNASeq read alignment.

Gene expression quantification

The effects of size-selective harvesting were assessed by comparing the gene expression profiles of the large-harvested and random-harvested fish following five generations of harvesting (Fig. 1). As all treatment replicates had by then likely experienced similar
(inadvertent) selection pressures associated with captive rearing, the random-harvested treatment can be considered as a control with respect to size-selective harvesting. To assess the effect of six generations of no harvesting following five generations of harvesting, we compared the gene expression profiles of large- and random-harvested treatments at the F_{11}-generation (Fig. 1). We additionally studied the change in the gene expression profile of the large-harvested fish by comparing the changes in gene expression from the F_{2}- to F_{11}-generation to changes from the F_{2}- to F_{5}-generation (Fig. 1). By doing so our aim was to assess whether there were genes where the expression effects of size-selective harvesting remained following the six generation no-harvest period.

A second comparison aimed to assess the effects of captive breeding. The random-harvested treatment was not intentionally selected for any obvious phenotypic trait but it likely adapted to the laboratory environment during the 11 generations of laboratory rearing in addition to being affected by the effects of genetic drift. Therefore, comparing differences in gene expression profiles within the random-harvested treatment across generations (F_{2} vs F_{11}) can provide an insight into the magnitude of gene expression changes induced by captive breeding (Fig. 1). To obtain a more conservative estimate of the gene expression changes induced by captive breeding, changes in gene expression profiles across generations in large- and random-harvested fish were compared and genes where expression changed in both harvest treatments were identified (Fig. 1).

**Test for gene expression differences**

Read count files were analyzed in R using the limma-voom (v2.9.8) linear model as implemented in the R/Bioconductor package (Anders & Huber 2010). Read counts were normalized for the analysis using a “remove unwanted variation” (RUVg) normalization
strategy (Risso et al. 2014). The RUVg approach uses negative control genes that are assumed to be not differentially expressed among the treatments of interest (Risso et al. 2014). Therefore, to normalize expression levels, comparisons between harvest treatments in all generations were combined, and the 5,000 least differentially expressed genes were selected as the normalization gene set. Library sizes were normalized by using the limma-package.

To detect differentially expressed genes between the harvest treatments, treatment replicates and among generations, a design matrix including all explanatory variables (harvest treatment, generation, their interaction and treatment replicate) was created. A linear model was then fitted using the design matrix, and comparisons for main and interaction effects between harvest treatments and among generations were extracted. All the explanatory variables were treated as fixed effects. When identifying differentially expressed genes, the $P$-value (Benjamini-Hochberg adjusted for multiple testing) threshold was set to 0.05. In addition, differentially expressed genes were identified using the empirical false discovery rate (FDR). Differences in absolute fold change (i.e., non-directional) of differentially expressed genes between the harvest treatments and harvesting periods were tested with a Wilcoxon signed-rank test as the differences were not normally distributed. To visualize differences in overall gene expression patterns across generations, we conducted a principal component analysis (PCA) on all transcripts. Three most important principal components were analyzed separately for each generation using a linear model with the principal component as a response variable and harvest treatment, treatment replicate, individual standard length and wet mass as explanatory variables. In the linear model
$y \sim a + b + c + d$

$y$ was the principal component, $a$ was the harvest treatment, $b$ was the treatment replicate, $c$ was the standard length (mm), and $d$ was the wet mass (g). All explanatory variables were treated as fixed effects.

We used a hypergeometric test to study explicitly whether differentially expressed genes overlapped between certain comparisons (e.g., between the harvest treatments across generations) more than expected by chance. More information about the empirical FDR and testing for gene expression differences can be found in SI Materials and Methods: Test for gene expression differences.

Functional enrichment analysis was conducted using the Database for Annotation, Visualization and Integrated Discovery (DAVID v6.7) to identify gene ontology (GO) terms that were over-represented in the genes expressed differently between harvest treatments or between different generations of the same treatment (Huang et al. 2009). Although the knowledge of overrepresented GO terms is helpful, it may only partially resolve the molecular mechanisms relevant to the explored genes (Antonov et al. 2009). Therefore, the functional analysis was extended to study protein-protein interaction (PPI) networks. We submitted a gene list to Ingenuity Pathway Analysis (IPA) to study the most important interaction networks of the differentially expressed genes and identify their relevant upstream regulators. More details of identification of GO terms and gene networks can be found in SI Materials and Methods: Functional enrichment and gene network analysis.

Gene expression variance
The variance in expression level for each gene was estimated across the four individuals in each treatment replicate (inter-individual variance) by using RUVg-normalized and scaled read counts (Risso et al. 2014). The median of the gene-specific variances across all genes for each treatment replicate was then calculated. Gene expression variance within each generation and treatment replicate was bootstrapped 10,000 times with 1,000 random genes in each bootstrap. After that, the change in the estimated median gene expression variance from the F2- to F5-generation and from the F5- to F11-generation was calculated. We further tested whether the difference in gene expression variance was significant within the harvesting (from F2 to F5) and the no-harvest periods (from F5 to F11) between the harvest treatments using a paired t-test.

