Title: Contrasting genomic shifts underlie parallel phenotypic evolution in response to fishing

Authors: Nina O. Therkildsen[1*], Aryn P. Wilder[1†], David O. Conover[2], Stephan B. Munch[3], Hannes Baumann[4], and Stephen R. Palumbi[5]

Affiliations:

[1]Department of Natural Resources, Cornell University, Ithaca, NY, USA

[2]Department of Biology, University of Oregon, Eugene, OR, USA

[3]Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA, Santa Cruz, CA, USA

[4]Department of Marine Sciences, University of Connecticut, Groton, CT, USA

[5]Department of Biology, Stanford University, Hopkins Marine Station, Pacific Grove, CA, USA.

*Correspondence to: Nina Overgaard Therkildsen (nt246@cornell.edu)

[†]Current address: San Diego Zoo Institute for Conservation Research, Escondido, CA, USA

Abstract: Humans cause widespread evolutionary change in nature, but we still know little about the genomic basis for rapid adaptation in the Anthropocene. We tracked genomic changes across all protein-coding genes in experimental fish populations that evolved striking shifts in growth rates due to size-selective harvest over only four generations. Comparisons of replicate

lines show parallel allele frequency shifts that recapitulate responses to size-selection gradients in the wild across hundreds of unlinked variants concentrated in growth-related genes. Yet, a super-cluster of genes also rose rapidly in frequency and dominated the evolutionary dynamic in one replicate line, but not in others. Parallel phenotypic changes thus masked highly divergent genomic responses to selection, illustrating how contingent rapid adaptation can be in the face of strong human-induced selection.

One Sentence Summary: Fish harvest drive rapid evolution in growth rates through inconsistent genomic shifts.

Main Text: Human actions cause rapid evolutionary change in many species (*1, 2*), but the underlying genomic basis remains poorly understood. Prime examples of human-driven evolution come from fisheries, where the selection pressure imposed by intense harvest has caused pronounced shifts in growth rates and reproductive timing in many stocks, potentially reducing yields and impeding recovery from overfishing (*3, 4*). Fishing has been shown to change gene expression, genetic diversity, and allele frequencies at candidate markers (*5-8*), but the overall magnitude and extent of genomic change and the repeatability of response patterns remain unclear, hampering our ability to predict the consequences for fish stock resilience to continually changing fishing regulations and new challenges imposed by climate and other environmental change.

Theory and empirical studies suggest that rapid adaptation can occur through two broadly different paths. In classic sweep scenarios, phenotypic shifts are caused by large allele frequency changes at one or a few loci (*9, 10*), which may quickly erode the genetic variation needed for reversal or

other adaptive responses. In contrast, complex quantitative traits have traditionally been assumed to evolve through small allele frequency shifts across many loci (*11*), which should better retain functionally important variation and to a lesser extent compromise the ability of populations to adapt to future conditions. Both types of response patterns have been observed across experimental evolution studies in small mammals, insects, plants, and microbes (*12, 13*). Yet, despite the important implications for the future evolutionary potential of species, we know little about how predictable genomic responses to pervasive human-induced selection are in large populations that harbor high levels of standing genetic variation, which can serve as raw material to fuel rapid adaptation.

We have returned to an influential experiment that demonstrated rapid evolution in response to size-selective fishing (*14*). The experiment subjected six populations each of ~1,000 adult Atlantic silversides (*Menidia menidia*, a small estuarine fish) to 90% harvest every generation. In two replicate populations, the individuals left to reproduce were the smallest 10% (in body length, hereafter the "down-selected" lines). In two other populations, the largest 10% were left to reproduce each generation (hereafter the "up-selected" lines), and another two populations were controls (the 90% mortality was random with respect to size). After only five generations, fish from the up-selected lines were on average 25% longer than fish from the down-selected lines (Fig. 1), resulting in an almost two-fold difference in average weight (*14*).

We used low-coverage (~1.3x per individual) whole genome sequencing of frozen fish from the experiment to examine the genomic basis underlying these phenotypic shifts. We sequenced 75 individuals from the source pool used to establish the experiment (offspring from hundreds of wild-caught parents, generation 0) and 48-50 individuals from each of the six populations in generation 5 (Fig. S1, Table S1). The species is distributed along the east coast of North America,

where a shorter growing season has driven evolution of faster growth in northern regions (*15*, *16*). For comparison, we also sequenced 42-50 individuals from each of four wild populations across this natural cline (Table S2). In the absence of a reference genome, we mapped the genomic reads to a comprehensive reference transcriptome (details in (*17*)) to examine exome-wide patterns of change across 2.36 million single nucleotide polymorphisms (SNPs) (*18*).

As expected, we observed a clear reduction in genetic diversity in all captive populations. However, the four populations subjected to size-selection (two up-selected and two down-selected) all suffered a significantly greater loss of diversity (23-27% loss of polymorphic sites and 7-9% loss of nucleotide diversity) than either of the two control lines (17-20% of polymorphic sites and 5% of nucleotide diversity, Table S1, $P \leq 0.028$, one-tailed t-test). Using a likelihood ratio test that accounts for genotype uncertainty given our low-coverage data, we also observed more significant allele frequency shifts in the selected lines (529-19,258 SNPs), compared to the control lines (nine SNPs in each; Table S1), despite identical sequencing effort. Thus, the size-selection treatments consistently led to accelerated rates of genomic change.

Higher levels of relatedness (Fig. S2) and fewer unique mitochondrial haplotypes in the selected lines (Table S1) suggest that this acceleration was partially caused by a selection-induced reduction in effective population sizes, which in turn increased genetic drift genome-wide. However, drift will cause random fluctuations in allele frequencies, whereas we expect selection to change allele frequencies at genes affecting the targeted phenotype in the same direction under identical harvest regimes and in opposite directions under opposing harvest regimes. We used the statistic diffStat (*19*) to quantify parallel divergence as the minimum allele frequency difference in the same direction between up- and down-selected replicates. Simulations under a neutral scenario showed on average 4,952 SNPs with diffStat >0.3, whereas our data show 10,523 (Figs. 2A, S3),

suggesting that size-selection caused consistent patterns of parallel divergence across more SNPs than expected by chance.

SNPs with diffStat values in the top 1% (greatest parallel allele frequency divergence between selection regimes, n=23,648) cumulatively show a clear correlation with individual body length across populations ($P<10^{-6}$, linear mixed-effects model, Fig. 2B), supporting their involvement in a polygenic response to selection. A subset of the same top 1% diffStat SNPs also exhibits highly elevated levels of differentiation among the four wild populations (1,596 with geographic $F_{ST}>0.25$ and 357 with $F_{ST}>0.5$ against an exome-wide median F_{ST} of 0.02). For the majority (83.2%) of the 357 SNPs that show the greatest geographic structure, the allele that is most common in the north (where short summers select for fast growth) became more frequent in the up-selected (fast-growing) experimental lines. Conversely, the northern allele decreased in frequency in the down-selected (slow-growing) lines for 89.9% of these SNPs (Fig. 2C). The pronounced bias in directionality of change suggests a shared genomic basis for growth rate divergence in the experiment and in the wild, implying that the rapid response to selection tapped into a reservoir of standing genetic diversity across hundreds of unlinked (Fig. S4) variants maintained by long-term clines in natural selection. Functional enrichment analysis revealed significant overrepresentation (FDR-corrected $P<0.05$) of these variants in 11 biological process categories, including "biomineral tissue development", "notochord morphogenesis", "bone mineralization", and "regulation of heart rate", all possibly linked to growth (Table S3).

However, our initial focus on parallelism revealed only part of the response to selection. Of SNPs that changed the most across generations (in the top 1% of temporal F_{ST} (Gen0 to Gen5) for each population) only 0.9% are the same between the two down-selected replicates, and only 2.3% between the two up-selected (Fig. S5). Alignment to the medaka (*Oryzias latipes*) genome further

indicates that SNPs with significant allele frequency changes are concentrated on different sets of chromosomes in the four different size-selected lines, indicating highly idiosyncratic responses (Fig. 3).