**SNP calling and annotation**

All individuals were used for single nucleotide polymorphism (SNP) identification and calling. Only SNPs with a minimum allele frequency of at least 0.1 were included. We used SAMtools (v.0.1.19) and the corresponding BCFtools to extract SNPs within a coverage range of 10 – 10,000, and filtered out SNPs within 10bp distance from indels. In addition to the default parameters in SAMtools, additional filtering steps were conducted: removal of potential indels, SNPs with an overall locus quality of lower than 999 (in SAMtools, a quality of 999 refers to a very good variant quality), and SNPs occurring outside chromosomes (mitochondrial and unassigned scaffold regions). SNPs where at least 43 (out of 48) individuals in each generation were successfully genotyped with at least 15X coverage (i.e., at least 43 individuals had at least 15 reads covering a valid SNP position) were retained. Following these filtering steps, 58,217 SNPs remained. To visualize differences in allele
frequency patterns across generations, we conducted a PCA on SNPs retained in the analysis and with all individuals having a valid genotype (21,181 SNPs; Fig. S1).

**Association between gene expression and SNP allele frequency changes**

To study the association between gene expression and SNP allele frequency changes in order to demonstrate a genetic basis for changes in gene expression levels, we first conducted an expression quantitative locus (eQTL) analysis. eQTLs are regions of the genome containing DNA sequence variants (e.g., SNPs) that influence the expression level of one or more genes. To identify SNPs that are located close to genes and then link variation in expression values to SNP genotypes, we used the MatrixEQTL R package (Shabalin 2012).

We performed linear regression of RUVg-normalized expression values and SNPs with generation and treatment replicate as covariates (included to account for population stratification). SNPs with more than four genotypes missing were filtered out and gene expression values were quantile normalized to fit a standardized normal distribution to avoid the overdispersion effect resulting from outlier individuals (resulting in excess of false positives). Although eQTLs for a certain gene can be found in any chromosome (trans-eQTLs), we focused on local (cis) eQTLs which map to the approximate location (i.e., < 1M bp) of the gene.

We further compared the allele frequency changes (from F2- to F5-generation) between the large- and random-harvested fish 1) across all 58,217 SNPs, 2) in 14,500 gene-associated SNPs (one association per SNP extracted by the MatrixEQTL based on the lowest P-value) in differentially expressed genes and in genes that were not differentially expressed (detected at the F5-generation), and finally 3) in SNPs assigned as eQTLs and not assigned as eQTLs. The observed average allele frequency changes were compared to a permuted allele
frequency change distribution. The permuted distribution consisted of 10,000 mean values, which were sampled from a dataset where we pooled the allele frequency changes of 1) all SNPs in random- and large-harvested fish, 2) gene-associated SNPs occurring in differentially expressed genes and in genes that were not differentially expressed separately for both harvest treatments, and 3) all SNPs assigned and not assigned as an eQTL separately for both harvest treatments.

Estimating the contribution of genetic drift

To determine the relative contribution of neutral (drift) and adaptive (selection) processes during gene expression evolution, the variation in gene expression level explained by the replicates (drift) and the harvest treatment (selection) were estimated using MANOVA. For each generation, the amount of variation in gene expression explained by the two models was compared: one with harvest treatment and treatment replicate as explanatory variables and one with only treatment replicate as an explanatory variable. The input data for the MANOVA analysis were the three first principal components, which explained most of the variation in expression count data between the harvest treatments and generations.

All analyses were conducted in R (version 3.1.3; R Core Team 2016) packages lme4, Base Stats, limma, RUVnormalize, and MatrixEQT.

Results

Broad scale differences in gene expression and reduction of gene expression variance after five generations of size-selective harvesting

Expression differences were detected between the large- and random-harvested treatments in 509 transcripts (2.80% of all expressed genes) at the F2-generation (Fig. 2A). Three
generations later, the number of differentially expressed genes had increased by more than ninefold, with expression differences between large- and random-harvested fish in 4,310 genes (23.7%; Fig. 2A). The absolute fold change of these differentially expressed genes was significantly higher among large-harvested (1.06 ± 0.02; median ± 95% confidence interval) than random-harvested fish (0.92 ± 0.01; W=1,0146,000, P < 0.001; Fig. S2A). Fold change was also significantly higher during the harvesting period (F₂ – F₅) than during the no-harvest period (0.80 ± 0.01; F₅ – F₁₁) among the large-harvested fish (W=1,0567,000, P < 0.001; Fig. S2B). An individual-level PCA revealed that the gene expression differences were explained by both selection and drift, as variation in principal component 1 (PC1) was explained by harvest treatment (F = 22.5, P < 0.001), in principal component 2 (PC2) by harvest treatment (F = 80.1, P < 0.001) and treatment replicate (F = 25.2, P < 0.001), and in principal component 3 (PC3) by harvest treatment (F = 13.3, P = 0.004). A PCA figure visualizes that the gene expression profiles of the large-harvested and random-harvested individuals were not markedly diverged at the F₂-generation (Fig. 3A). However, after an additional three generations of size-selective harvesting (F₅-generation), the gene expression profiles of large- and random-harvested fish had clearly diverged and formed distinct clusters (Fig. 3B).

The differentially expressed genes between large- and random-harvested individuals at the F₅-generation contributed to various major biological processes, such as translation and transcription (P < 0.001), lipid and steroid biosynthesis (P = 0.003) and energy metabolism (P = 0.034; Table S2). The central genes of the three most significant gene networks were \textit{ELAVL1}, \textit{EBP} and \textit{MSMO1} (Fig. S3A-C; Table S3A). In addition, five putative upstream regulators were implicated (Table S3B).
Reductions in gene expression variance between individuals within harvest treatments were evident in both size-selective (large-harvested) and non-selective (random-harvested) harvesting regimes during the harvesting period (Fig. 4A; Fig. S4). The reduction in variance was particularly clear in one of the replicates of both harvest treatments while in the other replicate the variance reduction was less substantial (Fig. 4A).

The average allele frequency change during the harvesting period in more than 58,000 SNPs was significantly ($P < 0.001$) higher among large-harvested fish (0.140) compared to random-harvested fish (0.125; Fig. 5A). The average allele frequency change in gene-associated SNPs was generally higher in differentially expressed genes than in genes that were not differentially expressed ($P < 0.001$; Fig. 5B). Although this was the case among both harvest treatments, the average SNP allele frequency change was higher in large-harvested than in random-harvested fish in both differentially expressed genes (large-harvested 0.174; random-harvested 0.152) and in genes that were not differentially expressed (large-harvested 0.159; random-harvested 0.128; Fig. 5B). Overall we identified 221 eQTLs in 140 genes. eQTLs were present in 24 differentially expressed genes (out of 4,300). The average allele frequency changes of SNPs not assigned as eQTLs were well within the permuted distributions in both harvest treatments (Fig. 5C) but unlike in random-harvested fish ($P = 0.721$), the average allele frequency change of eQTL SNPs in large-harvested fish (0.177) was significantly higher than in SNPs not assigned as eQTLs (0.140; $P < 0.001$; Fig. 5C).

Gene expression response to a no-harvest period

After six harvest-free generations that followed five generations of size-selective or non-selective harvesting, differences in gene expression between large- and random-harvested
fish had eroded to some extent, but there still remained a large number of differentially expressed genes (3,171 genes, 17.4%; Fig. 2A). Variation in PC 1 was weakly explained by harvest treatment ($F = 3.62, P = 0.078$) and treatment replicate ($F = 3.77, P = 0.074$). Variation in PC 2 was explained by none of the variables, and in PC 3 by harvest treatment ($F = 45.6, P < 0.001$), individual standard length ($F = 14.7, P = 0.003$) and wet mass ($F = 11.9, P = 0.005$). A consistent gene expression difference between the transcriptome profiles of harvest treatments in the F$_{11}$-generation was not evident in the two first principal components (Fig. 3C), but it was evident in the third principal component (Fig. 3F). It appears that the effect of harvest treatment on gene expression differences was weaker and that the random processes played a more important role following the no-harvest period (F$_{11}$) compared to following the harvesting period (F$_{5}$).

Many of the genes differing between the harvest treatments were enriched for biological processes related to RNA processing and metabolism ($P < 0.001$), protein catabolism ($P = 0.004$), ribosome biogenesis ($P = 0.012$) and nitrogen compound metabolism ($P = 0.016$; Table S4). Central genes of the key networks were $HNF1A$, $ESR1$ and $MDM2$ (Fig. S5A-C; Table S3C). Also several significant upstream regulators were implicated (Table S3D). Out of the more than 3,000 differentially expressed genes between the harvest treatments in the F$_{11}$-generation only around a quarter (826 genes) were in common with genes differentially expressed between the large- and random-harvested lines at the F$_{5}$-generation ($P < 0.001$; Fig. 2A).

A more conservative approach to study the effect of no-harvest period on a size-selectively exploited population is to compare the gene expression profiles of the large-harvested fish across generations (i.e., from F$_{2}$ to F$_{11}$). In that comparison, we identified 3,007 differentially
expressed genes (Fig. 2B) out of which 1,159 (38.5%) were the same genes that were affected by size-selection during the harvesting period ($P < 0.001$). These over 1,000 genes were largely different compared to the differentially expressed genes shared between the harvesting and no-harvest period in the random-harvested fish (75 genes in common; $P = 0.006$). The genes that were differentially expressed in large-harvested fish during the harvesting period and remained differentially expressed after the no-harvest period were enriched for biological processes associated with protein transport and localization ($P = 0.015$), and for genes in the insulin signaling pathway ($P = 0.045$; Table S5).

Expression variance during the no-harvest period ($F_5 – F_{11}$) had differing patterns of change for the size-selected and non-selected harvest treatments compared to patterns during the harvesting period ($F_2 – F_5$). Interestingly, gene expression variance decreased by 25% in large-harvested fish after the cessation of harvesting, but increased by 34% in random-harvested fish (Fig. 4B; Fig. S4).

**Genetic drift**

In the $F_5$-generation, the MANOVA model with the harvest treatment and the treatment replicate (full model) as explanatory variables explained substantially more of the variation (54%) in the three principal components as opposed to a model with only a treatment replicate as an explanatory variable (reduced model; 28%; Table S6), suggesting that differences in gene expression were better explained by the effects of both selection and genetic drift rather than by genetic drift alone.