The most extreme change occurred in a block of 9,348 SNPs on chromosome 24 that shifted from an initial frequency of <0.05 to ~0.6 in generation 5 in only one of the two down-selected populations ("Down2", Fig. 3F). Almost all of these SNPs, held together in strong linkage disequilibrium (Fig. S6), are either fixed or nearly fixed for opposite alleles between wild silverside populations in the north and south of the distribution range (Fig. 4A), indicating that this haplotype block has been under strong divergent selection across the latitudinal growth rate cline in the wild. The region showed no recombination over the course of the experiment (Fig. 4B, S7B), yet may span much of the chromosome if synteny is conserved between medaka and the Atlantic silverside (Fig. S7B). It stretches across 415 genes that show overrepresentation of four biological process categories relating to muscle contraction and calcium sequestration (Table S4). Furthermore, these genes have 7% more non-synonymous variants than exome-wide proportions (significant enrichment, χ^2=49.49, df=2, P<10^{-6}), suggesting long-standing impacts of natural selection between chromosomal variants.

Consistent with phenotypic patterns, it was the southern (naturally slow-growing) linked haplotype that increased in Down2, and it correlated negatively with individual body length (linear mixed-effects model P=0.038, Fig. 4C), explaining 7.9% of the phenotypic variation in Down2. Its frequency also reverted back towards initial levels after size-selection had been relaxed in a five-generation continuation of the experiment (Fig. S7), mirroring a phenotypic reversal to faster growth rates in the down-selected lines (*20*).

The much greater allele frequency shift than expected under drift alone (Fig. S8) followed by this reversal and the association with both growth divergence patterns in the wild and individual length in the experiment indicate that chromosome 24 played an important role in the response to selection in the Down2 population. Yet, we do not observe elevated allele frequency change in this genomic region in any of the other selected lines, including Down1 (Fig. 3). This may be because the experiment was established with fish from the middle of the species range, where the southern haplotype on chromosome 24 is very rare (only one copy out of 150 sampled haplotypes in Gen0, Fig. 4). Thus, the southern haplotype may have been absent or lost early in the experiment in most populations, leaving little opportunity for selection to act on it. Nonetheless, the very similar phenotypic responses to size-selection across other replicates that did not have this haplotype available (Fig. 1) indicate that alternate genomic pathways must have been involved.

Similarly, in Up1 change is concentrated on chromosomes 1, 9 and 15 (Fig. 3A), coinciding with elevated linkage disequilibrium in these regions and strong drops in genetic diversity, indicating selective sweeps (Fig. S9). Over 39% of SNPs in the top 1% of temporal F_{ST} (between generation 0-5) and over 66% of SNPs in the top 0.1% map to these three chromosomes. The SNPs in the top 0.1% also cumulatively show a significant correlation with individual body length (linear mixed-effect model, P=0.016, Fig. S9C) and are enriched for non-synonymous SNPs (χ^2=13.26, df=2, P=0.0013, Fig. S10), supporting their functional role. The other up-selected replicate (Up2) shows small shifts on chromosomes 1 and 9, but notably not on 15 (Fig. 3B).

Taken together, our results thus show three major patterns. First, size-selective fishing caused greater loss of genetic diversity compared to size-independent fishing. Findings of reduced genetic diversity in wild, overharvested fish populations (*6*) have generally been interpreted as

consequences of population bottlenecks. Our results suggest that fisheries-induced selection may also have contributed.

Second, we see parallel allele frequency shifts among selection lines in hundreds of unlinked genes associated with growth variation in the wild. Such repeatable polygenic responses across many loci follow classic quantitative genetic predictions about the effects of selection on complex traits based on pre-existing genetic variation (*11*) and suggest that natural variation maintained across environmental mosaics facilitates rapid responses to anthropogenic selection.