**The effects of captive rearing on gene expression**
To identify gene expression changes potentially induced by captive rearing, we compared the expression patterns of the random-harvested fish at the F\textsubscript{2}- and F\textsubscript{11}-generations, and identified 4,978 differentially expressed genes (27.4%). These genes contributed to biological processes such as DNA metabolic processes ($P = 0.001$), oxidative phosphorylation ($P = 0.029$) and protein catabolism ($P = 0.029$; Table S7). The central genes of the three most significant networks were \textit{CHAF1B}, \textit{TRIP13} and \textit{CYB5R2} (Fig. S6A-C; Table S3E). More than 20 upstream regulators were implicated (Table S3F). Out of the almost 5,000 differentially expressed genes between the F\textsubscript{2}- and F\textsubscript{11}-generations in random-harvested fish, 924 were in common with those differentially expressed between the F\textsubscript{2}- and F\textsubscript{11}-generations in large-harvested fish ($P < 0.001$; Fig. 2B). Approximately 3\% of these genes were associated with insulin signaling pathway ($P = 0.006$).

\textbf{Discussion}

Our experimental approach to understand the molecular consequences of harvest-induced evolution revealed that broad-scale changes in gene expression and SNP allele frequencies arose after just five generations of size-selective harvesting. This suggests that size-selective harvesting can rapidly induce large gene expression changes in exploited populations, and the extent of such changes cannot be explained by genetic drift alone. A no-harvest period resulted in the expression divergence of a largely different set of genes indicating that although selection and drift also occur in the absence of size-selective harvesting, the targets of these processes are different and ‘recovery’ may be unlikely to occur, even under a fishing moratorium. Exploited populations, particularly in freshwaters, are often supplemented with hatchery-reared individuals. Therefore, we also studied the rate and
magnitude of gene expression changes induced by captive breeding. We demonstrated substantial gene expression changes in random-harvested fish suggesting that although not selected for any obvious phenotypic trait, captive breeding and rearing alone might induce large unintentional genetic changes in fish populations.

Gene expression differences after five generations of size-selective harvesting

After five generations of size-selective harvesting, the number of differentially expressed genes between large- and random-harvested fish encompassed approximately 24% of all expressed genes investigated, although relatively few differences (<3%) were present early in the selection experiment at the F₂-generation. During the harvesting period, the magnitude of change in differentially expressed genes was also higher among large-harvested than random-harvested fish (Fig. S2A). Such broad-scale gene expression differences are unusual compared to many earlier studies investigating gene expression differences under domestication selection. For example, a comparison of brain gene expression levels between domesticated vs wild dogs, pigs, or rabbits revealed expression differences in less than 1% of all expressed genes (Roberge et al. 2006; Albert et al. 2012).

Similarly, differences in gene expression between wild and farmed Atlantic salmon (Salmo salar) subjected to four to seven generations of selection aimed at increasing growth rate revealed a difference in less than 2% of genes (Roberge et al. 2006; Debes et al. 2012). One potentially important difference between our study and the domestication studies is the fact that those studies focused on gene expression changes in the brain, rather than the liver, and this could have limited the number of genes differing significantly in their expression (Jobling et al. 2014). A notable exception is a recent study on fish domestication that demonstrated large expression differences similar to what we found between wild and
hatchery-reared steelhead trout (*Oncorhynchus mykiss*) after only one generation of hatchery rearing (Christie *et al.* 2016). The authors identified 723 differentially expressed genes (4.6%) between wild and first-generation hatchery offspring reared in common environment.

In addition to domestication studies, other artificial selection experiments have documented divergence in gene expression among differently selected lines. For instance, foxes (*Vulpes vulpes*) selected for aggressive or tame behavior in a long-term breeding experiment differed in their gene expression in 335 genes (Kukekova *et al.* 2011), and the fraction of genes that diverged under directional selection on body weight in chicken (*Gallus domesticus*) was shown to be on average 13% (Resnyk *et al.* 2015). Selection for improved residual feed intake resulted in gene expression differences in less than 4% of all the expressed genes in beef cattle (*Bos taurus*; Weber *et al.* 2016) and in pigs (*Sus scrofa*; Vincent *et al.* 2015), and in 41 genes in chickens with divergent feed efficiency (Yi *et al.* 2015). While aggression, body weight and feeding behavior are also polygenic traits similar to body length on which selection was operating on in our study, the larger gene expression differences we observed due to size-selective harvesting may be partly related to the differences in tissues sampled (e.g., brain, muscle or intestine vs liver), differences in gene expression analyses and therefore in sensitivity to detect differentially expressed genes (microarray vs RNASeq) or to the fact that many of these studies were conducted on domesticated animals which may have already lost genetic diversity. Nevertheless, our results demonstrate that size-selective harvesting (and not drift alone) can rapidly induce gene expression changes on a broader scale than reported in many of the previous domestication studies or selection experiments conducted on vertebrates. Further research
is required to identify the specific mechanisms causing the gene expression differences.

However, given the common-garden environment, which minimized environmental
differences between the harvest treatments, along with substantially higher average allele
frequency changes in differentially expressed genes compared to genes that were not
differentially expressed in both harvest treatments (Fig. 5B) and in eQTL SNPs in large-
harvested fish compared to SNPs not assigned as eQTLs (Fig. 5C), it is likely that at least
some of the expression differences we report have a genetic basis.

Earlier research investigating the effects of harvesting on life-history traits in the same
selection lines as used in the current study showed that while individuals from large- and
random-harvested treatments had similar pre-maturation growth rates, large-harvested
individuals reached a significantly lower adult body size than random-harvested individuals
(Uusi-Heikkilä et al. 2015), exhibited a lower condition factor (Fig. S7) and a lower age-
specific maturation probability (Uusi-Heikkilä et al. 2015). Although the individuals analyzed
in the present study were immature and there were no differences in body size between the
large- and random-harvested fish at the time of sampling (Table S1), gene expression
profiles can nevertheless potentially differ at earlier developmental phases between
zebrafish lines selected for differing growth patterns (Amaral & Johnston 2012), for instance
when fish begin to allocate energy to maturation and post-maturation growth. Indeed, the
differentially expressed genes between large- and random-harvested fish were enriched in
biological processes that are related to energy allocation and growth, albeit at a very
general level. Thus, these differences could relate to differences in adult body size which
have previously been found to evolve in response to size-selective mortality (Uusi-Heikkilä
et al. 2015).
Protein turnover is a fundamental biological process related to somatic growth (Reeds 1988). We detected a significant proportion of the differentially expressed transcripts between large- and random-harvested fish to be enriched for processes associated with protein synthesis and metabolism, such as transcription and translation. Another important process differing significantly between the harvest treatments was oxidative phosphorylation, which is the metabolic pathway forming ATP that stores and supplies energy, for instance for protein and lipid metabolism (Alberts et al. 2002). Harvest treatments also differed in expression of genes associated with steroid synthesis. Steroid hormones are associated with many major biological functions in fish, including feeding behavior, aggression, stress and oocyte maturation (Wingfield et al. 1998; Nagahama & Yamashita 2008). Steroids are also known to promote growth in fish by enhancing anabolic processes (Matty 1986). Central genes in key gene networks were associated with hypoplasia, growth failure and biosynthesis of cholesterol, which serves as a precursor for the biosynthesis of steroid hormones. Furthermore, genes that were differentially expressed and up-regulated during the harvesting period in large-harvested fish (Fig. S8) were tightly associated with maturation (Table S8). It could be speculated that differences between harvest treatments found earlier in maximum body size, exploration behavior (Uusi-Heikkilä et al. 2015) and condition factor between the harvest treatments are consistent with the observed differences in expression of genes that can be broadly associated with feeding, circadian rhythms, somatic growth and maturation. However, it is good to keep in mind that while our data represent interesting signals and gene ontology enrichments, follow-up studies are required to establish more direct links between the gene expression differences observed here, and the phenotypic differences observed previously (Uusi-Heikkilä et al. 2015).
Gene expression differences following a no-harvest period

The number of differentially expressed genes between the harvest treatments decreased during the no-harvest period from 24% to 17% but the differences in gene expression profiles were still substantial after this period and expression differences remained in genes that are linked to growth-related processes at the physiological and behavioral level.

Differentially expressed genes detected between the harvest treatments after the no-harvest period were enriched in biological processes that are important in energy release and cell growth. Among the most highly expressed regulators, hypocretin receptor 2 (HCRTR2) is known to be associated with appetite and feeding behavior (De Lecea et al. 1998).

The genes that were differentially expressed between the harvest treatments at the F_{11}-generation were generally different to those that differed at the F_5-generation: only a quarter (26%) of the differentially expressed genes between the large-harvested and random-harvested stocks were identical between the harvesting and the no-harvest period. Thus, the proportion of genes that were affected by size-selective harvesting had decreased while a new set of genes, likely affected by another selection pressure and/or drift, were identified. The random-harvested fish were under selection for captive rearing during the entire experiment while the large-harvested fish experienced an intensive size-selection pressure for five generations (in addition to selection for captive rearing) and after that, selection solely for captive rearing that likely began to favor very different characteristics than size-selection (e.g., large, fecund and aggressive individuals; Roberge et al. 2006; Devlin et al. 2009). This might suggest that the large-harvested fish were set on a different evolutionary trajectory. In addition, the magnitude of the expression of differentially
expressed genes was significantly lower during the no-harvest period than during the harvesting period among the large-harvested fish (Fig. S2B). Indeed, the lack, or only very slow recovery, of phenotypes after a phase of intensive harvesting has been repeatedly demonstrated in individual-based eco-genetic models (Enberg et al. 2009; Dunlop et al. 2015; Marty et al. 2015).

Even after six generations of no harvesting (i.e., at F_{11}), significant differences in expression levels remained in one third of the genes observed to have differential expression patterns after the five generation harvesting period. These approximately 1,000 genes were enriched for biological functions such as protein transport and localization, and insulin signaling pathway. One could argue that these genes might have been under selection to captive rearing but this seems unlikely because out of these, only 75 genes were in common with the genes that were differentially expressed among random-harvested fish in the equivalent generation-level comparison. Although the differentially expressed genes were mostly different between the harvesting and no-harvest period, a small group of reproduction- and maturation-related genes were up-regulated among large-harvested fish during the harvesting period and down-regulated during the no-harvest period (Fig. S8). Large-harvested fish have been shown to invest a relatively high amount of energy into reproduction (Uusi-Heikkilä et al. 2015), which could at least partly explain the up-regulation of these genes during the harvesting period. Selection pressure for reproductive investment likely relaxed after the harvesting was halted. Admittedly, age could have confounded the between-generation comparisons as the fish were sampled in each generation at different age (in days). However, the expression of maturation-related genes might not have been confounded by the sampling design as fish were always sampled at the same developmental
stage (i.e., they were all immature). Although expecting the populations to reverse back to
the early-harvest state might not be entirely realistic, at least not in the current laboratory
setting where the fish were under selection for captive rearing, both above mentioned
approaches (i.e., comparison between harvest treatments and generations) generally lead
to the same conclusion: a component of the effects of size-selective harvesting still
remained despite six generations of no harvesting.