Third, in contrast, we also observe idiosyncratic signatures of strong, but highly non-parallel, selection on large blocks of tightly linked genes within some replicates. Linked gene complexes play an important role for local adaptation in other marine species (*21*, *22*), and multiple well-known cases of rapid adaptation have been attributed to large allele frequency shifts at one or a few key loci (*23-25*). This contrasts with systems where adaptation is mediated exclusively by small shifts across many loci (*26*, *27*). Here, we see a combination, with similar phenotypic responses to selection (Fig. 1) underpinned by parallel polygenic shifts across all populations, but large changes in linked genetic variation in only some of them.

The juxtaposition of these different modalities of genome evolution and the distinct evolutionary trajectories in each of our replicates illustrate how observations at the phenotypic level may provide impressions of parallelism even when part of the underlying genomic shifts are, in fact, highly divergent. While this pattern is sometimes found among wild populations (*28*), it counters prior reports of very similar genomic responses to replicated experimental evolution from standing variation in outbred sexual organisms (*13, 29*).

Although the selection intensity used here was somewhat higher than in most fisheries and exact responses to selection always are influenced by the particular experimental design, our results clearly indicate that the genomic underpinnings of fisheries-induced evolution, which have been invisible to us until now, are not predictable from phenotypic patterns alone. They also show that fishing may rapidly cause genomic changes that are comparable to what exists across geographic populations in nature. These findings likely hold true for many other species that, like the Atlantic silverside, harbor a diverse reservoir of adaptive standing genetic variation (*30*), enabling multiple pathways to the same rapid evolutionary response. We now for the first time have the capability to reveal and monitor such responses at the genome level, allowing fisheries and wildlife managers to more comprehensively assess human impacts, and improving our understanding of the speed, consequences and reversibility of complex adaptations as we continue to sculpt the evolutionary trajectories of the species around us.

References and Notes:

1. S. R. Palumbi, Humans as the world's greatest evolutionary force. *Science* **293**, 1786–1790 (2001).

2. A. P. Hendry, K. M. Gotanda, E. I. Svensson, Human influences on evolution, and the ecological and societal consequences. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **372**, 20160028 (2016).

3. F. W. Allendorf, J. J. Hard, Human-induced evolution caused by unnatural selection through harvest of wild animals. *Proc. Natl. Acad. Sci. U.S.A.* **106 Suppl 1**, 9987–9994

(2009).

4. M. Heino, B. Diaz Pauli, U. Dieckmann, Fisheries-induced evolution. *Annu. Rev. Ecol. Evol. Syst.* **46**, 461–480 (2015).

5. S. J. van Wijk, M. I. Taylor, S. Creer, C. Dreyer, F. M. Rodrigues, I. W. Ramnarine, C. V. Oosterhout, G. R. Carvalho, Experimental harvesting of fish populations drives genetically based shifts in body size and maturation. *Front. Ecol. Environ.* **11**, 181–187 (2013).

6. M. L. Pinsky, S. R. Palumbi, Meta-analysis reveals lower genetic diversity in overfished populations. *Mol. Ecol.* **23**, 29–39 (2014).

7. S. Uusi-Heikkila, A. R. Whiteley, A. R. Whiteley, A. Kuparinen, S. Matsumura, P. A. Venturelli, C. Wolter, J. Slate, C. R. Primmer, T. Meinelt, S. S. Killen, D. Bierbach, G. Polverino, A. Ludwig, R. Arlinghaus, The evolutionary legacy of size-selective harvesting extends from genes to populations. *Evol. Appl.* **8**, 597–620 (2015).

8. S. Uusi-Heikkilä, T. Sävilammi, E. Leder, R. Arlinghaus, C. R. Primmer, Rapid, broad-scale gene expression evolution in experimentally harvested fish populations. *Mol. Ecol.* **26**, 3954–3967 (2017).