The effect of harvesting and a no-harvest period on gene expression variance

During the harvesting period, gene expression variance decreased in both harvest
treatments. This could be due to reduced population sizes but it is possible that also
random-harvested fish were under selection pressure due to high harvest rate or intrinsic
fecundity selection (Uusi-Heikkilä et al. 2015). However, the reduction in gene expression
variance was not entirely consistent within the harvest treatments (i.e., one of the
treatment replicate in both harvest treatments showed less reduced variance) possibly due
to the limited number of biological replicates.

We show not only that gene expression variance was reduced after five generations of size-
selective and non-selective harvesting but it continued to decrease among large-harvested
fish during the no-harvest period which mimicked a harvest moratorium. However, among
random-harvested fish gene expression variance increased during the no-harvest period.
This might suggest that the combined effects of selection and drift were stronger in the
large-harvested treatment and/or the response to captive rearing was relaxed in the
random-harvested treatment, but not in the large-harvested line after the cessation of
harvesting. Thus, it is possible that large- and random-harvested fish responded differently
to captive rearing. Further, it has been suggested that adaptive (life-history) evolution can
be very rapid during early phases of selection but it may also cease rapidly (Reznick et al. 1997). It is plausible that the large-harvested fish experienced two types of selection: first size-selection favoring small body size and then selection to captive breeding potentially favoring large body size (Heath et al. 2003; Devlin et al. 2009). These two opposite selection forces could have decreased gene expression variance in large-harvested fish compared to the random-harvested fish. Finally, based on the expression (Fig. 3A-C) and SNP PCA figures (Fig. S1), it appears that during the early selection, the individuals from the different harvest treatments were already genetically different but their gene expression profiles were not markedly different. Thus, at the F2-generation they seemed to exhibit strong canalization which broke down at the F5-generation potentially due to selection, and gene expression remained decanalized at the F11-generation. Decanalization could uncover hidden gene expression variance (Chen et al. 2015). If decanalization was stronger among random-harvested fish, this could have led to increased gene expression variance compared to the large-harvested fish. However, this remains speculative as quantifying decanalization among harvest treatments was beyond the scope of this study.

Gene expression variation represents a source of variability that can improve fitness in varying environments and under stressful conditions or varying selection pressures (Whitehead & Crawford 2006; Papakostas et al. 2014). Although sometimes considered as costly noise (Wang & Zhang 2011), inter-individual gene expression variation has been shown to contribute substantially to physiological performance among individuals, thus it can be biologically relevant (Li et al. 2010). Therefore, loss of gene expression variation could be detrimental for exploited fish populations because it can reduce adaptive potential and hamper their recovery.
Gene expression response to captive rearing

Today, many declining wild populations, especially freshwater fishes, are supplemented with captive-reared individuals and this can have many detrimental ecological and genetic effects on existing populations (Laikre et al. 2010; Lorenzen et al. 2012). The genetic and ecological concerns include direct genetic effects caused by introgression or hybridization, genetic changes in hatchery stocks brought about by selection and drift, and lowered survival and reproductive success of hatchery-reared individuals (Araki et al. 2007; Christie et al. 2014). Introgression of genetic material from captive-reared fish that are maladapted in the wild may indeed cause negative fitness effects in the wild populations, alter the gene pools of local stocks and negatively affect population productivity (Chilcote et al. 2011). Although it is known that hatchery fish may have lower relative reproductive success (Araki et al. 2007; Christie et al. 2014) and survival rate (Lorenzen et al. 2012) than their wild counterparts, the systematic functional genetic effects of captive rearing remain unclear (but see Christie et al. 2016).

We showed that nine generations of rearing and breeding in captivity affected the expression of a large number of genes. Over 27% of all expressed genes investigated were differentially expressed between the $F_2$- and $F_{11}$-generations in random-harvested fish. However, as the fish at the $F_2$-, $F_5$- and $F_{11}$-generations were sampled at different ages due to the fact that age at maturity was evolving during the experiment one cannot rule out the possibility that age differences could have also contributed to between generation comparisons within harvest treatments. Most earlier studies that have shown gene expression changes induced by captive rearing have focused on a relatively small set of candidate genes (Roberge et al. 2006; Debes et al. 2012), and few studies thus far have
examined this at a genome-wide level (Christie et al. 2016). Our results demonstrate that captive rearing affected biological processes that can be broadly related to growth, such as energy production, protein catabolism, and fatty acid metabolism, despite random-harvested fish were not directly selected for growth or large (or small) body size. Our findings of the functional genetic effects of captive rearing are broadly in line with salmon domestication studies (Roberge et al. 2006; Tymchuk et al. 2009; Debes et al. 2012; Christie et al. 2016), although many of the earlier studies compared wild fish and fish under intense growth selection and therefore cannot be directly compared to our results.