9. J. Maynard-Smith, J. Haigh, The hitch-hiking effect of a favourable gene. *Genet. Res.* **23**, 23–35 (1974).

10. J. Hermisson, P. S. Pennings, Soft sweeps and beyond: understanding the patterns and probabilities of selection footprints under rapid adaptation. *Methods Ecol. Evol.* **8**, 700–

716 (2017).

11. N. H. Barton, P. D. Keightley, Understanding quantitative genetic variation. *Nat. Rev. Genet.* **3**, 11–21 (2002).

12. S. F. Bailey, T. Bataillon, Can the experimental evolution programme help us elucidate the genetic basis of adaptation in nature? *Mol. Ecol*. **25**, 203–218 (2016).

13. A. Long, G. Liti, A. Luptak, O. Tenaillon, Elucidating the molecular architecture of adaptation via evolve and resequence experiments. *Nat. Rev. Genet.* **16**, 567–582 (2015).

14. D. O. Conover, S. B. Munch, Sustaining fisheries yields over evolutionary time scales. *Science* **297**, 94–96 (2002).

15. D. O. Conover, T. M. C. Present, Countergradient variation in growth rate: compensation for length of the growing season among Atlantic silversides from different latitudes. *Oecologia* **83**, 316–324 (1990).

16. D. O. Conover, S. A. Arnott, M. R. Walsh, S. B. Munch, Darwinian fishery science: lessons from the Atlantic silverside (*Menidia menidia*). *Can. J. Fish. Aquat. Sci.* **62**, 730–737 (2005).

17. N. O. Therkildsen, S. R. Palumbi, Practical low-coverage genomewide sequencing of hundreds of individually barcoded samples for population and evolutionary genomics in nonmodel species. *Mol. Ecol. Resour.* **17**, 194–208 (2017).

18. Materials and methods are available as supplementary materials at the Science website.

19. T. L. Turner, A. D. Stewart, A. T. Fields, W. R. Rice, A. M. Tarone, Population-based

resequencing of experimentally evolved populations reveals the genetic basis of body size variation in *Drosophila melanogaster*. *PLoS Genet*. **7**, e1001336 (2011).

20. D. O. Conover, S. B. Munch, S. A. Arnott, Reversal of evolutionary downsizing caused by selective harvest of large fish. *Proc. R. Soc. Lond. Ser. B*. **276**, 2015–2020 (2009).

21. J. Hemmer-Hansen, E. E. Nielsen, N. O. Therkildsen, M. I. Taylor, R. Ogden, A. J. Geffen, D. Bekkevold, S. J. Helyar, C. Pampoulie, T. Johansen, FishPopTrace Consortium, G. R. Carvalho, A genomic island linked to ecotype divergence in Atlantic cod. *Mol. Ecol.* **22**, 2653–2667 (2013).

22. A. M. Westram, M. Rafajlović, P. Chaube, R. Faria, T. Larsson, M. Panova, M. Ravinet, A. Blomberg, B. Mehlig, K. Johannesson, R. Butlin, Clines on the seashore: The genomic architecture underlying rapid divergence in the face of gene flow. *Evolution Letters* **2**, 297–309 (2018).

23. R. H. ffrench-Constant, P. J. Daborn, G. Le Goff, The genetics and genomics of insecticide resistance. *Trends Genet.* **20**, 163–170 (2004).

24. N. M. Reid, D. A. Proestou, B. W. Clark, W. C. Warren, J. K. Colbourne, J. R. Shaw, S. I. Karchner, M. E. Hahn, D. Nacci, M. F. Oleksiak, D. L. Crawford, A. Whitehead, The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. *Science* **354**, 1305–1308 (2016).

25. A. E. V. Hof, P. Campagne, D. J. Rigden, C. J. Yung, J. Lingley, M. A. Quail, N. Hall, A. C. Darby, I. J. Saccheri, The industrial melanism mutation in British peppered moths is a transposable element. *Nature* **534**, 102–105 (2016).

26. S. P. Egan, G. J. Ragland, L. Assour, T. H. Q. Powell, G. R. Hood, S. Emrich, P. Nosil, J. L. Feder, Experimental evidence of genome-wide impact of ecological selection during early stages of speciation-with-gene-flow. *Ecol. Letters* **18**, 817–825 (2015).

27. M. Laporte, S. A. Pavey, C. Rougeux, F. Pierron, M. Lauzent, H. Budzinski, P. Labadie, E. Geneste, P. Couture, M. Baudrimont, L. Bernatchez, RAD sequencing reveals within-generation polygenic selection in response to anthropogenic organic and metal contamination in North Atlantic Eels. *Mol. Ecol.* **25**, 219–237 (2016).

28. D. I. Bolnick, R. D. H. Barrett, K. B. Oke, D. J. Rennison, Y. E. Stuart, (Non)parallel evolution. *Annu. Rev. Ecol. Evol. Syst.* **49**, 303–330 (2018).

29. J. L. Graves, K. L. Hertweck, M. A. Phillips, M. V. Han, L. G. Cabral, T. T. Barter, L. F. Greer, M. K. Burke, L. D. Mueller, M. R. Rose, Genomics of parallel experimental evolution in *Drosophila*. *Mol Biol Evol*. **34**, 831–842 (2017).

30. L. Bernatchez, On the maintenance of genetic variation and adaptation to environmental change: considerations from population genomics in fishes. *J. Fish Biol.* ***89***, 2519–2556 (2016).

31. L. A. Hice, T. A. Duffy, S. B. Munch, D. O. Conover, Spatial scale and divergent patterns of variation in adapted traits in the ocean. *Ecol. Letters*. **15**, 568–575 (2012).

32. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Meth.* **9**, 357–359 (2012).

33. T. Korneliussen, A. Albrechtsen, R. Nielsen, ANGSD: Analysis of next generation

sequencing data. *BMC Bioinformatics* **15**, 356 (2014).

34. H. Li, A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).

35. R. Nielsen, J. S. Paul, A. Albrechtsen, Y. S. Song, Genotype and SNP calling from next-generation sequencing data. *Nat. Rev. Genet.* **12**, 443–451 (2011).

36. C. A. Buerkle, Z. Gompert, Population genomics based on low coverage sequencing: how low should we go? *Mol. Ecol.* **22**, 3028–3035 (2013).

37. M. Fumagalli, Assessing the effect of sequencing depth and sample size in population genetics inferences. *PLoS ONE* **8**, e79667 (2013).

38. A. Smit, R. Hubley, P. Green, Smit, AFA, Hubley, R & Green, P. *RepeatMasker Open-4.0*, (available at http://www.repeatmasker.org).

39. D. H. E. Setiamarga, M. Miya, Y. Yamanoue, K. Mabuchi, T. P. Satoh, J. G. Inoue, M. Nishida, Interrelationships of Atherinomorpha (medakas, flyingfishes, killifishes, silversides, and their relatives): The first evidence based on whole mitogenome sequences. *Mol. Phylogenetics Evol.* **49**, 598–605 (2008).

40. T. J. Near, R. I. Eytan, A. Dornburg, K. L. Kuhn, J. A. Moore, M. P. Davis, Wainwright, P.C., M. Friedman, W. L. Smith, Resolution of ray-finned fish phylogeny and timing of diversification. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 13698–13703 (2012).

41. D. Campanella, L. C. Hughes, P. J. Unmack, D. D. Bloom, K. R. Piller, G. Orti, Multi-

locus fossil-calibrated phylogeny of Atheriniformes (Teleostei, Ovalentaria). *Mol. Phylogenetics Evol.* **86**, 8–23 (2015).

42. A. Amores, J. Catchen, I. Nanda, W. Warren, R. Walter, M. Schartl, J. H. Postlethwait, A RAD-tag genetic map for the platyfish (*Xiphophorus maculatus*) reveals mechanisms of karyotype evolution among teleost fish. *Genetics* **197**, 625–641 (2014).