The key gene networks of the differentially expressed genes were associated with RNA repair, post-transcriptional modification and cholesterol biosynthesis (Table S3E). Some of the significant upstream regulators are known to be involved in reproduction through mediating estrogen and progesterone production, steroidogenesis and cell proliferation (Table S3F). Thus, adaptation to captive environment may have not only affected growth-related but also maturation-related processes. In fact, we showed that maturation- and reproduction-related genes were up-regulated among the random-harvested fish (from F5 to F11), and we have shown earlier that the random-harvested fish have a higher age-specific maturation probability than large-harvested fish (Uusi-Heikkilä et al. 2015). This is in agreement with other studies showing that fish held in captivity tend to mature later than their wild conspecifics (Debes & Hutchings 2014). Thus, it is possible that in captivity, traits under selection are related to body condition (body fat content), growth, and potentially maturation. Although we do not assess fitness consequences directly in this study, earlier research has demonstrated the negative fitness consequences for wild populations of captivity-induced modification of reproductive traits (reviewed in Christie et al. 2014).
A second approach for identifying genes potentially affected by captive rearing was to compare the gene expression profile divergence of large-harvested fish between the F$_2$- and F$_{11}$-generations to that of random-harvested fish between the same generations. The comparison of gene expression profiles of F$_2$- and F$_{11}$-generations in large-harvested fish revealed 3,000 differentially expressed genes and in random-harvested fish almost 5,000 genes (Fig. 2B). Out of these two sets of differentially expressed genes, only 924 genes were in common between the two harvest treatments. These roughly 900 genes potentially include those predominantly affected by selection for captive rearing, although it is possible that some of the remaining almost 4,000 genes are as well, but have been differently affected by the alternative selection regimes. A significant proportion of the 900 differentially expressed genes in common between the harvesting treatments were enriched within insulin signaling pathway. Insulin is tightly associated with growth, thus indirectly with gonad development. The insulin (and insulin-like growth factor) signaling pathway has also been shown to play a major role in the control of longevity in vertebrates (van Heemst 2010) and in the balance of the functioning of the immune system. Traits mediated by the insulin signaling pathway, such as growth, maturation and disease resistance, are known to be important traits for artificial breeding programs (Gjedrem & Thodesen 2005).

These results add an important functional genetic element to the genetic concerns of stocking: the divergent gene expression profiles of captive-reared individuals reported here and elsewhere (Christie et al. 2016) likely contribute to the reduced fitness of captive-reared individuals in the wild (Araki et al. 2007), as well as captive-wild hybrids (Christie et al. 2014), which has been shown to reduce population productivity in salmonids (Chilcote et al.
Therefore, it can be questioned whether hatchery-reared fish should be used intensively for enhancing wild populations that are still naturally reproducing as hybridization might have far-reaching consequences in terms of adaptive capacity and genomic integrity of wild populations without necessarily increasing fisheries yield (Lorenzen et al. 2012).

Conclusions

We have shown that both size-selective harvesting and captive rearing can induce rapid and substantial changes in gene expression profiles in experimentally exploited fish populations. Gene expression profiles did not fully converge after the cessation of harvesting. Our results thus suggest that the evolutionary response to size-selective harvesting can be broad, rapid and potentially difficult to reverse. This can be undesirable for fishing because phenotypes (and genotypes) favored most by natural selection and by fishers (e.g., the large fish) are removed. It can also be harmful from the evolutionary perspective because of reduced gene expression variance, thus potentially reduced adaptive potential of the exploited populations. Reduction in gene expression variance could thus be a factor potentially contributing to the lack of genetic (and phenotypic) recovery of size-selectively exploited fish populations in the wild. Hence, our results reinforce the recommendation of applying evolutionary principles to management and promote management that maintains large and diverse breeding populations to foster the full range of phenotypes and genotypes that natural selection can act upon (Schindler et al. 2010). Inadvertent selection due to captive rearing alone also resulted in large changes in gene expression profiles, suggesting that the use of hatchery-reared fish for supplementing wild populations might affect the adaptive capacity and genomic integrity of wild populations.
Acknowledgements

We thank Karena Kuntze, Asja Vogt, Marcus Ebert, Sylvia Werner, Sarah Becker, Yvonne Klaar, Theresa Arlt, Julie Menard, and David Lewis for fish husbandry, care taking, and data collection; Sanna Pausio for assistance in RNA extraction; and Henrik Zwadlo for technical assistance. We also thank reviewers for the excellent feedback. This research used computing resources of the CSC - IT Center for Science, Espoo, Finland. Funding for this study was received through the AXA Research Fund, ICES, and Kone Foundation to SUH, Adaptfish and BTypes grants by the Leibniz Community to RA, and Academy of Finland to CRP. Finally, we thank three reviewers for excellent suggestions and constructive feedback.

References


Chen J, Nolte V, Schlötterer C (2015) Temperature stress mediates decanalization and

Chilcote MW, Goodson KW, Falcy MR (2011) Reduced recruitment performance in natural
populations of anadromous salmonids associated with hatchery-reared fish. *Canadian
Journal of Fisheries and Aquatic Sciences*, **68**, 511-522.


Christie MR, Marine ML, Fox SE, French RA, Blouin MS (2016) A single generation of
domestication heritably alters the expression of hundreds of genes. *Nature
Communications*, doi: 10.1038/ncomms10676.

Conover DO, Munch SB, Arnott SA (2009) Reversal of evolutionary downsizing caused by

Darimont CT, Carlson SM, Kinnison MT, Paquet PC, Reimchen TE *et al.* (2009) Human
predators outpace other agents of trait change in the wild. *Proceedings of the National
Academy of Sciences*, **106**, 952-954.

Debes PV, Normandeau E, Fraser DJ, Bernatchez L, Hutchings JA (2012) Differences in
transcription levels among wild, domesticated, and hybrid Atlantic salmon (*Salmo salar*)


Kukekova A, Johnson JL, Teiling C, Li L, Oskina IN et al. (2011) Sequence comparison of prefrontal cortical brain transcriptome from a tame and an aggressive silver fox (*Vulpes vulpes*). *BMC Genomics, 12*, 482.


of gene networks for residual feed intake in Angus cattle using genomic prediction and RNA-


**Data Accessibility.** The raw data from each library are available at the NCBI Sequence Read Archive (SRA) under the accession number SRP105243.

**Figure legends**

**Fig. 1.** A conceptual figure of the experimental design and the research questions. The effects of size-selective harvesting on gene expression were studied by comparing the expression patterns of large- and random-harvested fish after five generations of harvesting (red line). The effects of a no-harvest period on gene expression were studied in two ways,
first by comparing the expression patterns of large- and random-harvested fish after six
generations of no harvesting (dashed orange line) and second, by comparing the overall
expression changes (from $F_2$ to $F_{11}$: thick solid orange line) to expression changes during the
harvesting period (from $F_2$ to $F_5$: thin solid orange line) among the large-harvested fish. The
effects of captive breeding on gene expression were studied by comparing expression
patterns of the $F_2$- and $F_{11}$-generations among the random-harvested fish and comparing
the changes in both harvest treatments during the whole experimental period (from $F_2$ to
$F_{11}$; grey lines).

**Fig. 2.** Number of differentially expressed genes between the harvest treatments and
between generations within the harvest treatments. (A) Number of differentially expressed
genes between large- and random-harvested fish in each generation (bold) and genes in
common between generations (italics). (B) Number of differentially expressed genes
between generations within a harvest treatment (bold) and genes that are shared between
harvest-treatments for the same generation comparison (italics).

**Fig. 3.** A principal component analysis (PCA) based on transcript abundances of all 18,192
expressed genes. PC1 and PC2 contrasted in (A) the $F_2$-generation, (B) the $F_5$-generation,
and (C) the $F_{11}$-generation. PC2 and PC3 contrasted in (D) the $F_2$-generation, (E) the $F_5$-
generation, and (F) the $F_{11}$-generation. Red diamonds represent large-harvested and grey
circles random-harvested treatments. Different shades depict individuals from the two
replicates within both harvest treatments.
Fig. 4. Change in gene expression variance during the (A) harvesting (from $F_2$ to $F_5$) and (B) no-harvest period (from $F_5$ to $F_{11}$) in both harvest-treatment replicates. There was no significant difference in gene expression variance between the harvest treatments during the harvesting period ($t = 0.486, P = 0.627$). During the no-harvest period the difference in variance between the harvest treatments was significant ($t = -618.1, P < 0.001$). Error bars show the 95% confidence intervals. Different shades of red represent the two replicates within large-harvested fish and grey within random-harvested fish.

Fig. 5. The distribution of permuted SNP allele frequency changes compared to those observed (marked by arrows) in large- and random-harvested fish (from $F_2$- to $F_5$-generation). (A) Across all SNPs, (B) across gene-associated SNPs in differentially expressed genes (solid arrows) and in genes that are not differentially expressed, referred to as “other genes” (dashed arrows), and (C) across SNPs assigned as eQTLs (solid arrows) and SNPs not assigned as eQTLs (dashed arrows). Red bars and arrows represent large-harvested and grey bars and arrows random-harvested fish.
Generation

Large-harvested

$F_2$ $F_5$ $F_{11}$

Random-harvested
### A

<table>
<thead>
<tr>
<th>Harvest Treatment</th>
<th>F&lt;sub&gt;2&lt;/sub&gt;</th>
<th>F&lt;sub&gt;5&lt;/sub&gt;</th>
<th>F&lt;sub&gt;11&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between generations</td>
<td>509</td>
<td>142</td>
<td>4,310</td>
</tr>
<tr>
<td></td>
<td>826</td>
<td>3,171</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>111</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### B

#### Between generations

<table>
<thead>
<tr>
<th>Harvest Treatment</th>
<th>F&lt;sub&gt;2&lt;/sub&gt; vs F&lt;sub&gt;5&lt;/sub&gt;</th>
<th>F&lt;sub&gt;5&lt;/sub&gt; vs F&lt;sub&gt;11&lt;/sub&gt;</th>
<th>F&lt;sub&gt;2&lt;/sub&gt; vs F&lt;sub&gt;11&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large-harvested</td>
<td>3,112 414</td>
<td>2,564 801</td>
<td>3,007 924</td>
</tr>
<tr>
<td>Random-harvested</td>
<td>2,083</td>
<td>5,764</td>
<td>4,978</td>
</tr>
</tbody>
</table>