43. B. E. Warkentine, C. L. Smith, J. W. Rachlin, A reevaluation of the karyotype of the Atlantic silverside, *Menidia menidia*. *Copeia* **1987**, 222 (1987).

44. F. Cunningham, M. R. Amode, D. Barrell, K. Beal, K. Billis, S. Brent, D. Carvalho-Silva, P. Clapham, G. Coates, S. Fitzgerald, L. Gil, C. G. Girón, L. Gordon, T. Hourlier, S. E. Hunt, S. H. Janacek, N. Johnson, T. Juettemann, A. K. Kähäri, S. Keenan, F. J. Martin, T. Maurel, W. McLaren, D. N. Murphy, R. Nag, B. Overduin, A. Parker, M. Patricio, E. Perry, M. Pignatelli, H. S. Riat, D. Sheppard, K. Taylor, A. Thormann, A. Vullo, S. P. Wilder, A. Zadissa, B. L. Aken, E. Birney, J. Harrow, R. Kinsella, M. Muffato, M. Ruffier, S. M. J. Searle, G. Spudich, S. J. Trevanion, A. Yates, D. R. Zerbino, P. Flicek, Ensembl 2015. *Nucleic Acids Res*. **43**, D662–9 (2015).

45. C. Camacho, G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, T. L. Madden, BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).

46. E. T. Domyan, Z. Kronenberg, C. R. Infante, A. I. Vickrey, S. A. Stringham, R. Bruders, M. W. Guernsey, S. Park, J. Payne, R. B. Beckstead, G. Kardon, D. B. Menke, M Yandell, M. D. Shapiro, Molecular shifts in limb identity underlie development of feathered feet in two domestic avian species. *Elife* **5**:e12115 (2016).

47. P. E. Jorde, N. Ryman, Unbiased estimator for genetic drift and effective population size. *Genetics* **177**, 927–935 (2007).

48. T. Maruki, M. Lynch, Genotype-frequency estimation from high-throughput sequencing data. *Genetics* **201**, 473–486 (2015).

49. M. S. Ackerman, P. Johri, K. Spitze, S. Xu, T. G. Doak, K. Young, M. Lynch, Estimating seven coefficients of pairwise relatedness using population-genomic data. *Genetics* **206**, 105–118 (2017).

50. F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J. D. Thompson, D. G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).

51. C. Aktas, haplotypes: Haplotype inference and statistical analysis of genetic variation. R package version 1.0, (available at http://CRAN.R-project.org/package=haplotypes).

52. P. W. Messer, SLiM: simulating evolution with selection and linkage. *Genetics* **194**, 1037–1039 (2013).

53. B. C. Haller, P. W. Messer, SLiM 2: Flexible, interactive forward genetic simulations. *Mol Biol Evol*. **34**, 230–240 (2017).

54. A. Conesa, S. Götz, J. M. García-Gómez, J. Terol, M. Talón, M. Robles, Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674–3676 (2005).

55. P. Jones, D. Binns, H.-Y. Chang, M. Fraser, W. Li, C. McAnulla, H. McWilliam, J. Maslen, A. Mitchell, G. Nuka, S. Pesseat, A. F. Quinn, A. Sangrador-Vegas, M. Scheremetjew, S.-Y. Yong, R. Lopez, S. Hunter, InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240 (2014).

56. S. Myhre, H. Tveit, T. Mollestad, A. Laegreid, Additional gene ontology structure for improved biological reasoning. *Bioinformatics* **22**, 2020–2027 (2006).

57. B. J. Haas, A. Papanicolaou, M. Yassour, M. Grabherr, P. D. Blood, J. Bowden, M. B. Couger, D. Eccles, B. Li, M. Lieber, M. D. Macmanes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman, T. William, C. N. Dewey, R. Henschel, R. D. Leduc, N. Friedman, A. Regev, De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512 (2013).

58. S. Tang, A. Lomsadze, M. Borodovsky, Identification of protein coding regions in RNA transcripts. *Nucleic Acids Res*. **43**, e78–e78 (2015).

59. P. Cingolani, A. Platts, Le Lily Wang, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, D. M. Ruden, A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly* **6**, 80–92 (2012).

60. J. Gillis, M. Mistry, P. Pavlidis, Gene function analysis in complex data sets using ErmineJ. *Nat. Protoc.* **5**, 1148–1159 (2010).

61. J. P. Sinnwell, T. M. Therneau, D. J. Schaid, The kinship2 R package for pedigree data. *HHE* **78**, 91–93 (2014).

62. K. Bartoń, Multi-Model Inference (MuMIn). R package v1.42.1, (available at https://cran.r-project.org/web/packages/MuMIn/).

63. T. Maruki, M. Lynch, Genome-wide estimation of linkage disequilibrium from population-level high-throughput sequencing data. *Genetics* **197**, 1303–1313 (2014).

64. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, 1000 Genome Project Data Processing Subgroup, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

65. J. H. Shin, S. Blay, B. McNeney, J. G., LDheatmap: an R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. *J Stat Soft*. **16**, Code Snippet 3 (2006).

Acknowledgments: We want to thank Sergey Kryazhimskiy for advice on library preparation, Devon Pearce for facilitating sample processing, Beth Sheets for assistance in the laboratory, Robin Waples for helpful discussions, and the Therkildsen lab at Cornell University along with the Editor and four anonymous reviewers for valuable comments that helped improve the manuscript. **Funding:** This work was funded by National Science Foundation (US) grants to SRP (OCE-1434325) and NOT (OCE-1756316), a Villum Foundation (DK) postdoctoral fellowship to NOT, and Cornell University. **Author contributions:** NOT and SRP conceived of and designed the study, NOT generated the data, NOT and APW analyzed the data with input from SRP, SBM and DOC conducted the original experiment, HB recovered the samples and participated in the study design, NOT, APW and SRP wrote the manuscript with input from all

authors. All authors approved the manuscript before submission. **Competing interests:** None declared. **Data and materials availability:** The genomic sequence data are archived in the NCBI Short Read Archive under Bioproject ID PRJNA376564.

Supplementary Materials:

Materials and Methods

Figures S1-S10

Tables S1-S5

References (*31-65*)

Figure legends:

Fig 1. Observed shifts in adult size. Trends across generations in mean length at harvest (standardized as the difference from the mean of the control populations within each generation) +/- the standard deviation in up-selected (blue shades), down-selected (yellow/orange), and control populations (green shades).

Fig. 2. Parallel genomic shifts among replicate treatments. A) The number of SNPs with diffStat>0.3 (red arrow) was much higher than observed in any of 1,000 simulated datasets (grey histogram). B) Individual standardized length at harvest was negatively correlated with the number

of putative slow-growing alleles found in each fish across the top 1% diffStat SNPs. Regression lines show the slope within each population. C) The distribution of change in frequency of the northern (putatively fast-growing) allele in each population across the 357 top-diffStat SNPs also differentiated (F_{ST}>0.5) in the wild.

Fig. 3. Genomic distribution of allele frequency changes within each population. Log_{10}-transformed FDR-corrected *P*-values for allele frequency change between generation 0 and 5 in the six experimental populations (A-F) at each SNP ordered along the 24 medaka chromosomes.

Fig. 4. Expansive linkage disequilibrium block on chromosome 24 shows fixed differences along the geographic cline and correlation to body length. A) Histograms of the minor allele frequency distributions in four wild Atlantic silverside populations for the 9,348 chromosome 24 SNPs that changed significantly in the Down2 population. B) The most likely genotype at the subset of these SNPs (columns) with geographic F_{ST}>0.9 (n=5,447) for each individual (rows) in generation 0 (top heatmap) and in Down2 generation 5 (bottom heatmap). The single heterozygous individual in generation 0 is highlighted with an arrow. C) Length distributions among individuals carrying either 2, 1, or zero copies of the southern chromosome 24 haplotype in the Down2 population in generation 5.